The Pennsylvania State University

The Graduate School

The Mary Jean and Frank P. Smeal College of Business Administration

ORDERING DECISIONS AND COORDINATION IN SUPPLY CHAINS: A BEHAVIORAL PERSPECTIVE

A Thesis in

Business Administration

by

Yan Wu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December, 2006

The thesis of Yan Wu was reviewed and approved* by the following:

Elena Katok
Associate Professor of Supply Chain and Information Systems
Thesis Adviser
Chair of Committee

Gary Bolton
Professor of Business Economics

Daniel Guide
Associate Professor of Operations and Supply Chain Management

A. Ravi Ravindran
Professor of Industrial and Manufacturing Engineering

John E. Tyworth
Professor of Supply Chain Management
Head of the Department of Supply Chain and Information Systems

*Signatures are on file in the Graduate School.

ABSTRACT

This dissertation investigates the impact of various behavioral factors on supply chain management, using a laboratory approach. A series of experiments is designed to challenge the assumptions of several basic theoretical models used in supply chain management that have ignored behavioral dynamics. We use controlled experimental settings to confront decision makers (mostly college students) with decision tasks conformed to the assumptions being tested. Our human subjects are motivated by real financial incentives based on their performance in the simulated game. The dissertation includes three interrelated essays. Chapters I and II provide an empirical test of the supply chain contracting theory. This theory assumes that supply chain members behave in a way that maximizes their own expected profits. However, our experimental results in Chapter I show that retailers have decision biases and often deviate from the optimal newsvendor solutions, whereas suppliers tend to behave risk-aversely rather than risk-neutrally. In addition, the results in Chapter II suggest that channel members have preferences over both their pecuniary payoff and their relative payoff standing. In other words, they are not purely self-interested but rather fairness-concerned. In Chapter III, we investigate mechanisms to reduce the *bullwhip behavior*, using the well-known *Beer Distribution Game*. And we find that when participants have obtained system-wide training experience and are allowed to communicate with their channel partners, the variability of orders in supply chains can be greatly decreased. This dissertation contributes to bridging the gap between two traditionally divergent fields of behavioral theory and operations management by identifying behavioral factors that have significant impact on the predictions of operational models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the enthusiastic supervision of Dr. Elena Katok during this work. I also thank Dr. Gary Bolton and Dr. Daniel Guide for their generous help and insightful suggestions.

I would like to thank the National Science Foundation (Doctoral Dissertation Grant # SES-0412900), Smeal College of Business, Penn State University (Doctoral Dissertation Research Grant) and the Center for Supply Chain Research (CSCR) for funding this dissertation. I am also grateful to the technical support from the Laboratory of Economics Management and Auctions (LEMA) at Penn State.

Finally, I am deeply indebted to my parents Zhengguang Wu and Yanqing Ma, and my husband Fei Wu for their understanding, endless patience and encouragement when it was most required.

INTRODUCTION

Over the past few years there has been growing interest in the incorporation of behavioral theory and dynamics into discussions of operations management (OM). Such interest has been demonstrated by an increase in the calls for publications devoted to this interface, from both individual researchers and leading journals in operations management.

For example, Hopp (2004), in his review of the last 50 years of *Management Science*, emphasizes that understanding the nature of a firm's operations "does not just require a theory of human motivation and a theory of material flow; it also requires a means for describing the interaction between the two" (pp. 5). He further speculates that behavioral factors could be the source of the next *paradigm shift* in the science of management to help explain behavioral *anomalies* that contradict the existing theories.

At a more detailed level, Boudreau et al. (2003) propose a unifying framework for identifying new research opportunities in the interaction of operations and human resources management. They argue that operations and human resources, the two traditionally separated fields, are intimately related at a fundamental level: OM provides contextual insights that help explain the effects of human resources activities, such as training and communication, while HRM provides behavioral insights that help explain variations of human responses in OM systems. Therefore, by probing their intersection, the precision and rigor of models in both fields can be improved significantly.

Similarly, Bendoly et al. (2005) developed a framework that divides the types of behavioral assumptions typically made in analytical OM models into three broad categories: intentions, actions, and reactions. These authors provide a literature review on behavioral research published within the last twenty years that uses an experimental approach to test or generate an OM theory. Their review shows that behavioral issues arise in a wide range of operational subfields, including production control, supply chain management, quality management, and operations technology, etc. They also point out several future research directions in the emerging field of *behavioral operations management*.

Coincidently, leading research journals such as the *Journal of Operations Management, Manufacturing and Service Operations Management* and *Decision Science* have created special issues to promote research work that merges elements of behavioral theory and OM. And research work that addresses such relationships is encouraged to take a number of forms and a variety of approaches.

In answering the call for behavioral research, this dissertation focuses on the impact of human factors on the management of supply chains, a sub-discipline of OM. This investigation has been undertaken at two levels. The first examines the managerial decision-making processes of *individual* supply chain members, identifying decision biases and ways to overcome them; the second looks at the *organizational* behavior of the integrated supply chain, testing various mechanisms to enhance channel coordination.

The main research method for this study was a laboratory approach that was first developed by experimental economists to validate game theory models (Kagel and Roth 1995). It uses controlled laboratory settings to confront decision makers (usually college students) with relatively simple decision tasks designed to conform to the assumptions of the theoretical models being tested.

Subjects in experiments are typically motivated by real financial incentives based on their decision quality. In other words, participants have the appropriate economic incentive to take the simulated decision task seriously. This methodology has been applied to studies in economics and psychology as well as business. And the close connection between game theory and experimental methods in economics suggests that experiments can be an important tool for testing theoretical results in OM.

This dissertation includes three interrelated essays in which a series of experiments are designed to simulate the operations of a supply chain. Specifically, Chapters I and II provide an empirical test of the supply chain contracting model. This model uses a simple setting with a two-member supply chain in which the retailer faces the newsvendor problem, the supplier has no capacity constraints, the delivery occurs instantaneously, and no inventory or lost sales are carried over to the next period. The behavioral assumptions commonly made in the model are that supply chain members behave in a way that maximizes their own expected profits, which implies that decision makers are *rational*, *risk-neutral* and *self-interested*.

Chapter I examines how individual retailers and suppliers respond to various formats of supply chain contracts. Human participants in the experiments take the role of a retailer or a supplier. They are asked to interact with a computer simulated counter-partner under either a wholesale price contract, or a buyback contract or a revenue-sharing contract. Our experimental results in Chapter I show that retailers have decision biases and often deviate from the optimal newsvendor solutions, whereas suppliers tend to behave risk-aversely rather than risk-neutrally.

The experiments described in Chapter II are designed to investigate the strategic interactions between supply chain members as the contract type varies. Pairs of human participants are asked to act as partners repeatedly over a finite period of time. The results suggest that channel members have preferences over both their pecuniary payoff and their relative payoff standing. In other words, they are not purely self-interested but rather fairness-concerned.

In Chapter III, we investigate mechanisms to reduce the *bullwhip behavior*, using the well-known *Beer Distribution Game*. The bullwhip effect refers to a phenomenon that the variability of orders in supply chains increases as one moves closer to the source of production. The Beer Game involves four interrelated supply chain members managing their inventory levels as a team over multiple periods (i.e., inventory or lost sales are accumulated over time). And we find that when participants have obtained system-wide training experience and are allowed to communicate with their channel partners, the variability of orders in supply chains can be largely decreased.

To summarize, the dissertation is organized in terms of the increasing complexity of the experimental settings, from single decision makers in a single-period repeated game (Chapter I), to multiple decision makers in a single-period repeated game (Chapter II), and to multiple decision makers in a multi-period game (Chapter III). It systematically examines the effects of several important behavioral issues on the performance of supply chains and contributes to bridging the gap between two traditionally divergent fields of behavioral theory and operations management.

CHAPTER I: IMPACT OF CONTRACT FORMATS ON BEHAVIORS OF INDIVIDUAL SUPPLY CHAIN MEMBERS

1. Introduction and Motivation

There has been a great deal of recent interest in supply chain coordination in general, and more specifically, in contracting mechanisms that can be used to coordinate supply chains. Cachon (2003) reviews the basic models that start with the observation that whenever a supplier charges a wholesale price to a retailer in excess of his own production cost, *double marginalization* causes sub-optimal supply chain performance. Spengler (1950) was the first to write about the double marginalization problem. The essence of this problem is that the wholesale price above production cost creates incentives for the retailer to order less than the profit-maximizing order quantity for the entire supply chain.

We do not attempt to present an exhaustive literature review here, but mention several studies that are representative of the scope of recent contracting literature (we direct the reader to Cachon 2003 for a comprehensive review of the literature). Theoretical studies include Donohue 2000, who studies supply chain contracts with demand updating in fashion industries. Choi et al. (2004) develop a supply contract menu that combines supplier's service level and expected backorders. Kamrad and Siddique (2004) show the flexibility of supply contracts can be Pareto improving. Corbett et al. (2004) examine the value to a supplier of being able to offer contracts that are more general than wholesale price, including two-part linear and nonlinear schemes. Attention has also been paid to procurement contract (Wu and Kleindorfer 2005), option contract (Burnetas and Ritchken 2005, Kleindorfer and Wu 2003), warranty contract (Balachandran and Radhakrishnan 2005), and target-rebate contract on false failure returns (Ferguson et al. 2005).

Lariviere and Porteus (2001) analyze the impact of market size and relative variability on the performance of the wholesale price contract and consider factors such as retailer power that may lead the manufacturer to set a wholesale price below the double marginalization prediction. Cachon (2004) analyzes how three types of wholesale price contracts of push, pull and advance-purchase, determine the allocation of inventory risk and the supply chain efficiency.

Empirical studies have been relatively rare, probably because of the difficulties in collecting field data. Mortimer (2004) analyzes the effect of using revenue-sharing contract on supply chain members' profits and consumer welfares in video-rental industries. Gopal et al. (2003) study contract choice for offshore software development projects. Azoulay and Shane (2001) use evidence from business format franchising industries to argue that different information about contracting of entrepreneurs determines the efficiency of the contract implemented.

In the conclusion to his review article, Cachon calls for a dialogue between empirical and theoretical work through empiricism with a strong theoretical backing:

> As a first step towards wider implementation, …[supply chain contracting] … research needs to develop an empirical-theoretical feedback loop. …[T]he literature contains a considerable amount of theory, but an embarrassingly paltry amount of empiricism. Thus we have little guidance on how the theory should now proceed. …If we observe that firms choose noncoordinating contracts, than we need an explanation. Irrational or incompetent behavior on the part of managers is a convenient explanation, but it is not satisfying to build theory on irrational behavior. A theory is interesting only if it can be refuted and irrational behavior cannot be refuted. A better approach is to challenge the assumptions and analysis of the theory. With some empiricism we should be able to identify which parts of the theory are sound and which deserve more scrutiny. (Cachon 2003, pp. 330-331).

Our approach is to begin the dialogue with theory by comparing the wholesale price, the buyback, and the revenue-sharing contracts in the controlled setting of the laboratory[1]. The

[1] This method of starting with basic theory and developing an empirical-theoretical feedback loop has been explored by experimental economists; see the many examples in Kagel and Roth 1995. Articles in a special issue of *Interfaces* describe how this approach has been applied to a number of institutions and problems in business practice (Bolton and Kwasnica 2002). Roth 2002 discusses how theory together with laboratory experiments contributed to the design of FCC spectrum auctions and the design of the American Medical Association algorithm for matching interns with hospitals.

objective of our study is to investigate the effect of behavioral assumptions on the ability of mechanisms to coordinate the supply chain. A better understanding of the behavior will in turn help provide a direction for future theoretical work, completing the empirical-theoretical feedback loop.

Cachon (2003) points to two empirical questions that can potentially be critical for successful practical design of supply chain coordination mechanisms:

1. Given that many different mechanisms can coordinate the supply chain, we would like to better understand why certain contracts are (or should be) adopted over certain other contracts.
2. Given that theoretical analysis relies heavily on the assumption of expected profit-maximizing behavior, we would like to better understand what effect deviations from this behavior have on mechanism performance.

We organize our paper into two studies. The first study focuses on the retailer behavior when faced with the wholesale-price, buyback and the revenue-sharing contracts. There is a substantial amount of evidence that laboratory participants do not order optimally in the newsvendor setting with a simple wholesale price contract (see section 1.3), but coordinating contracts have not been previously studied. The second study looks at the supplier behavior, and their willingness to offer coordinating contracts. In the next section we summarize the analytical results from the supply-chain contracting literature that are pertinent for our study and describe the basic experimental design. In sections 1.3 and 1.4 we describe methods specific to each study, formulate research hypothesis, and present results of the two studies. In section 1.5 we summarize our finding, point to directions for future research, and discuss managerial implications of our work.

2. Analytical Background and Laboratory Implementation

Although some of the recent theoretical studies (several cited above) consider more complex environments, the simplest setting analyzed theoretically, and the one we investigate in this paper, is one in which the retailer faces the classic newsvendor problem, and orders from a supplier. The supplier has no capacity constraint and delivers instantaneously. In the baseline model, the *wholesale price contract*, the retailer faces an exogenous stochastic demand and a market price, and suffers losses whenever his actual order quantity differs from the realized demand. In contrast, the supplier incurs no risk because he simply makes a profit on the entire retailer's order. To *coordinate* the supply chain, a contract must induce the retailer to order the same amount that would be optimal in a centralized setting. There are many types of contracts that can do this, and our focus is on two of the simplest *risk-sharing*[2] contracts: the *buyback* and the *revenue-sharing* contracts.

In the buyback contract the supplier assumes some of the risk associated with over-ordering by providing the retailer with a rebate for all units unsold at the end of the selling season. In the revenue-sharing contract, the supplier induces higher retailer order through a lower wholesale price, but in return he receives a portion of the gross revenue. It turns out that the two contracting mechanisms are mathematically equivalent, meaning that for each instance of the buyback contract, it is possible to construct an equivalent revenue-sharing contract that can induce the supply chain optimal order quantities and identically allocate supply chain profits.

[2] In both of these contracts the supplier coordinates the supply chain by assuming some appropriate amount of risk associated with demand uncertainty. This is in contrast to the wholesale price contract, in which the retailer assumes all the demand risk. See Cachon 2003 for a detailed description of these and many other contracting mechanisms.

2.1 Analytical Background

In this section we summarize analytical results about the optimal behavior of suppliers and retailers (see Cachon 2003 for details). Let p be the exogenously-determined market retail price per unit, c be the supplier's production cost per unit, q be the retailer's order quantity, and D be the customer demand with distribution of $F()$ and the density function of $f()$. If the retailer orders q, let $S(q) = q - \int_0^q F(y)dy$ denote the expected sales and $H(q) = q - S(q)$ denote the expected left over inventory. Since our game is single-period, no lost sales or excess inventory will be carried to the next period. For simplicity, we assume no penalty for lost sales (although it is straightforward to add the lost sales penalty to the model).

We are looking at three types of contracts, the Wholesale Price (*WP*) contract, the Buyback (*BB*) contract, and the Revenue-Sharing (*RS*) contract. Each contract includes a wholesale price, so let w_{WP}, w_{BB} and w_{RS} be the wholesale price for the WP, BB and RS contract. Also let b be the amount the supplier pays the retailer for unsold units in the buyback contract, and let r be the amount the retailer pays to the supplier for each unit sold in the revenue-sharing contract. The retailer's decision is the order quantity q (which we will subscript with WP, BB or RS to indicate the contract type). Each player is maximizing his expected profit: retailer always with respect to q, the supplier with respect to w and additionally with respect to b and r where appropriate. In Table 1.1 we summarize explicit expressions for expected profit for both players under the three contracts.

	Contract Type		
Decision-Maker	Wholesale price (*WP*)	Buyback (*BB*)	Revenue-sharing (*RS*)
Retailer	$pS(q_{WP}) - w_{WP}q_{WP}$	$pS(q_{BB}) - w_{BB}q + bH(q_{BB})$	$(p-r)S(q_{RS}) - w_{RS}q_{RS}$
Supplier	$(w_{WP} - c)q_{WP}$	$(w_{BB} - c)q_{BB} - bH(q_{BB})$	$(w_{RS} - c)q_{RS} + rS(q_{RS})$

Table 1.1: Expected Profit Functions for Both Players under Different Type of Contract

For the remainder of the paper, to improve exposition, we will drop the subscripts on q and w when the context makes it transparent which contract is being discussed. The well-known solution to the newsvendor problem is that at optimality, the decision-maker is indifferent between the expected shortage and overage costs. The quantity (q^*) that maximized the retailer's expected profit should in general satisfy

$$F\left(q^*\right)=\frac{p-w-r}{p-r-b} \tag{1.1}$$

Equation (1.1) is called the *critical fractile*, and for the entire supply chain it is

$$F\left(q_{SC}^*\right)=\frac{p-c}{p} \tag{1.2}$$

where q_{SC}^* is the first-best order quantity that maximizes the profit of the supply chain. If $w > c$, the first best solution q_{SC}^* is higher than the solution of the wholesale price contract q^*—double marginalization occurs.

The solution that optimizes suppliers' objective function for the wholesale price contract depends on the distribution of D, the customer demand. If D follows a uniform distribution between A and B, $D \sim U(A, B)$, the optimal wholesale price w^* is given by

$$w^* = \min\left\{p, \frac{c}{2}+\frac{p}{2}\frac{B}{B-A}\right\} \tag{1.3}$$

For a buyback contract to coordinate the supply chain—to induce the retailer to order q_{SC}^*—pairs of parameters $\{w, b\}$ need to satisfy

$$\begin{aligned} p-b&=\lambda p \\ w-b&=\lambda c \end{aligned} \tag{1.4}$$

for $0<\lambda<1$, where λ can be interpreted as the retailer's share of the supply chain's profit. Note that there exist multiple such pairs of $\{w, b\}$, independent of the demand distribution. And the actual contract parameters that get implemented (or the realized λ) may depend on the bargaining powers of both parties (Cachon 2003).

Cachon and Lariviere (2005) show that the revenue-sharing contract {w, r} is equivalent to the buyback contract {w, b} in achieving coordination when $r = b$ and $w_{RS} = w_{BB} - b$. Since $w - b = \lambda c$ and $\lambda < 1$, we can see that wholesale price w_{RS} that coordinates the supply chain for the revenue-sharing contract is required to be below supplier's production cost, meaning that in order to coordinate the supply chain with a revenue-sharing contract, the supplier will have to incur an initial loss, and make it up through the sharing of the revenue.

2.2 Laboratory Implementation

In the laboratory the customer demand distribution is discrete uniform and we use two demand levels, high and low. In the low demand condition (DLOW) $D \sim U$ (0, 100) rounded up to the nearest integer, and in the high demand condition (DHIGH) $D \sim U$ (50, 150) rounded up to the nearest integer. We varied the demand distribution in this way to investigate the effect loss aversion (Kahneman and Tversky 1979) has on behavior: retailers can lose money under the wholesale price contract in both demand settings, but suppliers cannot. However, in the DLOW condition retailers and suppliers can also lose money under a coordinating contract, but in the DHIGH condition they cannot. In all our treatments the supplier's production cost is $c = 3$ and the retail price is $p = 12$, which is a setting in which potential supply chain profit is high[3].

We will refer to the setting with the human retailer and automated supplier as the *Retailer Game*, and to the setting with the human supplier and automated retailer as the *Supplier Game*.

[3] Schweitzer and Cachon (2000) used two profit conditions, in the high-profit condition the critical fractile is greater than ½, and consequently the optimal order is above average demand. In the low-profit condition, the critical fractile is below ½, so the optimal order is below average demand. Bolton and Katok (2006) used the same two profit settings and note that it may be that the low-profit condition is somewhat unnatural. We focus on the high profit condition in this paper because it is the setting with higher potential gains from coordination.

We describe methods specific to the two games in sections 1.3.1 (Retailer Game) and 1.4.1 (Supplier Game). Instructions and screen shots used in both games can be found in the Appendix A.

Each study includes three sets of treatments. In the *Buyback* treatments all participants played 100 periods in the wholesale price contract setting and 100 periods in the buyback contract setting, half participants started with the wholesale price setting, and the rest started with the buyback setting. Similarly, in the *Revenue-Sharing* treatments each subject played 100 periods in the wholesale price contract setting and 100 periods in the revenue-sharing contract setting, with the order varied across subjects. In both studies, we repeated the Buyback and Revenue-Sharing treatments in DLOW and DHIGH demand conditions. The third set of treatment, that we label *Same-Frame*, included each subject playing 100 periods in the Buyback contract setting and 100 periods in the Revenue-Sharing setting (with the order switched for half the subjects). These treatments were conducted in the DHIGH demand condition, but the customer demand was presented to the subjects as 50 units guaranteed and an additional number of units uniformly distributed from 0 to 100. The retailer's decision task was described to the subjects as deciding on the number of units to order in addition to the 50 guaranteed units. Thus, although the Same-Frame treatments used the DHIGH demand distribution that eliminates the possibility of negative profits, they also used the DLOW decision frame.

Having each subject make decisions under two contracts, and comparing decisions for each subject is called *within-subjects* design. The main advantage of using the within-subjects design is that it increases statistical power by automatically controlling for individual differences across subjects (Camerer 2003, pp. 41-42). A disadvantage of the within-subjects design is that laboratory sessions last longer, and since participants have to complete two different tasks, it is important to test for the *order effects*. Order effects refer to the possibility that experience in the first task might bias the behavior in the second task, and the standard methods for checking for the order effects is to vary the order of the tasks for different subjects, and then compare the outcomes of a task for the participants who performed it first to the participants who performed it second (Camerer 2003, p. 40). We employed this method in all of the treatments and found no evidence of order effects in any of them.

In summary, we have a $2\times3\times3$ design that manipulates the decision-maker's role (Retailer Game and Supplier Game), the two contract types used for a within-subject comparison (Wholesale Price and Buyback, Wholesale Price and Revenue-Sharing, and Buyback and Revenue-Sharing), and customer demand distribution (DLOW, DHIGH, Same-Frame). We summarize the design and the sample sizes in Table 1.2.

		Contract Combination		
	Demand Condition	WP / BB	WP / RS	BB / RS
Retailer Game	DLOW	15	15	
	DHIGH	17	16	
	Same-Frame			20
Supplier Game	DLOW	17	18	
	DHIGH	18	20	
	Same-Frame			19

Table 1.2: Experimental Design and Sample Sizes

In total 175 subjects participated in our study. Each session lasted for approximately 75 minutes and average earnings, including a $5 participation fee, were $18. All sessions were conducted at the Laboratory for Economic Management and Auctions (LEMA) at Penn State Smeal College of Business during the summer of 2005. Participants were Penn State students, mostly undergraduates, from a variety of majors, recruited through a web-based recruitment system, with earning cash being the only incentive offered. The software we used was web-based and was built using PHP and mySQL.

3. Study 1: The Retailer Game

3.1 Methods

The main purpose of this study is to test the theoretical predictions about the behavior of the retailer in the three contracts. We set the levels of w in a way that would maximize the supplier's own expected profit, using (1.3):

$$w_{WP}^{DLOW} = \min\left\{12, \frac{3}{2} + \frac{12}{2}\frac{100}{100-0}\right\} = 7.5$$

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

For the buyback contract, we use (1.4) and set $\lambda = 1/3$ so that both parties can benefit from coordination, to get $w = 9$ and $b = 8$. We then construct the equivalent revenue-sharing contract $r = 9$ and $w = 1$. Given those 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 the Same-Frame conditions.

3.2 Research Hypothesis

Our first two hypotheses follow directly from the theory. Under the wholesale price contract the expected profit-maximizing retailer will order less than the first-best order quantity q_{SC}^* due to double marginalization, so we have:

Hypothesis 1A (*Double Marginalization*): the average order quantity under the wholesale price contract is below the first-best order quantity q_{SC}^*. Specifically, theory implies q = 37.5 in the DLOW condition, which is below the first-best quantity of 75, and 62.5 in the DHIGH condition, which is below the first-best quantity of 125.

We expect both the buyback and the revenue-sharing contracts to induce higher orders from the retailer than the wholesale price contract, since both can coordinate the supply chain.

Hypothesis 2A (Coordination): the average order quantity under the coordinating buyback and the revenue-sharing contracts should be higher than the average order quantity under the wholesale price contract. Theoretical predictions are 75 for DLOW and 125 for DHIGH.

There is a substantial amount of empirical evidence that when laboratory participants solve the newsvendor problem, the average order quantity is biased towards average demand. Schweitzer and Cachon (2000) were the first to show that in a laboratory newsvendor setting similar to the wholesale price contract, the retailer's behavior of ordering on average between the optimal order and the average demand is inconsistent with many behavioral theories. They called the ordering pattern they observed "anchoring and insufficient adjustment." Bolton and Katok (2006) replicated this result and additionally showed that performance does improve over time with extensive experience, although slowly, and restricting decision-makers to placing standing orders[4] speeds up learning substantially. Lurie and Swaminathan (2005) report a similar finding, that too frequent feedback can degrade performance and slow down learning. Benzion et al. (2005) vary the demand distribution and find that orders are affected by both the average demand and the last period's demand, but this bias is weakened slowly over time—participants learn. In our third hypothesis we conjecture that the similar pattern of "anchoring and insufficient adjustment" will persist in all three contracts.

Hypothesis 3A (Anchoring and Insufficient Adjustment): The average order quantity for the wholesale price contract will be above the theoretical prediction; it will be between 37.5 and 50 in DLOW and between 62.5 and 100 in DHIGH; and the average order quantities for

[4] 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 the case of Bolton and Katok (2006) this was 10 periods).

coordinating buyback and revenue-sharing contracts will be below the theoretical predictions, specifically, they will be between 50 and 75 in DLOW and between 100 and 125 in DHIGH.

The last hypothesis speaks to the fact that the buyback and the revenue-sharing contracts are mathematically equivalent. This equivalence implies that we should not observe any differences in the performance of the two mechanisms.

Hypothesis 4A (Equivalence): Retailers' average orders and the resulting supply chain efficiency in the two coordinating contracts will be identical.

3.3 Results

We start by comparing retailers' average order quantities over 100 periods, and the total supply chain efficiency across the three contracts. Table 1.3 presents the sample means and standard deviations of average orders and efficiencies under different contracts. We calculate the efficiency as the ratio of the actual average supply chain profits to the average profits that would have been achieved with the first-best order, given the actual demand draw. All comparisons are done treating each subject as a single independent observation.

	Double Marginalization (DM)	First-Best (FB)	Wholesale Price (WP)	Buyback (BB)	Revenue-Sharing (RS)
DLOW					
Retailer's Order	37.50	75	42.11 (6.01)	54.55 (8.34)	64.84 (10.86)
Efficiency	0.77	1.00	0.75 (0.0789)	0.86 (0.0860)	0.94 (0.0660)
DHIGH					
Retailer's Order	62.5	125	81.32 (10.13)	113.18 (13.59)	102.25 (18.87)
Efficiency	0.71	1.00	0.83 (0.0612)	0.96 (0.0350)	0.92 (0.0677)

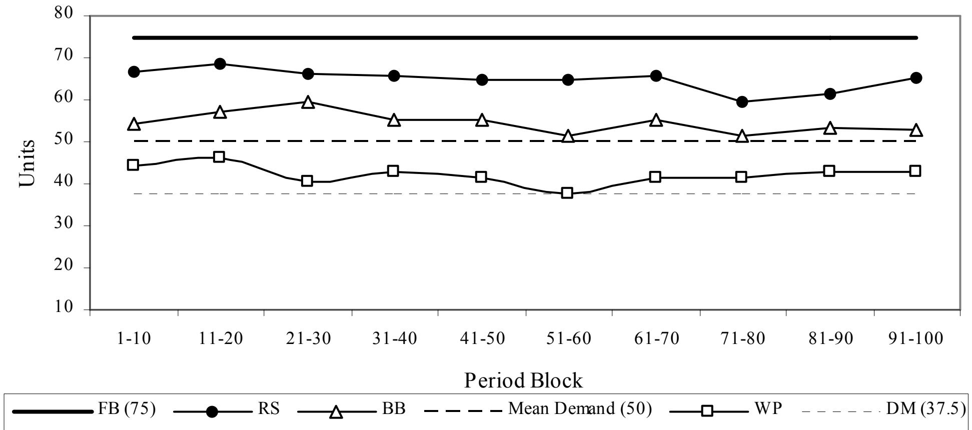
Table 1.3: Retailer Game Results

- *Result 1: Wholesale price contract performs as well or better than the theoretical prediction.* Wholesale price contract induces the average order quantities which are significantly higher than the double marginalization theoretical predictions (p-value < 0.001). The overall efficiency is not significantly different from what it would have been with the retailer-optimal order under DLOW (2-sided p-value = 0.4932), but it is significantly higher than with the retailer-optimal order under DHIGH (p-value < 0.001).

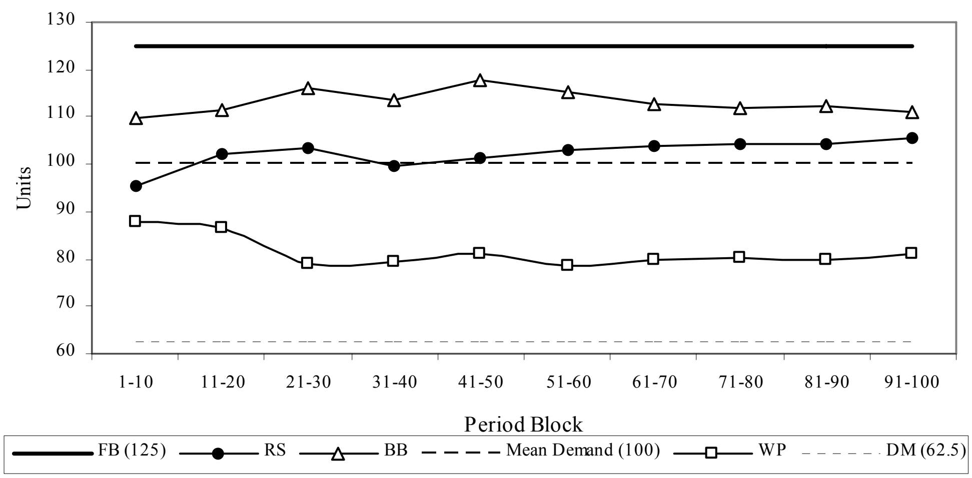
- *Result 2*: *Buyback and Revenue-sharing contracts induce higher orders and achieve higher efficiency than the Wholesale Price contract*. For these comparisons we use the matched-pair *t* test, since we have observations for each participant for the WP and one of the coordinating contracts. The p-values for all tests are below 0.001.
- *Result 3*: *Buyback and Revenue-sharing contracts fall short of achieving full supply chain coordination.* Although both of those contracts induce higher orders than the Wholesale Price contract does, they are significantly below the first-best levels. Resulting supply chain efficiency is also significantly below 100% (p-values < 0.001).

To examine retailers' behavior over time, we plot in Figure 1.1 average orders for the three contracts, aggregated into 10-period blocks to improve exposition. In the same figure we also plot the first-best order, the average demand, and the double marginalization order quantity under the wholesale price contract. So as to quantify any behavioral change, we further compare retailers' average orders in the first 50 periods with those in the second 50 periods for each treatment using matched-pair *t* tests. Results reported thereafter on the time trend are all 2-sided p-values.

- *Result 4*: *Adjustments over time are slow.* In DLOW, there is no change in retailer's orders over time under any of the three contracts (p-values are all above 0.1). In DHIGH, retailers tend to significantly decrease their orders over time in the wholesale price contract (p-value = 0.0488) and weakly increase their orders in revenue-sharing contract (p-value = 0.0542), both towards their own optimal benchmarks. However, no trend is found under the buyback contract (p-value = 0.5462). The ranking of average orders and efficiency levels do not change over time.



(a) DLOW condition



(b) DHIGH Condition

Figure 1.1: Average Orders Over Time in Retailer Game

We now turn to comparing the performance of the buyback and the revenue-sharing contracts. Recall that those two contracts, as implemented in our laboratory setting, are mathematically equivalent. However, if we compare the average retailer orders and supply chain efficiencies under the two contracts (see Table 1.3) we can generally reject the null hypothesis that they are the same (2-sided p-values comparing average orders are 0.0072 under DLOW and 0.0682 under DHIGH, and comparing efficiencies they are 0.0660 under DLOW and 0.0974 under DHIGH). So it appears that at least initially, the two contracts are not perceived as being equivalent by the human subjects.

To narrow down potential causes of these differences in perception we conducted an additional treatment. In the new treatment, the same group of subjects played 100 periods using the buyback contract and 100 periods using the revenue-sharing contract. This design allows us to do the within-subject comparison to see how the two contracts affect each individual's decision-making. The demand distribution in the new treatment was DHIGH, but we presented the ordering decision in the DLOW frame: Participants were told that the customer demand per round consists of two parts: a guaranteed demand of 50 units, and additional demand which is uniform from 0 to 100. We described the decision task as deciding on the amount to order in addition to the 50 guaranteed units. In other words, participants make the same decision (order from 0 to 100) as in the DLOW condition, but without the possibility of losing money. The new treatment included 20 participants, none of whom participated in any of the other treatments in this study. We summarize data on the average orders and efficiency levels for the two contracts in Table 1.4.

	Periods 1-100		Periods 51-100	
	Buyback (BB)	Revenue-sharing (RS)	Buyback (BB)	Revenue-sharing (RS)
Retailer's Order	106.07 (12.12)	111.60 (14.37)	109.94 (13.83)	110.84 (16.07)
Efficiency	0.95 (0.0393)	0.95 (0.0290)	0.95 (0.0480)	0.96 (0.0358)

Table 1.4: Equivalence between Coordinating Contracts in Retailer Game

In Figure 1.2 we plot average orders for the two contracts in 10-period blocks along with the average demand and the optimal order.

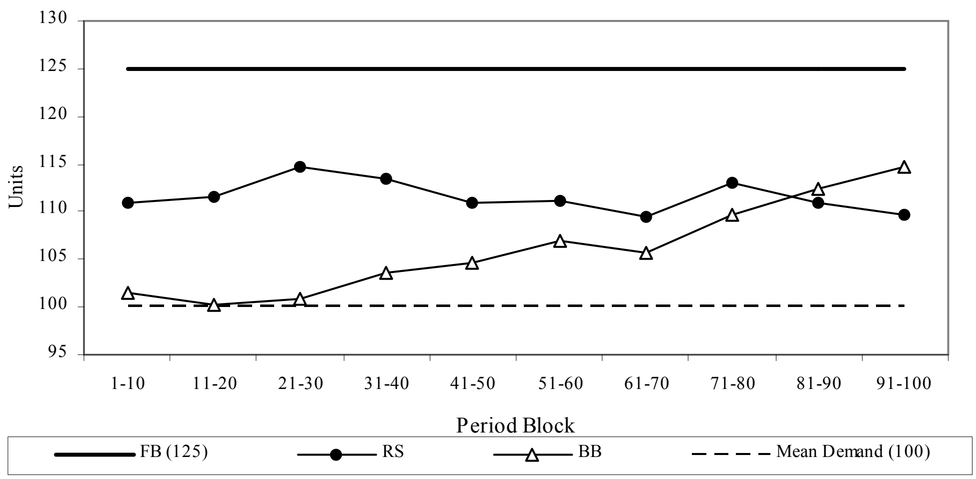


Figure 1.2: Retailer's Average Orders Over Time in Same-Frame

Although the buyback contract initially performs poorly, orders under the buyback contract increase over time, and by midway through the session they reach the same level as the orders under the revenue-sharing contract. Looking at all 100 periods, the average orders under the buyback contract are lower than the average orders under the revenue-sharing contract (p-value from the matched pair *t* test = 0.0062) but if we look at the last half of the session these differences disappear (p-value = 0.6286); efficiency levels are quite high and not significantly different either over all 100 periods (p-value = 0.7498) or over the last 50 periods (p-value = 0.1283). To summarize:

- *Result 5*: *Initial differences in performance between the two coordinating contracts are mitigated by experience.*

Since participants in the Same-Frame treatment have experience with both the buyback and the revenue-sharing contracts, we therefore include a post-game survey that questions them about their ordering policies used in the two coordinating contracts, and whether and why they would “prefer” one form to the other. Feedback from the post-game survey further confirms that retailers were able to detect the equivalence between the buyback and revenue-sharing contracts.

We asked subjects to rate their feelings about the statement that “whether it is a Buyback or Revenue-sharing contract does not impact my ordering decisions”, on a scale of 1-7 (with 1 meaning strongly disagree and 7 meaning strongly agree). 11 out of the 19 participants who answered the survey rated 7 and the overall average score is 5.47. In terms of their preferences over the two contracts, 13 subjects said they had no preference while 5 subjects favored the buybacks and 1 preferred the revenue-sharing contract.

Thus interestingly, although participants tend to “order” more under the revenue-sharing contract at start, more decision makers tend to “prefer” the buyback contract. It may be because that the revenue-sharing contract can create some “pressure” on retailers to realize more sales by acquiring them to share revenue and it thereby effectively induces higher orders, while the buybacks seems to be more “friendly” in providing retailers with protection against harsh consequences and is therefore more preferred. As one of our subject wrote in the survey:

… Buyback, there is always a chance to put a floor on the winnings whereas you put a ceiling on the revenue-sharing contract….

The main finding in the Retailer Game is that coordinating contracts do induce higher orders than the wholesale price contract does, and with those higher orders comes higher supply chain efficiency. So the theory correctly predicts the qualitative shifts observed in the data as well as the detrimental effect double marginalization has on performance—orders induced by the wholesale price contract are substantially below the first best. The “anchoring and insufficient adjustment” heuristic organizes our data well, in that the actual average orders for each contract lie between the optimal order for this contract, and the average demand. So we find qualitative support for hypothesis 1A, 2A and 3A.

The theory falls short in its quantitative predictions. Due to "anchoring and insufficient adjustment," average orders in the wholesale price contract are significantly above the theoretical point predictions, and average orders in the coordinating contracts are significantly below the first best point predictions.

Our evidence about the equivalence of the two coordinating contracts is mixed. Subjects who were exposed to only one of the two contracts do not, on average, perform the same. But when we use within-subject design that also eliminates the possibility of losses, the initial differences in performance disappear over time, consistent with Hypothesis 4A.

4. Study 2: The Supplier Game

4.1 Methods

In the second study we look at the contracting setting from the supplier's perspective to better understand the extent to which suppliers are able and willing to offer coordinating contracts. The design includes an automated retailer programmed to act in accordance with the theory, and this feature helps control for any potential strategic interactions between suppliers and retailers. The retailer is programmed to place optimal orders given a contract offered, according to **Error! Reference source not found.**. Since $F()$ is $U(A,B)$, the automated retailer's order is given by

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

where $A = 0$ and $B = 100$ in the DLOW condition and $A = 50$ and $B = 150$ in the DHIGH condition. 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. And in the revenue-sharing contract we set $b = 0$, and participants selected w and r simultaneously.

After inputting contract parameters into the decision form, suppliers saw the implied automated retailer's order and the expected retailer's profit associated with that order. At that point participants could go back and change contract parameters, and repeat the process as many times as they wanted to before submitting the final decision. We used this procedure to make certain our participants had access to the relevant information the theory implicitly assumes they have, thus giving the theory its best shot.

4.2 Research Hypothesis

In formulating research hypothesis we are guided by the theory, which provides normative benchmarks about the families of contracts that *should* be implemented by rational and risk-neutral players. We also consider risk aversion, as a potential behavioral explanation for deviations from theory. Our first hypothesis deals with the wholesale price contract. In this contract, the supplier's decision does not involve any risk—all the risk is assumed by the retailer. Therefore, our first hypothesis reflects the belief that, absent risk, rational decision-makers will be able to find the profit-maximizing contract.

Hypothesis 1B (Wholesale Price Contracts): Suppliers will select profit-maximizing wholesale prices for the wholesale price contracts; 7.5 in DLOW and 10.5 in DHIGH.

Our second hypothesis deals with the coordinating contracts, and since the two we are studying are mathematically equivalent, we do not distinguish between them:

Hypothesis 2B (Coordinating Contracts): Suppliers will select coordinating contracts that will induce higher retailer orders (and correspondingly higher supply chain efficiency) than the wholesale price contract.

There is evidence that people are risk-averse, even when they play games in the laboratory for relatively modest stakes (see for example Halt and Laury 2002 and references therein) but according to Isaac and James (2000) the degree of risk aversion seems to vary depending on the setting. So the question of whether risk-aversion plays a role in supplier behavior, and if so, what effect it has on mechanism performance, is an important one for mechanism designers. If the suppliers are risk-averse, they will not be willing to assume as much risk as is required to fully coordinate the supply-chain, leading to our third hypothesis:

Hypothesis 3B: (Risk Aversion): Suppliers will assume less risk than is required to implement the first-best contract, leading to retailer orders below 75 in DLOW and below 125 in DHIGH.

Lastly, since the two coordinating contracts are mathematically equivalent, our final hypothesis speaks to this equivalence.

Hypothesis 4B (Equivalence): Buyback and revenue-sharing contracts will be equivalent in their performance: they will induce identical order quantities and result in the same profit divisions and efficiency level.

4.3 Results

We start by summarizing (Table 1.5) average wholesale prices under the wholesale-price contract (w), the average retailer order quantities (q) the contract induced, and the overall efficiency for each contract in DLOW and DHIGH conditions.

	Double Marginalization (DM)	First-Best (FB)	Wholesale Price (WP)	Buyback (BB)	Revenue-sharing (RS)
DLOW					
Wholesale Price	7.5		7.35 (0.34)		
Induced Order	37.5	75	39.07 (2.78)	57.85 (13.40)	49.72 (14.98)
Supplier's Profit Share	0.67	≈1.00	0.65 (0.0364)	0.69 (0.0994)	0.72 (0.0835)
Efficiency	0.77	1.00	0.77 (0.0323)	0.89 (0.0726)	0.85 (0.0897)
DHIGH					
Wholesale Price	10.5		10.01 (0.74)		
Induced Order	62.5	125	66.87 (6.12)	88.95 (20.19)	97.31 (10.96)
Supplier's Profit Share	0.85	≈1.00	0.79 (0.0745)	0.82 (0.0760)	0.78 (0.0772)
Efficiency	0.71	1.00	0.74 (0.0419)	0.86 (0.0803)	0.92 (0.0418)

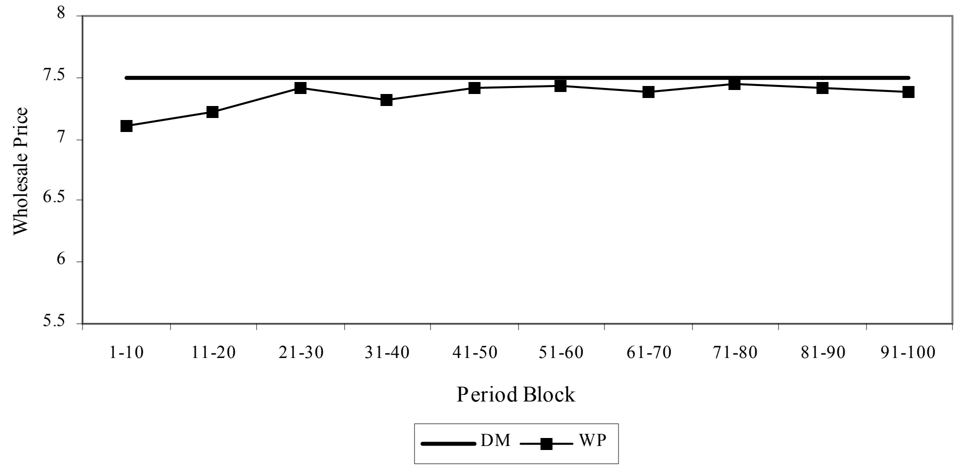
Table 1.5: Supplier Game Results

Additionally, we report in Table 1.5 the average profit share that the supplier obtained in the three contracts. We also present, for comparison purposes, the relevant predictions for the optimal wholesale price contract and for the first-best contracts. Figure 1.3 shows the average wholesale price *w* over time (grouped in 10-period blocks to improve exposition) and compares it to the optimal wholesale price levels of 7.5 for DLOW and 10.5 for DHIGH.

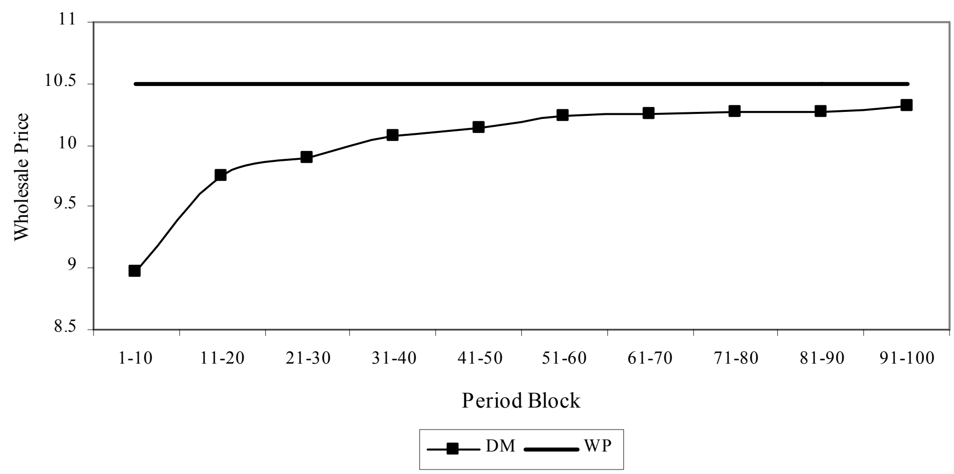
- *Result 6: Suppliers find optimal wholesale price contracts.*

In both DHIGH and DLOW conditions, the average prices chosen are slightly below optimal (p-value = 0.0300 in DLOW and < 0.001 in DHIGH), inducing average order quantities that are slightly above predicted (p-value = 0.004 in DLOW and < 0.001 in DHIGH). As a result, the supply chain efficiency is above predicted (p-value < 0.001 in DLOW and DHIGH) and the supplier's share of the profits is also below predicted (p-value = 0.0088 in DLOW and 0.0001 in

DHIGH). But looking at Figure 1.3, it is clear that most of the deviation happens early in the game, and over time participants are able to find near-optimal wholesale price contract.



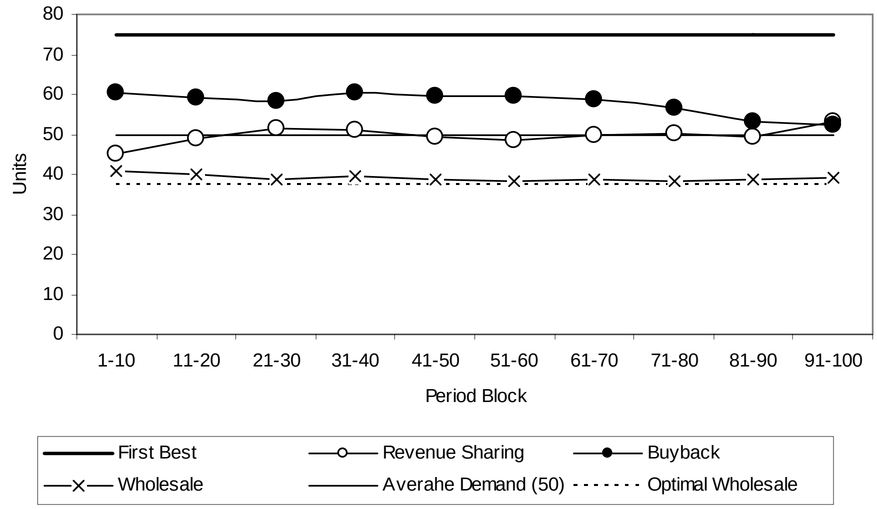
(a) DLOW



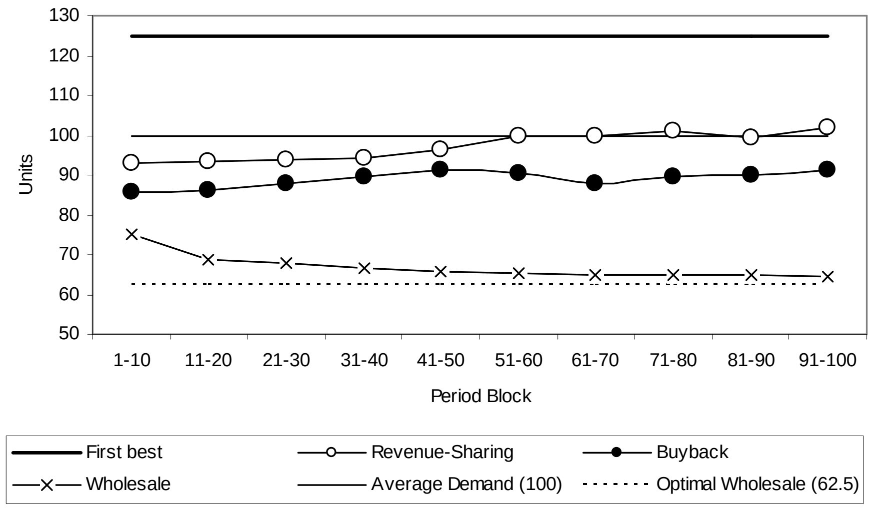
(b) DHIGH

Figure 1.3: Average Wholesale Prices Over Time

Figure 1.4 shows the average order quantities q induced by each of the three contracts as it changes over time. As before, periods are aggregated into 10-period blocks for clearer exposition.



(a) DLOW



(b) DHIGH

Figure 1.4: Induced Order Quantities Over Time in Supplier Game

- *Result 7: Order quantities induced by the coordinating contracts are higher than order quantities induced by the wholesale price contract* (all p-values < 0.005). The higher order quantities also lead to significantly higher efficiencies (all p-values < 0.001), so suppliers are able to extract value from coordinating contracts in both demand conditions. Interestingly, suppliers are not able to come close to extracting most of the profit from the supply chain; in fact, the suppliers' profit share is only slightly higher under the coordinating contracts than under the wholesale-price contract (1-sided p-values = 0.0612 for BB and 0.0014 for RS under DLOW and 0.0667 for BB and 0.4916 for RS under DHIGH).

Since coordinating contracts involve two parameters, to measure how close actual contracts offered are to some coordinating contract, we compute and report in Table 1.6 the average coordinating rebate given the actual wholesale price in the buyback condition ($b^* \mid w$) and the average coordinating wholesale price given the actual revenue share in the revenue-sharing condition ($w^* \mid r$).

	DLOW		DHIGH		Same-Frame	
	Buyback (BB)	Revenue-sharing (RS)	Buyback (BB)	Revenue-sharing (RS)	Buyback (BB)	Revenue-sharing (RS)
Wholesale Price (w)	8.90 (1.05)	3.86 (2.01)	10.27 (0.74	2.65 (1.07)	9.74 (1.36)	2.47 (1.77)
$w - b$	2.57 (1.13)		3.86 (2.63)		3.27 (1.98)	
Rebate or Revenue (b, r)	6.33 (1.91)	5.04 (2.67)	6.40 (2.83)	7.23 (1.57)	6.47 (2.75)	7.53 (1.96)
$w^*\|r$ and $b^*\|w$	7.86 (1.40)	1.74 (0.67)	9.69 (0.98)	1.19 (0.39)	8.99 (1.81)	1.12 (0.49)
Induced Order	57.85 (13.40)	49.72 (14.98)	88.95 (20.19)	97.31 (10.96)	100.04 (21.00)	101.09 (16.36)
Supplier's Profit Share	0.69 (0.0994)	0.72 (0.0835)	0.82 (0.0760)	0.78 (0.0772)	0.77 (0.1409)	0.79 (0.0744)
Efficiency	0.89 (0.0726)	0.85 (0.0897)	0.86 (0.0830)	0.92 (0.0418)	0.91 (0.0735)	0.92 (0.0737)

Table 1.6: Equivalence between Coordinating Contracts in Supplier Game

- *Result 8: Suppliers are not willing to share enough risk to coordinate the supply chain.* In the buyback contract, the actual rebate b is substantially below values needed to coordinate the supply chain (p-value = 0.0062 in DLOW and < 0.001 in DHIGH and 0.0022 in Same-Frame). The wholesale price w in the revenue-sharing contract is substantially higher than it would need to be to coordinate the supply chain (p-values <

0.001 in DLOW and DHIGH and 0.0042 in Same-Frame). This behavior is consistent with suppliers acting as if they were risk averse.

Like in the Retailer Game, we observe some differences in the performance of the buyback and revenue sharing contracts (see Figure 1.4 and the DLOW and DHIGH sections of Table 1.6). Even across subjects these differences disappear towards the end of the session in the DLOW condition, but not in the DHIGH condition. We compare the performance of the two contracts directly within subjects using the Same-Frame setting, and present the induced retailer's order over time in Figure 1.5, and descriptive statistics in the Same-Frame section of Table 1.6.

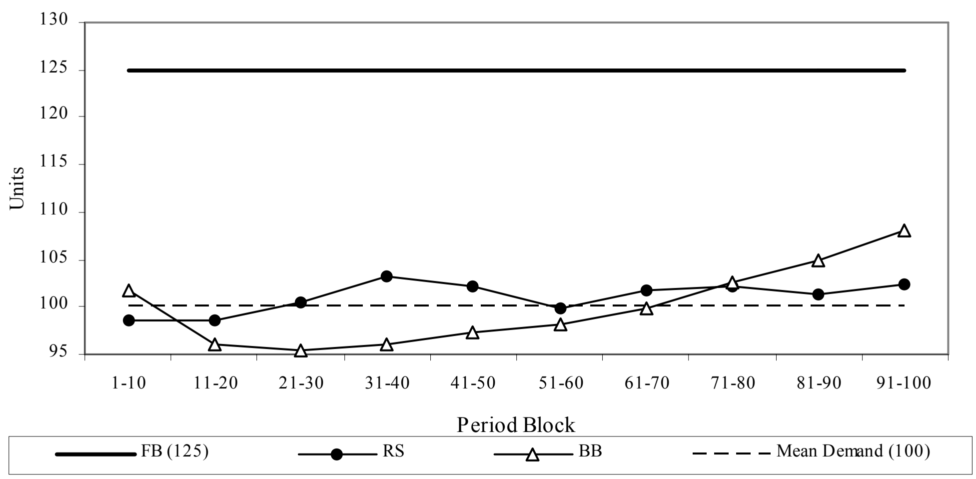


Figure 1.5: Supplier's Induced Orders Over Time in Same-Frame

- *Result 9: The two coordinating contracts do not differ significantly in their overall performance,* as indicated by the average orders induced (matched-pair *t* test 2-tailed p-value = 0.7871), supplier's profit share (p-value = 0.4826) and overall efficiency (p-value = 0.1623).
- *Result 10: The two coordinating contracts differ in how they are implemented.* For the two contracts to be equivalent, the difference between the wholesale price and rebate in

the buyback contract should equal to the wholesale price in the revenue-sharing contract ($w_{BB} - b = w_{RS}$) and the rebate has to equal the revenue share (b = r). However, average wholesale prices in the revenue-sharing contract are lower than the average differences between the wholesale price and rebate in the buybacks (p-value = 0.0084), while the average revenue share is higher than the average rebate (p-value = 0.0181).

Another interesting observation lies in answers to the post-game survey question about which contract the supplier prefers. Among the 16 samples we collected, 11 subjects said they would prefer the revenue-sharing contract, 4 of them said the buybacks and only 1 participant claimed no preference. And according to the survey, more decision makers found the revenue-sharing contract to be more "straightforward" such that they could find a combination that yields good payoff relatively quickly – if the supplier wishes a big slice of the retailer's revenue he will *have to* lower his wholesale price. In contrast, it is not as intuitive as in the buybacks that if the supplier wants to charge a high wholesale price he should accompany it with a high rebate.

We find that in the Supplier Game suppliers are able to take advantage of coordinating contracts and find contracts that induce higher retailer orders than the wholesale price contract. These higher orders lead to higher supply chain efficiency. While suppliers are able to find near-optimal wholesale price contracts, the coordinating contracts they choose involve significantly less risk sharing than is required to coordinate the supply chain. That is, suppliers are unwilling to share enough risk to coordinate the supply chain. So we find support for Hypothesis 1B and 3B, but only partial support for Hypothesis 2B. Suppliers fall significantly short of being able and willing to implement coordinating contracts.

The differences between the buyback and revenue-sharing contracts in the DLOW and DHIGH conditions are small and tend to lessen over time. When we compare the two coordinating contracts within subjects, differences in performance disappear, although there are still detectable differences in the way our participants choose to implement the two contracts. So overall, we conclude that we find some support for Hypothesis 4B.

5. Conclusions

We present a laboratory study in which we compare the actual performance of the wholesale price contract and two coordinating risk-sharing contracts: buyback and revenue-sharing. We compare these mechanisms in two ways: first, from the retailer's perspective, we look at how retailers respond to different mechanisms. Second, from the supplier's perspective, we look at the suppliers' willingness 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. This "anchoring and insufficient adjustment" behavior causes the wholesale price contract to perform better than it should according to the theory, because the higher-than-optimal order quantities counteract the effects of double marginalization. This persistent behavioral bias also causes coordinating contract to perform worse than the theory suggests. This occurs because actual order quantities do not increase by as much as they should as a result of coordinating contracts.

One important consequence of this finding is that it may well be worthwhile for suppliers who are considering using coordinating risk-sharing contracts to implement mechanisms that help retailers learn to overcome the demand following behavior that results in not ordering enough high-profit products. For example, Bolton and Katok (2006) find that restricting retailers to placing standing orders improves performance significantly through faster learning. Our present study indicates that those and potential other ways of helping retailers make better ordering decisions may improve the ability of contracting mechanisms to coordinate supply chains.

When we look at contracts from the suppliers' perspective, we find that suppliers do take advantage of coordinating contracts to some extent, but because they are unwilling to share enough risk, the contracts they offer do not fully coordinate the supply chain. One explanation for the suppliers' behavior is risk aversion—suppliers simply prefer contracts that are on average less profitable, but also less risky (see Gan et al. (2004) for theoretical analysis on coordinating risk-averse agents in supply chains). An alternative explanation may be that suppliers are simply unable to find profit-maximizing coordinating contracts. After all, these contracts have two parameters, and searching a two-parameter space is a significantly more complex task than searching a one-parameter space.

A promising direction for future research is to study the game in which both players are human. Such a game is likely to have strategic considerations (and this was the reason that we did not use it in the present study) that may be difficult to separate from other considerations. For example, Keser and Paleologo (2004) report a tendency for players to choose wholesale price contracts that split profits approximately equally when both players have equal market power. It would be interesting to see the extent to which players in the laboratory are able to take advantage of market power and implement wholesale price contracts that favor one side or the other.

CHAPTERII: FAIRNESS CONCERNS – THE IMPACT OF LONG-TERM RELATIONSHIP ON SUPPLY CHAIN CONTRACTS

1. Introduction and Related Literature

In Chapter I, we have observed laboratory evidence that the newsvendor retailers have demand-chasing decision biases and that the suppliers are unwilling to share a sufficient amount of risk to achieve coordination. These experimental results significantly deviate from what the supply chain contracting theory assumes regarding the channel members' *individual* behaviors. In this chapter, we will focus on testing another hypothesis commonly assumed in the contracting model, which is that both the retailer and the supplier are *self-interested* and care only about their own monetary payoffs.

The self-interest hypothesis is the basic assumption of most economic models. However, experimental economists have shown that in many situations, especially in those with repeated social interactions, people tend to be concerned with *fairness*. In other words, decision makers are motivated by both their pecuniary payoff and their relative payoff standing (Bolton and Ockenfels 2000). Recent economic research has thus developed theories to rigorously explain the observed behavioral anomalies (Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Cox et al. 2004).

Previous behavioral research suggests that in many cases, firms like individuals may have incentives to act in a fairness-minded manner (Kahneman et al. 1986). Field studies in the business literature also consistently show that in industries such as the fashion industry (Uzzi 1996) and the automobile dealership industry (Kumar 1996), fairness plays an important role in maintaining qualified business relationships.

Some quite recent theoretical work in the supply chain contracting literature has attempted to incorporate these social concerns. Cui et al. (2004) model fairness using an inequality-averse utility function and show how fairness may affect the performance of the wholesale price, two-part tariff and quantity discount contracts. Wu and Loch (2006) consider social preferences such as reciprocity, status and group identity, and examine their impact on the wholesale price contract over one-period and multi-period horizons. Both papers assume the linear inverse deterministic demand function and demonstrate that supply chain efficiency can be enhanced with the simple linear pricing schema if channel partners care about relative payoffs.

Besides the experiments presented in Chapter I, three other laboratory studies also provide empirical tests of the supply chain contracting theory. Unlike our Retailer Game and Supplier Game in which one supply chain member is automated by the computer, these experiments involve two human decision makers interacting with each other. Yet they still focus on the impact of individual decision biases rather than that on social concerns. Ho and Zhang (2004) and Lim (2004) investigate behaviors under both linear and nonlinear contracts in a one-shot game. In their experiments, the retailer faces a deterministic demand and sets the retail price. And participants are paired with a new partner in each round. Ho and Zhang (2004) find laboratory evidence that participants act consistent with loss aversion in the two-part tariff contract; while Lim (2004) shows that regret avoidance influences participants' decision in the quantity discount contract. Keser and Paleologo (2004) test a model where the retailer faces a stochastic demand and determines the order quantity. And the same retailer and supplier pairs interact repeatedly over 30 rounds under a wholesale price contract. Their results show that the suppliers tend to anchor their wholesale prices on the middle point between the retail price and the production cost, whereas the retailers order significantly less than their best responses to the proposed wholesale prices.

In this study, we will focus on the *strategic interaction* between the retailer and the supplier under different supply chain contracts. Specifically, we are interested in the influence of *distributive fairness* (Tyler and Lind 1992) on the performance of three types of contracts we studied before: the wholesale price contract and two risk-sharing coordinating contracts of buyback and revenue-sharing. Distributive fairness can be viewed as an evaluation of the channel partner's relative rewards or losses compared to its respective contributions (Frazier 1983), which therefore reflects supply chain members' strategic decisions on how to divide the benefits and burdens of the channel. Our research questions are: (a) whether or not fairness concerns are relevant to decision making in the supply chain contracting model, and (b) if fairness considerations do exist, whether they will help achieve higher supply chain efficiency as implied by prior theoretical work (i.e., Cui et al. 2004, Wu and Loch 2006).

2. Experimental Design, Hypothesis and Implementation

2.1 Game Design

Since the main interest of our study is the behavioral interaction between channel members, we simulate a supply chain that consists of a single retailer and a single supplier. In this game (which we refer as *Contracting Game* from now on), both roles are played by human subjects. To give the fairness hypothesis its best shot, the same pairs of decision makers are asked to play against each other repeatedly for 100 rounds (i.e., a *finite-repeated* game). Note that such repeated partnership is common in practice and may help foster reciprocal behaviors. We manipulate the contract formats used in the game, resulting in three treatments of Wholesale Price (WP), Buybacks (BB) and Revenue-sharing (RS). Following the experiments in Chapter I, a stochastic customer demand uniformly distributed between 50 and 150 (i.e., the DHIGH condition) is used to approximate the uncertainty in all three conditions. And the same sequence of random demand is used in all experiments. In addition, the supplier's production cost is set at $c = 3$ and the retail price is fixed at $p = 12$ as before.

In each round, the supplier moves first to specify the contract terms, depending on the treatment. In WP, the suppliers are asked to decide a linear wholesale price (w), whereas in BB or RS, they need to determine a combination of wholesale price and rebate (b) or revenue share (r) simultaneously. The retailer then responds to the contract offer by placing orders before the random demand of that period is realized. We restrict the retailer to have only three ordering options. The retailer is offered the ordering solution that maximizes his own expected profit given his supplier's offer. And this "optimal order" is automatically computed by the game software according to the contracting theory. Alternatively, the retailer can order the minimum possible demand which is 50. Such an option protects the retailer against the stochastic demand, but both parties' profits will be reduced comparatively. If the retailer is not satisfied with the contract, he can reject it (i.e., order zero), in which case both parties will have zero profit for the round.

We note that our game design has the same basic structure to that of Keser and Paleologo (2004), yet it differs in several important respects. First, the game software we implement computes automatically for both players the expected-profit-maximizing order quantity given a proposed contract. We allow the supplier to try as many different offers as he wants, and the computer will inform him of the corresponding optimal orders of the retailer for each attempt. After the supplier makes his final decisions, the computer will display for the retailer the terms of the current offer and his best response to it. The retailer then decides whether to place the optimal order, order 50, or reject the contract. Figure 2.1 provides an example of screen shots in the buyback contract for the retailers.

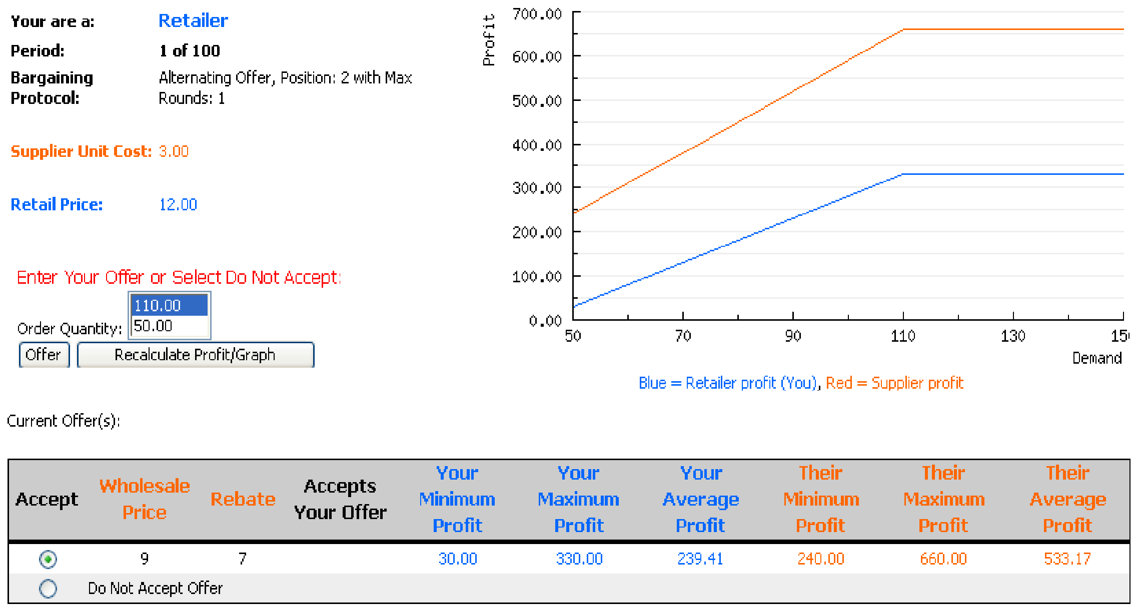


Figure 2.1: Retailer’s Decision Interface in the Buyback Contract

On the other hand, the decision makers in Keser and Paleologo (2004) were allowed to place any integer order that is nonnegative but had to do the calculations (if any) by themselves. We have already seen much evidence on the newsvendor retailers’ suboptimal behaviors in the Retailer Game and in other studies cited in Chapter I. In this Contracting Game, retailers have to adjust their ordering decisions according to the dynamic change of the suppliers’ contract in every round and thus their learning opportunities are not as salient as in previous experiments. By providing computational decision aids and limiting the retailers’ ordering options, we can control for individual decision biases and help accelerate learning so that participants can focus on the strategic division of the supply chain profits.

Moreover, our study provides direct and public information on the profit outcomes of each supply chain member. Over the game’s decision interface, the supplier is able to view the expected profits of both parties based on his current proposed contract. Similarly, the retailer can see how the expected profits of the two members change according to the ordering option he selects. As illustrated in Figure 2.1, the above information is summarized by simple statistics in a table and the expected profit distribution for the entire demand range is plotted in a graph as well. The computer also shows both players the actual profits each member receives after the stochastic demand is revealed for the current round and for all previous rounds. In contrast, players in Keser and Paleologo (2004) only observed their own realized profits. Although they had the information required to compute the other party's expected and realized profits, no such information was explicitly displayed in their experiments.

Lastly, we use a different payment scheme to motivate our subjects. Our participants (no matter what role they play) are paid according to their own total profits accumulated over 100 rounds. In Keser and Paleologo (2004), the payoff of the retailer (supplier) was computed relative to the average profit of the other retailers (suppliers) who took part in the same session. In other words, their payment scheme models the competition between supply chains and can avoid extreme differences in the payments to suppliers and retailers, whereas ours mimics the competition within a supply chain and can potentially cause a large payment gap between pairs of players.

One may also find similarity between our two-person Contracting Game and the *Ultimatum Game* much studied in the economic literature (see references in Kagel and Roth 1995). In the ultimatum experiments, one player proposes how to split some fixed amount of money, while the other decides whether to take the deal or not. If the offer is rejected, both players get nothing. Even though the game theory solution suggests that the first mover offers the minimum possible share and that the responder always accepts it, quite often the outcome is a fair share and rejections of unfair offers frequently occur. And on average, the majority of proposers offer 40 to 50% of the total sum, and about half of all responders reject offers below 30% (Bolton and Zwick 1995, Kagel and Roth 1995).

Our contracting model has two unique features. First, the total supply chain profits to split are no longer fixed but rather random due to the stochastic demand. As a result, the retailer and the supplier need to divide the *expected* supply chain profits. This means that if the channel members do have any fairness concerns, *they may not only care about the actual profit that each receives but also the demand risk that each takes*. Further, instead of a 1/0 type of outcome, the retailer can influence the size of the pie by his order quantity. In our study, by ordering 50, the retailer keeps the flexibility of punishing the supplier without sacrificing all of his profit and in the meantime fully protects himself against the demand risk.

In a word, our game design and protocol reflect our need to elicit the strategic considerations between the two supply chain members. A detailed description of the game instruction with sample screen shots of our software interface is provided in Appendix B.

2.2 Research Hypothesis

Recall that in the Supplier Game only the suppliers were human players and the retailers had been pre-programmed in a way that would always place the expected-profit-maximizing order quantity if the contract allows him a positive expected profit (as the theory assumes). The human supplier knew all that in advance and therefore would have acted upon self-interest because their counter-partners, the computerized retailers, would not have any fairness concern at all.

In this study, we replace the artificial retailer with the human player and limit his ordering decisions. These are not fundamental changes in structure, and the theoretical predictions we discussed in Section 2 of Chapter I thus remain the same. If, however, the contracting theory does miss the role of fairness concerns, results from the two-person Contracting Game will deviate from those observed in the Supplier Game. Table 2.1 summarizes the design and sample sizes collected in this study compared with the Supplier Game.

DHIGH Demand Uniform (50, 150)	Contract Format		
	WP	BB	RS
Contracting Game (Two-person)	15	15	16
Supplier Game (One-person)	39	19	20

Table 2.1: Experimental Design and Sample Sizes of the Contracting Game[5]

According to the theory reviewed in Chapter I, the double marginalization (DM) will occur in the wholesale price contract. Given our DHIGH demand condition, the supplier will set a wholesale price of 10.5. In the two coordinating contracts, there are multiple pairs of contract parameters that maximize the total supply chain profits. Moreover, the retailer should not reject any offer that gives him positive expected profit under any type of contract. Our first two hypotheses follow directly from the theory that both the retailer and the supplier behave in a purely economic way. And we have:

Hypothesis 1A (*Self-interested Retailer*): The retailer will select the expected-profit-maximizing order quantity (q^*) whenever the proposed offer gives him a positive expected profit. No rejections or orders of 50 should be observed.

Hypothesis 2A (*Self-interested Supplier*): The supplier will set a wholesale price of 10.5 in the WP treatment; and under the buyback and revenue-sharing contracts, the supplier will choose contract parameters that can induce the retailer to order the first-best solution (q = 125) and give the supplier almost all of the supply chain profits.

[5] Outliers found in the contracting game are not counted in the table. See section 3 for details.

According to the results from the Supplier Game, the suppliers were able to find the optimal wholesale price in the WP treatment except for some deviations in the early rounds. In the buyback contract, the suppliers did not offer a rebate high enough to achieve coordination. In the revenue-sharing contract, they did not offer a wholesale price low enough to align the channel. Behaviors in the two coordinating contracts seemed to be in line with risk-aversion. So besides the point estimates indicated by the theory, the results from the Supplier Game might as well serve as our benchmarks.

If the supplier takes the first-mover advantage as the theory predicts, it can result in an extreme payoff difference between the two channel members with the supplier taking most of the channel profits. If the retailer views such outcome as unfair, rejections or lower order quantities of 50 will appear. If the supplier takes the retailer's concern for fairness into account, he may also behave differently in order to guarantee a successful channel relationship. Especially in the wholesale price contract, the retailer is forced to take all the demand risk and is subject to possible monetary loss due to the contract format and the stochastic demand. We hence speculate that the supplier may decrease his wholesale price to compensate the retailer at his own cost. Consequently, the retailer will choose to order an optimal quantity that is higher than the DM prediction, and so the overall supply chain will achieve higher efficiency. In the two coordinating contracts, the opportunity to share the demand risk is available. Although the supplier tends to behave risk-averse in the Supplier Game, we hypothesize that if the retailer has fairness concern over the demand risk he will push the supplier to assume the appropriate amount of risk so that full channel coordination can be realized.

In order to provide some benchmarks for possible fairness-minded behaviors, we apply the ERC model developed by Bolton and Ockenfels (2000, pp. 173) to our two-player game settings. In this model, the players are maximizing some utility function that consists of two weighted terms:

$$u_i(y_i, \sigma_i) = a_i y_i - \frac{b_i}{2}(\sigma_i - \frac{1}{2})^2, \ \ a_i \geq 0 \quad b_i > 0$$

The first component expresses the individual's preference over the standard pecuniary payoff of his own (y_i). The second term measures the distance between the player's relative standing σ_i (i.e., one's profit share in the game) and the *social reference point* of equal division (i.e., 1/2 in a two-person game). The minus sign in front of the second term shows that the further the allocation moves from player *i* receiving an equal split, the higher the disutility from this comparative effect. *a* and *b* are weights that a participant attributes to the pecuniary and relative components correspondingly. According to Bolton and Ockenfels (2000), strict narrow self-interest is the limiting case when the ratio of *a/b* reaches infinity, which will result in the predictions we presented in Hypotheses 1A and 2A. In contrast, *strict relativism* characterizes an individual's pure preference over his relative payoff and is represented by *a/b* = 0. It implies, in our game, that the supplier will offer a contract that divides the supply chain profits equally between the two parties; and the retailer will reject any offer that leads his (expected) profit share to below 50%.

Setting the ratio of *a/b* to zero and solving the ERC model numerically for our parameters, we obtain point estimates for strictly relative-payoff-concerned behaviors. Namely, the supplier will set a wholesale price of 6.92 in the WP treatment; will choose a wholesale price of 7.5 and a rebate of 6 in the BB treatment; and will offer a wholesale price of 1.5 with a revenue share of 6 in the RS condition. And the retailer will reject an offer if his expected profit share is lower than 50% in all three conditions. Given a wholesale price of 6.92, the supplier will receive a *fixed* payoff equal to the *expected* payoff of the retailer, and the efficiency of the wholesale price contract will be greater than the DM prediction. Using the above-mentioned pairs of parameters in the two coordinating contracts, the retailer and the supplier will share the production cost and sales revenue of the channel (and thus the demand risk associated) equally. Moreover, the retailer will order the first-best solution and the supply chain coordination will be completely reached.

However, the population is a mix of both types of self-interested and fairness-concerned decision makers as experimental results in the economic literature suggest. Thus, due to this heterogeneity, we do not expect the exact point estimates for fairness-minded behaviors to be realized in our experiments but rather the average performance of the participants to be along the *directions* for fairness concerns. And based upon the ultimatum game results reviewed previously, we have the following competing hypotheses:

Hypothesis 1B (*Fairness-concerned Retailer*): The retailer will reject an offer if it leads his expected profit share to be lower than 30% in all three conditions.

Hypothesis 2B (*Fairness-concerned Supplier*): The supplier will set a wholesale price between 6.92 and 10.5 in the WP treatment, and will choose pairs of contract parameters that allow him 50% to 60% of the *expected* channel profits in the buyback and revenue-sharing contracts. Overall, the supply chain efficiency will be higher with fairness-concerned channel members than with self-interested ones.

The next hypotheses speak to a fact that is independent of any fairness concern. Namely, coordinating contracts should perform better than the simple linear contract, and the buyback and the revenue-sharing contracts are mathematically equivalent.

Hypothesis 3 (Coordination): The average order quantity under the buyback and the revenue-sharing contracts should be higher than the one under the wholesale price contract, as will be the overall channel efficiency.

Hypothesis 4 (Equivalence): Retailers' average orders, profit division between the two parties and the supply chain efficiency in the two coordinating contracts will be identical.

2.3 Implementation

Upon arrival, each participant was assigned to play either as a retailer or as a supplier, and was randomly matched with another participant who took the other role. It was made publicly known that they would play with the same partner for the entire 100 rounds anonymously. Each subject participated in only *one* of the treatments in this study, and none of them had ever played in the experiments conducted in Chapter I before. Since participants have the tendency to anchor their decisions on their previous game experience, this *between-subjects* design controls for any possible order effects. Further, the same sequence of random demand draws was used in all three treatments for comparison purpose. After the game, subjects voluntarily filled out an online survey that asks what strategies they used in the game.

A total of 98 subjects participated in this study. Each experimental session lasted for approximately 90 minutes and average earnings, including a $5 participation fee, were $23. All sessions were conducted at the Laboratory for Economic Management and Auctions (LEMA) at Penn State University's Smeal College of Business during the summer of 2006. The participants were Penn State students, mostly undergraduates, from a variety of majors, recruited through a web-based recruitment system, with cash being the only incentive offered. The software we used was web-based and was built using PHP and mySQL.

3. Results

We treat each pair of retailers and suppliers as an independent observation and have collected 16, 16 and 15 sample teams respectively in WP, BB and RS conditions. We use the Grubbs procedure (see Grubbs 1969) to test for outliers in the data. One outlier team in each treatment is found and excluded from the main text of the statistical analysis and discussion.

3.1 Aggregate Treatment Effect and Hypothesis Testing

Table 2.2 presents the sample means and the standard deviations of the performance of the retailers, the suppliers and the overall supply chains in three different contracts, averaged over 100 periods. We also list the theoretical predictions and results from the Supplier Game for each corresponding condition as benchmarks in the same table. Statistical results on pair-wise treatment comparisons are summarized in Table 2.3. We use *t* tests with one-sample mean to compare with theoretical point estimates. For the rest of the comparisons, *t* tests with unequal variance are used. All p-values reported are 2-sided and the significance level is set at 5%.

- *Result 1*: *Rejections and orders of the minimum demand exist in all three contracts.*

Figure 2.2 shows the average percentage of both types of behavior observed in three treatments. As the graph indicates, the buyback contract seems to have the lowest percentages of both rejections and orders of 50, followed by the revenue-sharing contract, and then the wholesale price contract. The *t* test results further reveal that the percentages of rejections and orders of 50 are all significant from zero, except that the rejection rate of 3.47% in the BB treatment is weakly significant (p-value = 0.0987). Hence, Hypothesis 1A is rejected.

	Contracting Game			Supplier Game			Theory Prediction	
WP	Retailer	Supplier: $w \mid q$		Retailer	Supplier		Retailer	Supplier
q=0 (%)	9.00%	8.80						
	(0.06)	(2.46)						
q=50 (%)	23.36%	8.01						
	(0.15)	(0.87)						
q^*	86.49	7.62		66.87	10.01		62.5	10.5
	(6.35)	(0.76)		(6.12)	(0.74)			
Profit | q^*	313.01	393.22						
	(66.40)	(34.90)						
Efficiency | q^* (%)	89.17%							
	(0.04)							
Overall Profit	263.96	327.91		121.32	458.30		83.13	468.75
	(57.72)	(47.82)		(56.67)	(25.35)			
Overall Efficiency (%)	74.73%			74.21%			69.68%	
	(0.08)			(0.04)				
BB	Retailer	Supplier: $(w, b) \mid q$		Retailer	Supplier: $(w, b) \mid q$		Retailer	Supplier
q=0 (%)	3.47%	8.52	6.24					
	(0.08)	(0.72)	(1.69)					
q=50 (%)	7.67%	8.32	6.32					
	(0.10)	(0.67)	(1.04)					
q^*	121.35	8.12	6.40	88.95	10.27	6.40	125	
	(12.69)	(0.48)	(0.91)	(20.19)	(0.74)	(2.83)		
Profit | q^*	332.08	441.87						
	(49.35)	(40.45)						
Efficiency | q^* (%)	97.72%							
	(0.02)							
Overall Profit	307.27	412.52		124.88	558.12		≈ 0	≈ 792
	(45.79)	(60.02)		(60.08)	(57.79)			
Overall Efficiency (%)	90.88%			86.24%			100%	
	(0.08)			(0.08)				
RS	Retailer	Supplier: $(w, r) \mid q$		Retailer	Supplier: $(w, r) \mid q$		Retailer	Supplier
q=0 (%)	4.00%	2.29	6.96					
	(0.06)	(1.79)	(1.83)					
q=50 (%)	14.13%	3.04	5.16					
	(0.10)	(1.67)	(2.10)					
q^*	116.34	2.33	5.58	97.31	2.65	7.23	125	
	(8.57)	(0.98)	(1.50)	(10.96)	(1.07)	(1.57)		
Profit | q^*	341.66	434.88						
	(61.79)	(62.11)						
Efficiency | q^* (%)	98.05%							
	(0.01)							
Overall Profit	309.89	389.44		164.49	563.06		≈ 0	≈ 792
	(68.45)	(33.82)		(59.60)	(47.98)			
Overall Efficiency (%)	88.30%			91.86%			100%	
	(0.07)			(0.04)				

Table 2.2: Contracting Game Results

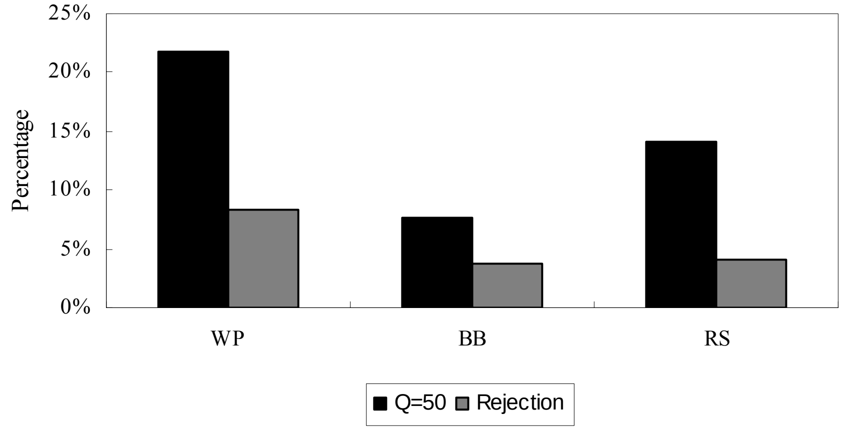


Figure 2.2: Average Percentage of Rejections and Orders of 50 in the Three Contracts

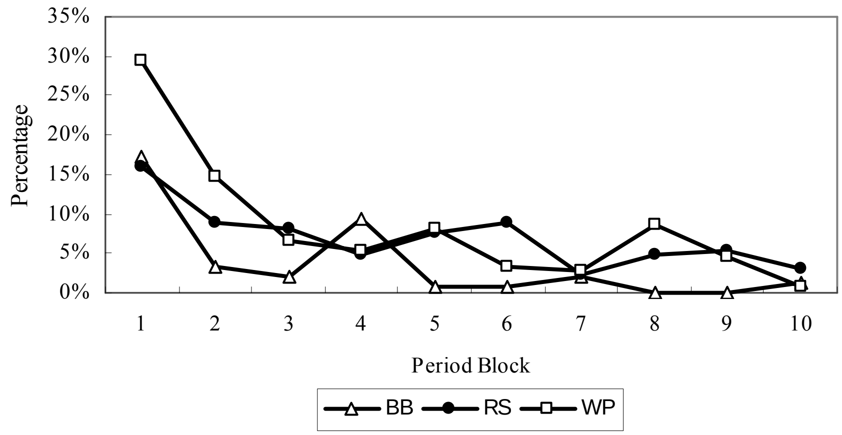


Figure 2.3: Rejection Rate Over Time in the Three Contracts

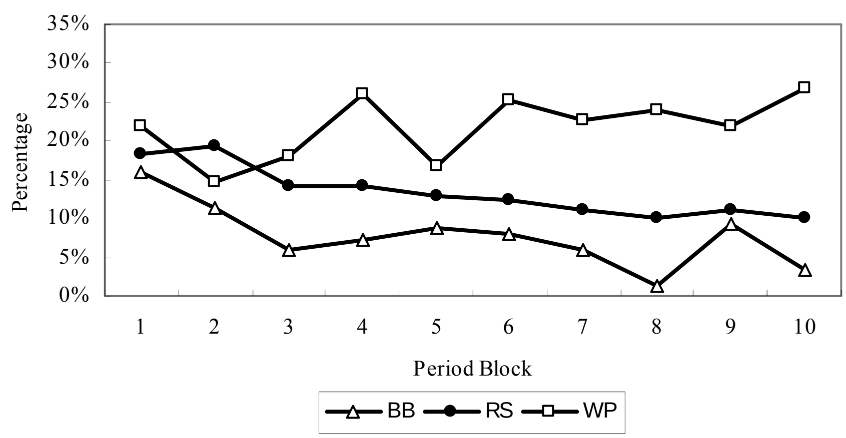


Figure 2.4: Percentage of Orders of 50 Over Time in the Three Contracts

To investigate the retailers' behaviors over time, we plot in Figure 2.3 the average rejection rate and in Figure 2.4 the average percentage of orders of 50 under three conditions, aggregated into 10-period blocks to improve exposition. As the graphs show, time trends seem to be apparent. We perform some logistic regression analysis to formally tests for the trends, and the results are discussed in the next section.

In Table 2.2, we also report the average decisions (i.e., w in WP, (w, b) in BB and (w, r) in RS that the supplier makes conditional on his paired retailer's corresponding actions ($q = 0$ for rejections, $q = 50$ for the minimum demand, and q^* for the optimal quantity). Note that offers being rejected or responded to by orders of 50 are generally those with relatively higher wholesale prices, lower rebates or higher revenue-shares. Since the theory and our Supplier Game results provide clear benchmarks for channels in which the retailer always chooses the best-reply solution, we will mainly focus on contract offers that induce the optimal order quantity (q^*) in the following analysis.

- *Result 2: Suppliers in WP choose a wholesale price that is significantly lower than the double marginalization prediction.*

Under the wholesale price contract, the overall average wholesale price the supplier offers is 7.75, whereas the average price that is accepted by q^* of the retailer is 7.62, both significantly below the self-interested benchmark of 10.5. As a result, the efficiency of the contract is improved, yet the supplier's profit share is reduced significantly[6].

Figure 2.5 shows the wholesale price the supplier proposes over time in the two-person WP treatment. In the same figure we also plot the self-interested prediction, the fairness-

[6] As mentioned in Chapter I, supply chain efficiency is calculated as total supply chain profits realized, divided by what could have been achieved if the retailer ordered the first-best solution given our random demand draws. The supplier's profit share is calculated as profits that the supplier received, divided by the total channel profits realized.

concerned prediction and results from the Supplier Game. We see that deviations from the benchmarks are quite dramatic at *initial*. Additionally, the average wholesale price that the suppliers proposed increased quickly when the retailers were automated to be self-interested but it decreased gradually when the retailers are human players.

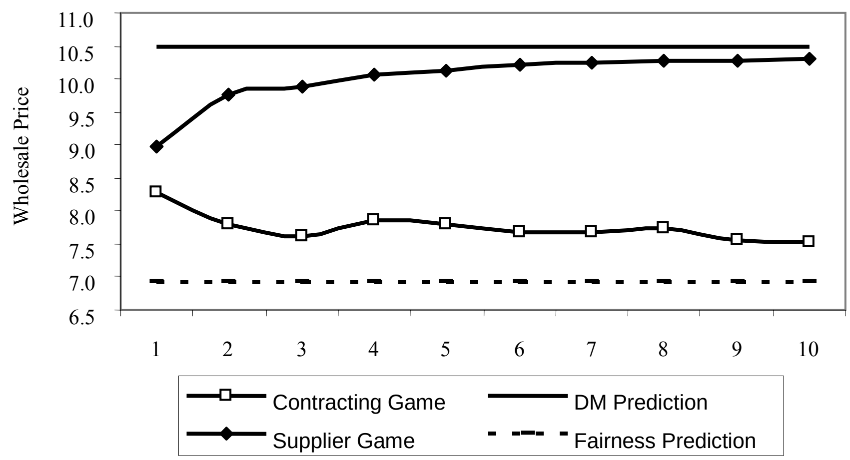


Figure 2.5: Average wholesale Price Over Time in the Contracting Game

From Figure 2.5, we can see that the average wholesale price that the retailer responds to q^* (7.62) is closer to predictions for fairness-concerned behaviors than for self-interested ones. Even though such a wholesale price is significantly higher than the benchmark of 6.92, it is insignificantly different from 7.5, the middle point between the retail price and the production cost of the game (p-value = 0.5507). The observation that the average wholesale price anchors to the middle point is in line with what Keser and Paleologo (2004) reported. At 7.5, the retailer's optimal order quantity is 87.5. If the retailer were able to sell all these units, he would obtain an expected profit that is the same as the fixed payment his supplier receives – a roughly equal outcome. The comparison with the Supplier Game designed in Chapter I allows us to conclude that it is the supplier's strategic consideration of fairness that drives him to stay at a low wholesale price throughout the game.

	Contract Theory		Supplier Game		BB		RS	
WP	T-stat	P-value	T-stat	P-value	T-stat	P-value	T-stat	P-value
q=0 (%)	5.35	0.0001			-1.96	0.0600	-2.10	0.0449
q=50 (%)	5.41	0.0001			-2.98	0.0066	-1.60	0.1227
Supplier's Profit (%)	14.65	< 0.0001	10.27	0.0001	0.63	0.5336	0.18	0.8593
Overall Efficiency (%)	2.40	0.0308	-0.24	0.8152	5.50	< 0.0001	4.96	< 0.0001
q^*	14.62	< 0.0001	-10.27	< 0.0001	9.51	< 0.0001	11.07	< 0.0001
$w \mid q^*$	14.62	< 0.0001	10.40	< 0.0001				
Supplier's Profit $\mid q^*$ (%)	14.83	< 0.0001	10.29	< 0.0001	0.47	0.6443	0.00	0.9996
Efficiency $\mid q^*$ (%)	17.79	< 0.0001	-11.65	< 0.0001	7.00	< 0.0001	7.71	< 0.0001
BB	T-stat	P-value	T-stat	P-value	T-stat	P-value	T-stat	P-value
q=0 (%)	1.77	0.0987					0.22	0.8267
q=50 (%)	3.08	0.0081					1.80	0.0831
Supplier's Profit (%)	28.29	< 0.0001	10.74	< 0.0001			-0.44	0.6657
Overall Efficiency (%)	4.45	0.0006	-1.69	0.1024			-0.96	0.3466
q^*	1.11	0.2844	-5.71	< 0.0001			-1.28	0.2127
$(w-b) \mid q^*$			3.31	0.0031			1.79	0.0834
$b \mid q^*$			0.01	0.9914			-1.85	0.0765
Supplier's Profit $\mid q^*$ (%)	28.18	< 0.0001	10.74	< 0.0001			-0.46	0.6471
Efficiency $\mid q^*$ (%)	4.22	0.0009	-5.98	< 0.0001			0.51	0.6155
RS	T-stat	P-value	T-stat	P-value	T-stat	P-value	T-stat	P-value
q=0 (%)	2.88	0.0114						
q=50 (%)	5.43	0.0001						
Supplier's Profit (%)	22.37	< 0.0001	8.21	< 0.0001				
Overall Efficiency (%)	6.68	< 0.0001	1.80	0.0858				
q^*	4.04	0.0011	-5.85	< 0.0001				
$w \mid q^*$			0.91	0.3688				
$r \mid q^*$			3.23	0.0028				
Supplier's Profit $\mid q^*$ (%)	22.16	< 0.0001	8.20	< 0.0001				
Efficiency $\mid q^*$ (%)	5.46	0.0001	-6.19	< 0.0001				

Table 2.3: The t Test Results on the Treatment Comparisons

To measure how close the actual contracts offered are to some coordinating contracts, we again compute and report in Table 2.4 the average coordinating rebate given the actual wholesale price in the buyback condition ($b^* \mid w$) and the average coordinating wholesale price given the actual revenue share in the revenue-sharing condition ($w^* \mid r$). Results of the t test comparing these contract parameters with various benchmarks are presented in the last column of the same table.

- *Result 3: The performance of the buyback and revenue-sharing contracts is close to coordination and better than the corresponding treatments in the Supplier Game.*

In the buyback contract, the average wholesale price and the rebate the supplier sets are not significantly different from the fairness anchoring point of (7.5, 6). The average optimal order quantity that the retailer accepts is 121.35, which is not statistically different from the first-best solution of 125. In the revenue-sharing contract, the average wholesale price is higher than the fair split estimate of 1.5, yet the average revenue share is insignificant from 6, and the retailer's average order still reaches 116.34. Further, the difference between the actual rebate and the coordinating rebate ($b - b^*|w$) under the BB condition of the two-person game is significantly smaller than that in the Supplier Game. So is the difference between the actual wholesale price and the coordinating wholesale price ($w - w^*|r$) under the RS condition. The results of our data analysis seem to imply that the existence of retailers who have some sort of bargaining power (e.g., the option to reject) may help push the channel towards coordination.

To conclude, we have observed evidence (Result 2 and 3) consistent with Hypothesis 2B, and Hypothesis 2A is thus rejected. The rest of our analysis focuses on comparisons among the three treatments in the Contracting Game. We start by plotting the average optimal orders that retailers place over time in each contract in Figure 2.6.

	Contracting Game		Supplier Game		*t* test		
Conditional on q^*	BB	RS	BB	RS	Hypothesis	T-stat	P-value
w	8.12	2.33	10.27	2.65	$w_{RS} = 1.5$	3.41	0.0039
	(0.48)	(0.98)	(0.74)	(1.07)	$w_{BB} - b = 1.5$	0.98	0.3436
$w - b$	1.73				$b = 6$	1.68	0.1155
	(0.90)				$r = 6$	1.13	0.2771
b or r	6.40	5.58	6.40	7.23	$b - b^*\|w = 0$	1.78	0.0975
	(0.91)	(1.50)	(2.83)	(1.57)	$w - w^*\|r = 0$	-4.40	0.0005
$b^*\|w$ or $w^*\|r$	6.83	1.61	9.69	1.19	Contracting vs. Supplier		
	(0.64)	(0.37)	(0.98)	(0.39)	Game: $b - b^*\|w$	-5.76	< 0.0001
$b - b^*\|w$	0.44		-3.29		Contracting vs. Supplier		
	(0.95)		(2.61)		Game: $w - w^*\|r$	3.12	0.0037
$w - b^*\|r$		0.73		1.45	Supplier's Share in BB		
		(0.66)		(0.73)	= 0.5	4.71	0.0003
Supplier's Profit (%)	57.16%	56.00%	82.00%	77.56%	Supplier's Share in RS		
	(0.06)	(0.08)	(0.08)	(0.08)	= 0.5	3.02	0.0085

Table 2.4: Performance of Coordinating Contracts

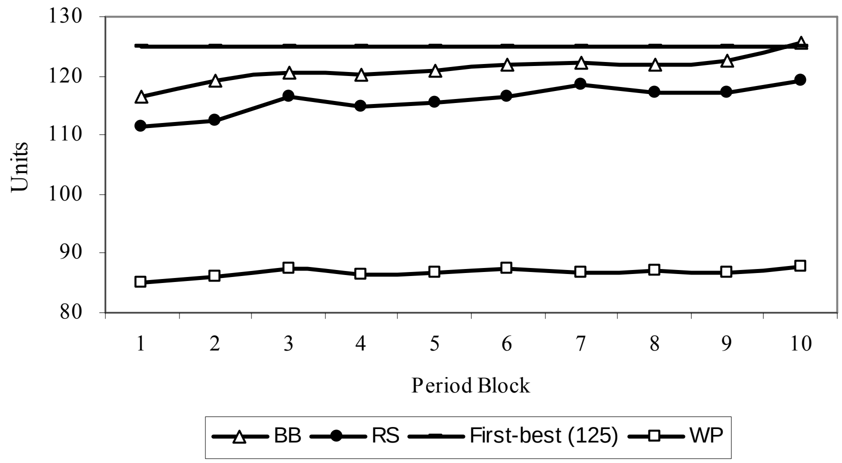


Figure 2.6: Average Order Quantities Over Time in Contracting Game

- *Result 4: The two coordinating contracts achieve higher efficiency than the wholesale price contract, but the supplier's profit share remains stable across the three treatments.*

From the graph in Figure 2.6, it is clear that the two coordinating contracts induce much higher orders than the wholesale price contract. The average profit share of the suppliers in WP, BB and RS are 56.00%, 57.16%, and 56.00%, respectively. These observations again are in line with the predictions on fairness-concerned suppliers of Hypothesis 2B. And quite interestingly, the average profit shares in the three contracts are not statistically different from each other, and they are all higher than the corresponding retailers' shares and the equal share of 50%, but they are all lower than those achieved in the Supplier Game.

Consequently, in each contract, the actual profit the suppliers/retailers received on average is significantly lower/higher than the corresponding ones in the Supplier Game (all p-values < 0.0001). Our results indicate that the suppliers when faced with human retailers lose part of their first-mover advantage but still manage to control a greater portion of the channel profit. The fact that the suppliers were unable to obtain a higher profit share through the coordinating contracts also suggests that both parties benefit proportionally from the coordination.

- *Result 5: The aggregate performance of the buyback and revenue-sharing contracts is equivalent, yet there is some subtle difference in how the two contracts are implemented.*

No statistical difference is found between the two coordinating contracts in terms of the retailer's order quantity, the supplier's profit share and the channel efficiency. As mentioned in Chapter I, the equivalence also implies that the difference between the wholesale price and the rebate in the buyback contract should equal the wholesale price in the revenue-sharing contract ($w_{BB} - b = w_{RS}$) and that the rebate has to equal the revenue share ($b = r$). However, average wholesale prices in the revenue-sharing contract are weakly higher than the average differences between the wholesale price and rebate in the buybacks (p-value = 0.0834), whereas the average revenue share is weakly lower than the average rebate (p-value = 0.0765), suggesting a slightly better performance of the buyback contract. Overall, we accept Hypotheses 3 and 4.

3.2 Regression Analysis

To formally examine how participants' decisions change dynamically, we perform a regression analysis for the retailers and suppliers separately. Recall that the retailers in our game have three ordering options. We therefore code the retailer's response with "1" for rejections, "2" for orders of 50, and "3" for expected-profit-maximizing orders and run a *multinomial logistic* model. We pick the category of optimal orders as the baseline reference and compare every other category with the baseline response. The model is defined as:

$$\log\left(\frac{p_{it}}{p_{3t}}\right) = \mu_i + \beta_1 t_i + \beta_2 D_{it-1} + \beta_3 F_{it} \qquad (2.1)$$

The dependent variable is the probability of preferring the *i*th (i = 1 or 2) category to the baseline category of optimal orders (i = 3) in each round. In other words, the model estimates the likelihood of deviations from the optimality. μ is the intercept and t refers to the period number from 1 to 100. D_{t-1} is the demand realization in the previous period. In order to check for any fairness-related response of the retailer, we measure the *inequality aversion* by F_t, the absolute difference between the retailer's *expected* profit share and the equal split of 1/2, if the retailer were to order the best-reply to the supplier's offer for the current round. F_t will range from 0 to 1/2, and the larger the score the more unequal the profit division. Our model does not control for any fixed effects. Because there are retailers who never reject or order 50 in our data, the maximum-likelihood estimates for those individuals will reach infinity and thereby bias the overall model.

As for the suppliers in the WP condition, we run an ordinary least squares (OLS) model with the wholesale price as the response variable. Under the BB and RS treatments, the suppliers need to make two decisions at the same time in every round. A *multivariate* regression model is used since we have more than one dependent variable which is correlated (Cohen et. al 2003). The model provides the least-square estimates of the joint linear effect of the set of predictors on the *set* of responses simultaneously. The general model is summarized as:

$$\begin{pmatrix} w_t \\ b_t \times BB \\ r_t \times RS \end{pmatrix} = s_i + \beta_1 t + \beta_2 D_{t-1} + \beta_3 QD_{t-1} + \beta_4 F_{t-1} \qquad (2.2)$$

BB and RS are indicator variables for the buyback and the revenue-sharing conditions respectively. The fixed effect for the supplier i is represented as s_i. To examine how the supplier adjusts to the retailer's response, we use QD_{t-1} to measure the difference between the optimal order and the actual order the retailer placed in the prior round (i.e., $q^*_{t-1} - q_{t-1}$), with 0 indicating that the retailer chose the optimal quantity. F_{t-1} measures the absolute difference between the retailer's realized profit share and 1/2 in the past period. The other variables and explanations remain unchanged as in model (2.1). We present regression estimates under three contracts in Table 2.5 for retailers and in Table 2.6 for suppliers. All our models are significant (p-values < 0.0001).

- *Result 6: The probability of preferring rejections decreases significantly over time in all three contracts; the probability of preferring orders of 50 decreases significantly over time in the BB and RS conditions but not in the WP treatment.*

The above result suggests that while successful relationships are managed to establish over time in all three contracts, retailers use the option of ordering 50 quite differently. And by examining the coefficient of D_{t-1}, we find that it may have something to do with the different amount of demand risk the retailer takes under each contract. The coefficient of D_{t-1} for the second category is significantly negative under the WP condition. It suggests that the retailer is using the option of ordering the minimum possible demand to avoid risk, and thereby the probability of ordering 50 is not decreasing over time in WP. The coefficients of D_{t-1} for the second category are insignificant in BB (p-value = 0.1344) and weakly positive in RS (p-value = 0.0968). It implies that orders of 50 are not much correlated to the random demand since portions of the risk are shared by the supplier in these two contracts. The coefficient of D_{t-1} for the first category also indicates that the probability of preferring rejections is independent of previous demand in each contract.

Variables	Level *i*	WP		BB		RS	
Chi-square		1557.82	**	847.77	**	1267.18	**
Intercept	1	-2.9164	**	-1.7172	**	-3.8695	**
		(0.4250)		(0.5085)		(0.5155)	
	2	-1.2975	**	-1.5284	**	-2.1803	**
		(0.2737)		(0.3859)		(0.3127)	
t	1	-0.0264	**	-0.0446	**	-0.0215	**
		(0.0040)		(0.0069)		(0.0053)	
	2	0.0021		-0.0154	**	-0.0086	**
		(0.0023)		(0.0036)		(0.0026)	
D_{t-1}	1	0.0049		-0.0024		-0.0020	
		(0.0032)		(0.0039)		(0.0041)	
	2	-0.0044	**	-0.0046		0.0041	*
		(0.0022)		(0.0031)		(0.0025)	
F_t	1	10.7727	**	4.1504	**	13.9303	**
		(1.0622)		(1.9355)		(1.1890)	
	2	4.4814	**	2.9522	**	4.8423	**
		(0.8026)		(1.4057)		(0.7731)	

** p-value < 0.05, * p-value < 0.1

Table 2.5: Regression Estimates for the Retailers' Decisions

- *Result 7: Retailers' orders are mostly affected by their perceptions of the fairness of the current contract.*

The F_t coefficients in all three conditions are positive and significant for both levels of the response variable, meaning that a contract that can potentially generate extreme difference between the two parties increases the likelihood of both rejections and orders of 50. This result together with Result 1 shows that although the retailers in our experiments do not strictly prefer the relative payoff, they do act in a fairness-concerned manner. Overall, we find some evidence supporting Hypothesis 1B.

- *Result 8: Suppliers respond strategically to the retailers' rejections and orders of 50 so as to maintain their first-mover advantage.*

The coefficients for *t* in the supplier's regression model show that over time the average wholesale prices decrease in WP and RS, while the rebate in BB and the revenue share in RS increase significantly. And the average wholesale price in WP is positively correlated with the previous demand, although a similar correlation is not observed in other conditions.

The most surprising result is in how the suppliers respond to the retailers' previous rejections or orders of 50. The coefficient estimates for QD_{t-1} in the three contracts suggest that when the retailers deviate from optimal orders in the last round, the suppliers significantly increase their wholesale price in WP and in BB, and weakly increase the revenue share in RS (p-value = 0.0932) – all in the direction that would enlarge the supplier's profit share. An opposite trend would have been more intuitive. Perhaps the following sample answers from our post game survey may shed light on how this happens:

> *If the retailer rejected my offer, I showed them who's boss and would raise my price per unit . . . they understand that they have to buy from me. Sort of a monopoly in this game, the strategy worked . . . .* — a supplier in the wholesale price contract.
>
> *. . . I didn't cave in because I knew he wouldn't stop if I did . . . .* — a supplier in the buyback contract.
>
> *. . . I would switch up a bit . . . . I wouldn't understand if my offer was rejected, because I think the offers were fair . . . .* — a supplier in the revenue-sharing contract.

It seems that the suppliers try being "tough" to retailers strategically so that they can maintain their first-mover advantage. Moreover, when the suppliers have already taken fairness into consideration in setting offers, but still receive rejections, they may punish the retailer for being too "greedy."

The last interesting observation is based on the estimates for F_{t-1} in Table 2.6. The coefficient is negative and significant in WP, suggesting that the suppliers reduce their wholesale price if previously realized profits are extremely unequal between the two parties. This result reveals that the suppliers are quite *sympathetic* to the retailers for taking all the demand risk, just as one supplier writes in the survey:

> *. . . When they profited negative, I reasoned and lowered my price so that they could return to the same profit as me!*

The corresponding estimates are insignificant in BB and RS, probably because the suppliers have already taken part of the risk. In summary, our regression results provide additional evidence on participants' concern over fairness.

	WP		BB				RS			
Variable	*w*		*w*		*b*		*w*		*r*	
F-value	141.27	**	106.41	**	98.02	**	67.80	**	106.05	**
R-square	0.6321		0.5640		0.5437		0.4991		0.5605	
Average Fixed Effect	7.8039	**	8.1223	**	5.8775	**	2.9945	**	4.9779	**
	(0.1061)		(0.0644)		(0.1302)		(0.1779)		(0.2123)	
t	-0.0045	**	0.0001		0.0068	**	-0.0109	**	0.0097	**
	(0.0006)		(0.0004)		(0.0008)		(0.0010)		(0.0012)	
Dt-1	0.0016	**	-0.0001		0.0010		-0.0006		0.0004	
	(0.0007)		(0.0004)		(0.0008)		(0.0011)		(0.0012)	
QDt-1	0.0038	**	0.0021	**	0.0010		-0.0008		0.0021	*
	(0.0007)		(0.0004)		(0.0009)		(0.0011)		(0.0013)	
Ft-1	-0.3116	**	0.1578		0.5639		0.0366		-0.1200	
	(0.1416)		(0.1717)		(0.3471)		(0.3824)		(0.4430)	

Table 2.6: Regression Estimates for the Suppliers' Decisions

4. Conclusion

We present a laboratory study designed to challenge the traditional assumption made in the supply chain contracting literature that channel members are self-interested. Our experiments extend the studies described in Chapter I to compare the performance of the wholesale price, buyback and revenue-sharing contracts in a setting that involves strategic interactions between two human players. We find that contrary to what the contracting theory assumes, decision makers do care about their relative payoffs when making decisions in our simulated Contracting Game. And being fair to each other has a crucial impact on maintaining a successful long-term relationship.

In the wholesale price contract, fairness concerns cause the supplier to stay at a low wholesale price. Consequently the overall efficiency is higher than predicted, even when rejections are counted in. This finding may help explain one phenomenon that the contracting theory has not been fully able to: despite the criticism the contracting theory makes of the wholesale price contract for creating the double marginalization, such a linear contract has been widely applied in many business settings: pharmaceutical industries and consumer goods industries (Cui et al. 2004), software industries (Robison 1994), and many others. Our results suggest that it may well be that channel member do not always act in a selfish way but care about their partners. The resulting reciprocal cooperation can help the supply chain stay reasonably efficient, and there may not be such a strong need to implement more complicated contracts that are difficult to administer.

In the two coordinating contracts, the comparison with results from the Supplier Game shows that fairness concern about how to divide the expected profits between the two parties helps improve the supply chain efficiency. Note that the Supplier Game is similar to the case where the supplier is the dominant player of a channel; whereas the two-person Contracting Game can be viewed as an approximation of a supply chain with relatively equal power players. Our result seems to imply that having a more powerful retailer (like Wal-mart) may be beneficial to the overall channel.

One way to formally test this conjecture is to introduce competition in future experiments. For example, we can run experiments with two retailers and one supplier or one retailer and two suppliers. In the present study, we only give the retailer options to punish the supplier for being unfair but no option to reward the supplier's fair behavior. Another direction for future research is to study a game in which the retailer can choose to order something larger than his expected-profit-maximizing solution.

CHAPTER III: LEARNING, COMMUNICATION AND THE BULLWHIP EFFECT

1. Introduction

Supply chain management is an example of a dynamic decision task that involves lagged feedbacks and multiple dependent decision makers. This task is known to be difficult for several reasons. According to Sterman (1989a), when decisions have indirect and delayed feedback effects decision-makers find it difficult to control the dynamics. Managing supply chains involves multiple agents whose performance depends on the quality of other supply chain members' decisions, and therefore is subject to coordination risk that may trigger instabilities in the system (Croson et al. 2004). One well-known source of inefficiency is the much studied *bullwhip effect*.

The bullwhip effect refers to the observation that the variability of orders in supply chains increases as one moves closer to the source of production. It was first noted by Forrester (1958), and has since been observed in many diverse settings. For example, Hewlett-Packard found that orders placed to the printer division by resellers have much bigger fluctuations than customer demands, and the orders to the company's integrated circuit division have even worse swings (Lee et al. 1997). A wide range of industries has experienced similar symptoms including computer memory chips (Fisher 1994), grocery (Fuller et al. 1993), and gasoline industry (Sterman 2000). The effect is costly because it causes excessive inventories, unsatisfactory customer service, and uncertain production planning. According to Lee et al. (2004), several industry studies such as ECR (Efficient Consumer Response) and EFR (Efficient Foodservice Response), report the bullwhip effect as most harmful to the efficiency of a supply chain.

Previous research on the bullwhip phenomenon thus focuses on understanding of its causes and ways to alleviate it. Two categories of explanations for the bullwhip effect have been advanced. Lee et al. (1997) identify four *operational* causes of the problem, including errors in demand signal processing, inventory rationing, order batching, and price variations, and recommend a number of operational strategies for dampening the effect.

The second category focuses on the *behavioral* causes of the effect. Behavioral causes are usually studied in the laboratory because it provides ways to eliminate operational causes, which is impossible to do in the field. The existence of the behavioral causes of the bullwhip effect in the laboratory has been demonstrated in a variety of settings and by many different researchers (see for example Sterman (1989a, 1989b), Kaminisky and Simchi-Levi (1999), Croson & Donohue (2003, 2004) and Croson et al. (2004)). These studies consistently show that participants do not adequately account for the time delays in making ordering decisions, and specifically, they tend to underweight the *supply line*, (orders placed but not yet received). Hence, the first behavioral explanation emphasizes the individuals' bounded rationality to make decisions in an environment with lagged, indirect and nonlinear feedbacks (Sterman 1989a). More recently, Croson et al. (2004) identifies another behavioral cause based on *coordination risk* — the uncertainty about the actions of others — and show that it often triggers instability.

The controlled environment of experiment also enables research to explore and isolate the impact of *institutional or structural* changes to the supply chain on mitigating the bullwhip behavior. Innovations such as reducing ordering and shipping delays (Steckel et al. 2004, Kaminisky and Simchi-Levi 1999), providing additional inventory information (Croson and Donohue 2004), sharing point-of sales information (Steckel et al. 2004, Croson and Donohue

2003), and adding excess inventory known as coordination stock to the system (Croson et al. 2004), all improve performance in the laboratory.

In this paper, we examine the impact of *training* and *communication* on reducing the bullwhip behavior in the laboratory. The idea is inspired by the unifying framework proposed by Boudreau et al. (2003), which calls for translation, experimentation and integration of operations and human resource management. They advocate that a successful interface between these two areas, in which human resource management provides behavioral insights for operation models and operations management provides contextual insights for the human resource models, will enhance the precision and rigor of research in both directions.

Our experiment involves the *Beer Distribution Game*—a serial supply chain with four links (see the next session for details). This game is popular in supply chain management classes and it has also been used extensively in research we cite above. We make two changes to the standard design in order to increase control: (1) we display to the participants their own outstanding orders (also called the *supply line*), and (2) in some of the sessions we have participants play the game twice, the first time to learn the rules and the dynamics of the game. One of our manipulations is the way the first game is structured: in a third of the sessions participants play in a specific role, the teams are reshuffled after the first game, but participants keep their roles in the second game. We call this training variation *role-specific training* since it allows participants role-constrained learning experience (March and Olsen 1975). In another third of the sessions each participant makes decisions for all roles in the supply chain (by literally walking between four computers placed in a row that represent the serial supply chain). We call this training variation *system-wide training* since it permits *systems thinking* (Checkland 1981, Senge 1990, Senge and Sterman 1992 and Jackson 1995). We also include sessions without the training game, as a benchmark.

The second manipulation in our experiment is communication. In half the sessions participants are allowed to communicate prior to the second game with the members of their team. The communication takes 10 minutes, and during that time the participants are asked to fill out a quiz testing their understanding of the game, and discuss their strategy for the up-coming game by themselves. In the other half of the sessions communication is not allowed, and participants are asked to sit quietly for 10 minutes, take the quiz individually, and reflect on the strategy for the upcoming game. This communication protocol controls for the amount of time participants are allowed to reflect on the game, so that any difference can be attributed to the availability of communication. In summary, we have a 3 x 2 design (the three types of training are none, role specific, and centralized, and the two communication protocols are with and without communication). Most of the six treatments involved eight teams of four participants, for the total of 192 participants. All sessions were conducted at the laboratory for Economic management and Auctions (LEMA) at Penn State, Smeal College of Business, between Fall 2003 and Fall 2004. Participants, mostly undergraduate business majors, were recruited using the on-line recruitment system, with cash the only incentive offered. Average earnings, including a $5 participation fee, were $22.

In the following sections, we present detailed experimental design and implementation (section 3.2), build up optimality benchmarks (section 3.3), report experimental results (section 3.4), and discuss managerial implications and conclude our study (section 3.5).

2. Experimental Design and Implementation

We follow the basic protocol of the "Beer Distribution Game" used in previous experimental studies. The game simulates a multi-echelon serial supply chain consisting of a *Retailer*, a *Wholesaler*, a *Distributor* and a *Factory/Manufacturer* with exogenous *Customer* demand. Each participant manages her own inventory by placing orders to the upstream supplier for replenishment so as to satisfy orders downstream over multiple periods. The decision task is complicated by the existence of lead-times/delays in the supply chain: *order processing delays* (two periods) and *shipment delays* (two periods) *or production delays* (three periods and only for the factory). Figure 3.1 (Croson et al. 2004) provides an illustration of the system.

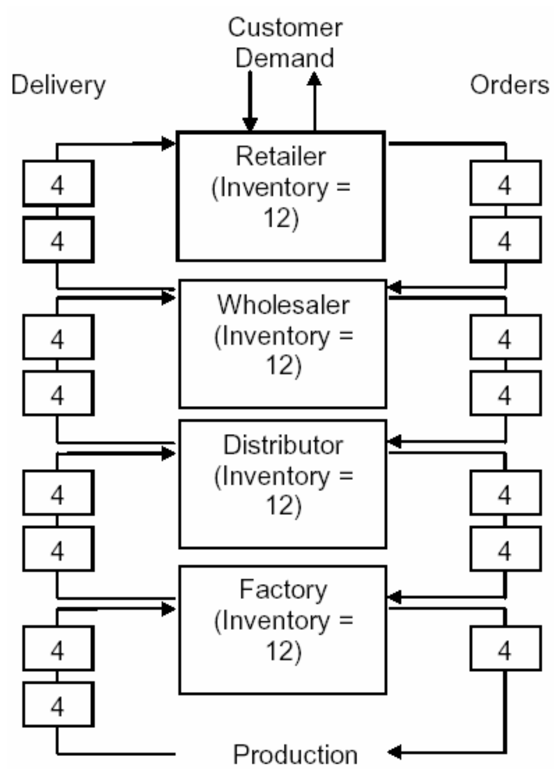


Figure 3.1: Supply Chain Settings in the Beer Game (Source: Croson et al. 2004)

Each period begins with the arrival of shipments from a participant's upstream supplier, which increases her inventory. Next orders placed by the downstream customer are received, which are either filled when inventory is available or become backlogged. Each participant then makes an ordering decision (variable of interest) and carries any remaining inventory/backlog over to the next period. See Croson and Donohue (2002) and Sterman (1989a) for further details of the game.

As in most previous work, we initialize orders and shipments in process to be 4 units, and the starting inventory at each echelon of the chain to be 12 units. Each team was given an initial endowment of 5000 tokens, and all participants were told they would incur inventory cost of 0.5 token per unit per week and backorder cost of 1 token per unit per week. Final team earnings in token were the difference between the initial endowment and the cumulative holding and backorder costs of all team members. At the end of the session the team earnings were converted to US dollars at a pre-determined exchange rate and split equally among the four team members. All experimental sessions last 48 periods/weeks, which is unknown to participants to avoid the end-of-game behavior. The game is conducted using the computer interface.

As mentioned in the introduction, our study differs from prior work in several aspects. First we provide participants with local *supply line* information in all treatments. We implement an automatic tracking system to update and display this information real time for participants with an interpretation of the concept in the instruction.

Second, a *training session* and a *study session* are added prior to the second game. The training session consists of 20 periods, in which participants either practice as teams to obtain role-specific experience, or train as a central planner making ordering decisions for four echelons sequentially. Results from training session are shown to the participants but do not affect their final earnings. The study session that followed lasts 10 minutes, and participants are encouraged to study the instruction of the experiment, finish a quiz, and think of strategies to be used in the actual game either independently, or through team meetings.

Finally, as in Croson and Donohue (2003, 2004), we control the customer demand distribution to be uniformly distributed from 0 to 8 units, independently generated between periods by the same random number seed across all treatments, and commonly known to all subjects. This publicly announced stationary customer demand, unlike those used in Sterman (1989a) and Steckel et al. (2004), removes the last operational cause corresponding to demand signal processing in the experimental setting. Table 3.1 below summaries our experimental design.

Uniform Demand Supply Line Visible	**Training Protocols**		
Communication	**None**	**Role-specific**	**System-wide**
No	Treatment NN (8)	Treatment RN (8)	Treatment SN (8)
Yes	Treatment NC (8)	Treatment RC (9)	Treatment SC (7)

Table 3.1: Basic Design (sample sizes in parenthesis)

Subjects are randomly assigned to computer terminals, which determined their roles and teams in the game when arrived. Once seated, participants are instructed to the rules and settings of the game. Then the 20-period training session begins, followed by the 10-minute study session, in which a quiz was given. After subjects are informed of the answers to the quiz, the actual 48-period game starts in which participants keep their pre-assigned roles and no further communication in any form is allowed. Finally, subjects are asked to fill in a questionnaire before receiving payments.

3. Theoretical Benchmarks

Before we analyze the data to examine the effectiveness of our manipulations, it is necessary to determine the theoretical performance benchmarks in our settings. We have eliminated operational causes of the bullwhip effect identified by Lee et al. (1997), and we now review several theoretical studies of optimal inventory policy for such systems.

3.1 Decentralized Optimal Ordering Policy

Karlin and Scarf (1958) set up inventory models with stationary and known customer demands and lead times. They have shown that it is optimal for the decision maker to follow an order-up-to policy, which requires one to keep his system stock, current inventory plus any on orders (supply line) at some constant levels. This optimal system stock level S is:

$$\Phi^{L+1}(S) = \frac{C_p - (1-\alpha)C}{C_p + C_H}$$

Where Φ is the customer demand distribution, L is the lead times, α is the discount factor, and C, C_H, C_p are the ordering, holding and shortage cost per unit respectively. The formula means that the optimal stock should cover demands during the time it takes to receive the current order plus the period it takes until the next order is received.

In the context of our specific game, a Retailer (R) faces $\Phi \sim$ Uniform (0, 8), L = 4 (lead-time is fixed under the assumption that its upstream supplier would have ample stock), $C_H = 0.5$, $C_p = 1$, while all the other parameters are zeros. Therefore $\Phi^5(S) = 2/3$, and S is the point that covers the five-fold convolution of the uniform customer demand two-third of the time, which gives $S_R^* = 22$. If Wholesaler (W) assumes that his downstream customer R would follow this optimal ordering policy so as to keep his system stock at some constant level, W would expect the same uniform customer demand to be passed in each period, and therefore $S_W^* = 22$. Similar logic applies to Distributor (D) and Manufacturer (M) except that M has 3-period lead times, resulting $S_D^* = 22$ and $S_M^* = 18$. If all members were able to follow this policy, the variance of orders at each echelon would equal to the variance of customer order and no amplification should exist.

We simulated this optimal policy ($S_R^* = 22$, $S_W^* = 22$, $S_D^* = 22$ $S_M^* = 18$) for 48 periods using the actual game customer demand. Results from order variance generated by role are shown in Figure 3.2, where supply chain members place zero orders at start to adjust to the optimal levels, and then converge to pass the order they received.

Although the deriving of this decentralized optimal policy that minimizes local supply chain costs does not require any global information of the supply chain, it is not intuitive and must be made under the assumption that all members are rational. Indeed, we did not observe any precise "pass-on-order" behavior in our treatments.

3.2 A Centralized Ordering Policy

Chen (1999) constructs a team model in which the division managers share a common goal to optimize the supply chain performance with lead times and known stationary demand. We follow his analysis to set up a centralized ordering strategy.

Note that the holding and backlog costs are the same for each echelon in our supply chain setting. By Chen's results, it would be unnecessary to hold any inventory at higher echelons but to have ample stocks at the Retailer to cover the stochastic customer demand. We simulated some similar but simple behaviors that work as follows: R keeps its initial inventory of 12 units and orders the average customer demand of 4 every period, while W, D and M first eliminate their excess initial inventory of 12 by ordering zeros (3 periods for W, 6 periods for D and 9 periods for M) and subsequently order 4 each period. If all members were able to follow this strategy, after 12 periods, there would be no inventory at W, D, M, but only R keeps inventory and uses 4s in the pipeline to average out the customer demand. As a result, order variance would be zero for all supply chain members when the system reached steady state. Figure 2 summarizes the simulated benchmarks of order variance.

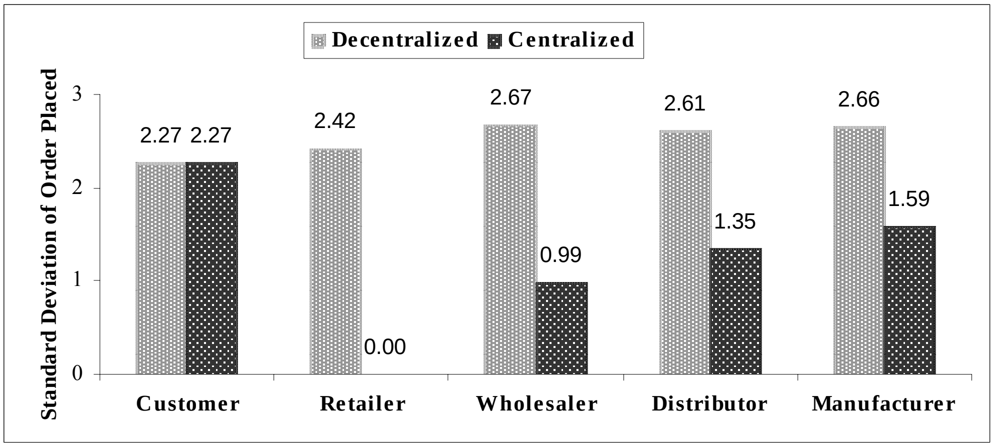


Figure 3.2: Decentralized and Centralized Benchmarks by Standard Deviation of Order Placed

This centralized ordering strategy is straightforward in a way that captures much of the initial status of the system (4s in the pipeline). However, in order to discover it, system understanding of the supply chain is required; in order to implement it successfully, coordination among chain members is also required. Quite excitingly, in the treatment with both system-wide training experience and coordination opportunities, 3 (out of 7) teams exhibit behaviors approaching this potential.

4. Experimental Results

We use the Grubbs procedure (see Croson et. al 2004 for more information) to test for outliers in the data we collected. One outlier was found and excluded from the main text of statistical analysis. The report of our results using the entire data set including the outlier is provided in Table 3.2.

4.1 Impact of Supply Line Visibility

Our data analysis will focus on standard deviation of subjects' ordering decisions in the experiments. Before we discuss the main results of our study, we make a methodological observation. Although underweighting the supply line has been cited as a major cause of the bullwhip effect, the supply line information is typically not available to the decision-makers directly (to be sure, decision makers have all the information required to compute it, but we do not know whether they do compute it, and if so, whether they compute it correctly). We provide participants with supply line information in all treatments, and compare our NN treatment (no training and no communication) with a similar treatment in Croson and Donohue 2004, to gage the effect of the supply line information.

Hypothesis 1: Direct available supply line information will (a) *remove order amplifications, and* (b) *decrease order variations.*

Figure 3.3 displays the experimental results of standard deviations of order for this treatment. The estimated median standard deviations of orders placed for each role are 2.07, 2.53, 2.97, and 3.43 for R, W, D and M respectively. The bullwhip effect seems to persist. We use a non-parametric sign test to examine order amplifications. If amplification exists, then the standard deviation of the i^{th} stage in the supply chain, σ_i, exceeds that of its immediate customer, σ_{i-1} (for all $i \in \{R, W, D, F\}$). If there were no amplifications, we would observe $\sigma_i > \sigma_{i-1}$ at the chance rate of 50%. Our data reveal that for 71% of the cases (17 out of 24), $\sigma_i > \sigma_{i-1}$, rejecting hypothesis 1 at p = 0.011. Hence, order amplification remains even when supply line information is visible.

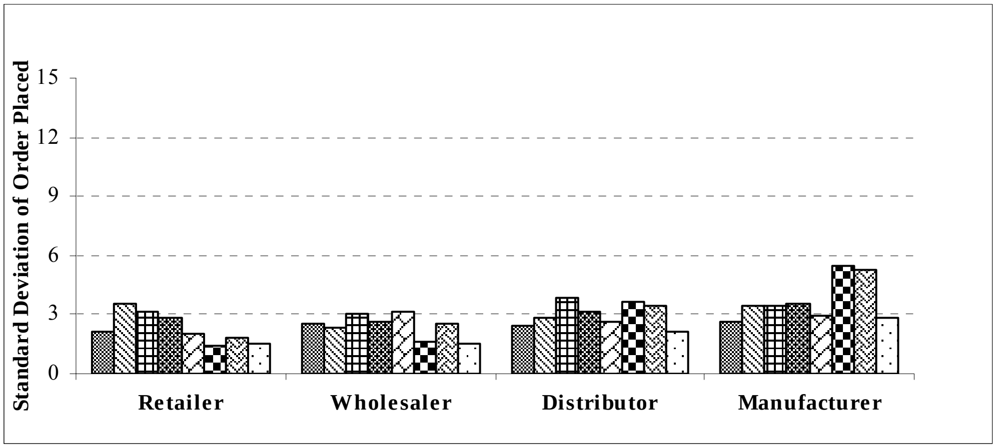


Figure 3.3: Standard Deviation of Order Placed by Role in Treatment NN

We then compare data in Treatment NN with that in the Baseline treatment of Croson and Donohue 2004, in which supply line information is hidden but other game settings are identical to ours. Figure 3.4 reports their data. We conduct a nonparametric Mann-Whitney U test to compare the variations of order placed and costs generated in these two experiments. The median standard deviations of orders in Treatment NN is 3.03, which is significantly less than the 4.20 in Croson and Donohue 2004 at $p < 0.0001$. We therefore cannot reject Hypothesis 1b that direct available supply line does have a stabilizing effect on the system.

Although making supply line visible decreases order variability, order amplifications and variations still remain, and it is not clear to us how subjects utilize their supply line information. In the post-survey, many subjects claimed that they did not find their supply line helpful and actually ignored it when placing orders. These comments may indicate that some participants do not understand how to effectively make use of the on order information, and giving the hands-on experience with the system may improve their understanding.

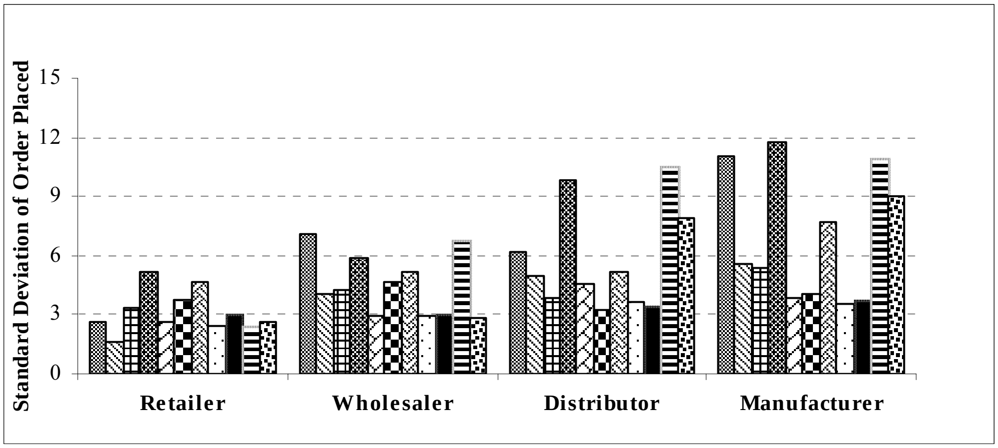


Figure 3.4: Standard Deviation of Order Placed by Role in Croson and Donohue (2004)

4.2 Effect of Training

A common way to improve the quality of decision-making is through training. Educating participants about the optimal ordering policy can reduce the bullwhip effects (see Croson et al. 2004). But education can be used only when ex-ante optimal strategies are known, whereas it is often the case in practice that the system is too complex, and changes too rapidly, to offer an optimal solution. In such situations, decision makers often rely on their ex-post experience to respond. In our treatment of training, we focus on this learning through experience.

4.2.1 Effect of Role-Specific Training

The supply chain setting we investigate is dynamic and inherently unstable. It is sensitive to human errors because a team with one member behaving erratically has little chance of performing well, even if the other three members understand the system. An early "unintentional" error by one person could cause the entire supply chain to get out of control. In most experimental studies involving the Beer Game reported in the literature, that we are aware of, participants play the game one time. This is a reasonable procedure because the studies investigate dynamics, and participants play one time in all treatments, so there is no priori reason to think that repeating the game (without any further intervention) should improve performance or account for a treatment effect. In our case, however, repeating the game is necessary because the type of training is one of our treatment variables. Additionally, our design provides a way to measure the extent to which the bullwhip behavior reported in the literature might be due to participants' misunderstanding of the game.

Treatment RN differs from Treatment NN in that participants play in a 20-period training session before the actual game. Participants keep the same role in both, the practice and the real game. We formally test the hypothesis that repeating the game improves performance: If the bullwhip behavior is induced by unsystematic human errors, order variations and amplifications should decrease in Treatment RN relative to NN.

Hypothesis 2: If the game is repeated (a) o*rder amplifications and* (b) *order fluctuations will decrease.*

Figure 3.5 reports the results of Treatment RN. Both order oscillations and amplifications clearly remain. The sign test suggests that order amplifications are highly significant: $\sigma_i > \sigma_{i-1}$ in 79% cases (19 out of 24) which is different from the 50% chance rate at $p < 0.001$, rejecting Hypothesis 2a. The median standard deviation of orders in RN is 3.72, which is not statistically different from that in NN by the Wilcoxon test (two-sided $p = 0.65$), so we reject both parts of Hypothesis 2.

So far we have demonstrated that the bullwhip behavior persists in the laboratory, in an environment that is more controlled than standard. We show that that role-specific training fails to induce learning effective enough to help alleviate misbehaviors. This experimental evidence is consistent with the observation that even highly experienced supply chain managers cannot avoid the bullwhip effect in practice.

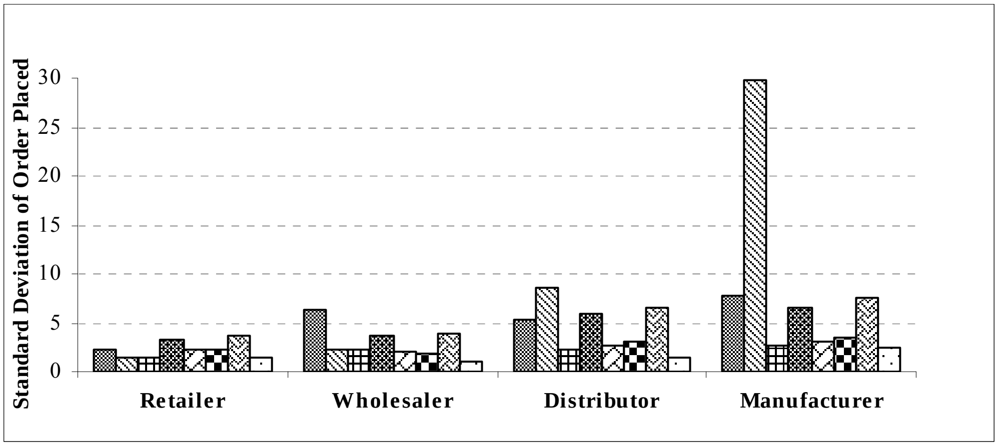


Figure 3.5: Standard Deviation of Order Placed by Role in Treatment RN

Our hypothetical explanation for this learning failure lies in participants' lack of direct feedbacks and "system knowledge". According to Sterman, "…learning from experience may be hindered by such misperception (Sterman 1989a, p.336)…" of origins of the dynamics due to delayed feedbacks. In the post-survey, most participants attributed their poor performance to the erratic orders they received that directly change their current inventory levels, yet few noticed that not taking a full account of their supply line in placing orders would affect their future inventory levels. In other words, subjects lose connection between their actions and the feedback they received because of the time lags in the system. Moreover, practicing only in specific roles may prevent participants from observing the interrelationship with other decision makers in the supply chain, and the long-term consequence of their own decisions thereafter (March and Olsen 1975, Senge 1990). For example, subjects in the experiments often failed to realize that unusually high orders they placed would likely knock their suppliers out of stocks, which will in turn make their own lead times unpredictable.

The fact that performance of the team depends so heavily on the decisions of each individual member makes the Beer Game difficult to understand and manage. We conjecture that the inability to see the system as a whole may be one reason for the poor performance, which leads us to our next treatment.

4.2.2 Effect of System-wide Training

The idea of system-wide training is common in practice. Managers in large corporations are often sent to work in different departments so as to better understand the system structure and interrelationship of the overall organization. Similarly, the research work by Senge and Sterman (1992) advocates that tools involving “learning laboratories” or “microworlds” where managers play roles in simulated organizations to experience the long-term, system-wide, dynamic consequences of decisions, will accelerate learning. A recent work by Hwang (2004) also argues that systems thinking should be incorporated into the design of training strategies. We apply the idea of “systems learning” to the Beer Game. Treatment SN is identical to Treatment RN, except that participants are not trained locally but as central planners for 20 periods, and we have the following hypothesis:

Hypothesis 3: System-wide training will (a) remove order amplifications, and (b) decrease order variations.

Figure 3.6 depicts the results of Treatment SN. Contrary to Hypothesis 3a, orders amplifications persist, $\sigma_i > \sigma_{i-1}$ in 83% cases (20 out of 24), and p = 0.0001. More surprisingly, contrary to Hypothesis 3b, the median standard deviation of 2.77 in SN is not significantly different from that in NN or in Treatment RN (p = 0.439, 0.221 respectively).

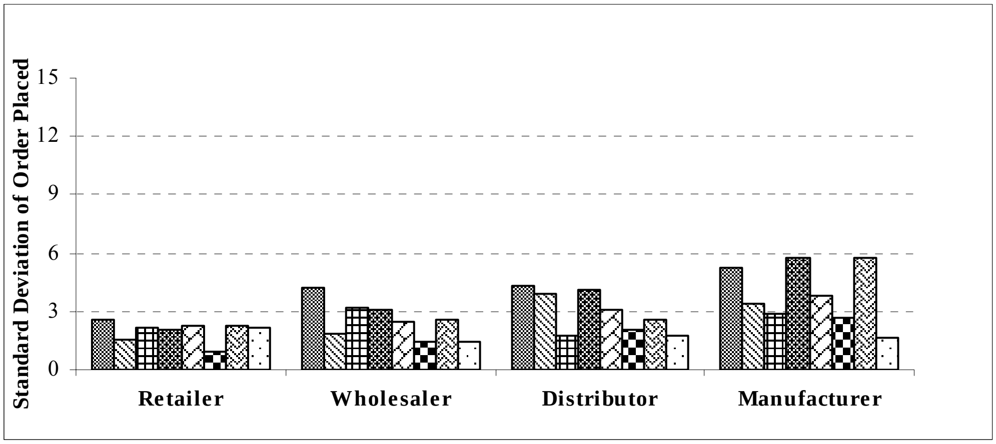


Figure 3.6: Standard Deviation of Order Placed by Role in Treatment SN

We undertook this treatment with strong expectation that we would be able to demonstrate significant behavioral improvement with system-wide training. Why do the data speak to the opposite conclusion? The post-game comments by subjects revealed that although many recognized the importance of smooth order flow in the system after system-wide training, they had to deviate from their original strategies in response to unanticipated teammates' misbehaviors (e.g., big orders from downstream or no shipments from upstream). Some participants reported regretting not being able to convey their ideas to the members of their team.

Based on the post-game comments, we speculate that system-wide training does improve system understanding for many individuals; however, since it is not equally effective for all individuals, and since the system is so susceptible to individual errors, the insights gained from the system-wide training cannot be effectively translated into improved performance. This observation is in line with the findings of Croson et al. (2004), that coordination risk, the lack of common knowledge about others' actions, triggers instability. The idea that coordination risk may be interfering with the improved performance leads us to the next set of treatments that involve communication.

4.3 Communication Effect and Its Interactions with Training

Cooperative learning, an approach that encourages group members to collaborate on a collective task through communication, has been demonstrated to successfully enhance learning in both education (Johnson, Johnson, and Smith 1991) and training literatures (Doyle 1991 and Schendel 1994), and in the following treatments we investigate its effect on the behavior in the Beer Game.

Treatments NC, RC and SC differ from their counterpart treatment NN, RN and SN only in terms of the availability of coordination opportunity. We control the form of communication to be a 10-minute team discussion (after training in RC and SC) prior to the game, in which group members are encouraged to study game materials, share training experience, and develop team strategies collectively.

We expect communication to reduce possible coordination risk in two ways. First, communication allows participants to exchange experiences and ideas, so that at the end of the discussion all team members have a similar level of understanding of the system. For instance, if any teammate recognizes the importance of supply line information on making ordering decisions, being able to share this idea with other members should help alleviate the overall tendency to underweight unfilled orders. Second, an explicit team strategy if developed successfully can be used to guarantee the knowledge of other teammates' actions. We conjecture that communication to help create *organizational rationality* (or referred as *organizational learning* in management literature, see Stata 1989 and Kim 1993) that copes with dynamic environment better despite of individuals' bounded rationalities. This leads to the following hypothesis:

Hypothesis 4: Communication, (a) *when combined with training, will improve performance (in terms of the decrease in order oscillations and amplifications), but communication* (b) *without training will have no effect on performance.*

Figures 3.7, 3.8 and 3.9 reflect data in Treatments NC, RC and SC. Team 8 in Treatment SC is detected as an outlier by the Grubbs procedure, and is therefore excluded from subsequent analysis. The basic statistics for these treatments and others are summarized in Table 3.2. Sign test results are summarized as well in Table 3.3, note that only in Treatment RC is the order amplification marginally significant. To test Hypothesis 4, we compare the variability of orders in NN and NC (the two treatments without training). The Wilcoxon test shows no significant variance decrease (one-sided p = 0.191), indicating that, consistent with our hypothesis, communication alone fails to alleviate the bullwhip effect. Next we compare RN with RC, as well as SN with SC, to examine the communication effect after training. Results show that communication after both types of training helps reduce order variability (weakly so for role-specific training, one-sided p = 0.084, and somewhat more significantly so for the system-wide training, one-sided p = 0.047). Overall, our data is consistent with Hypothesis 4.

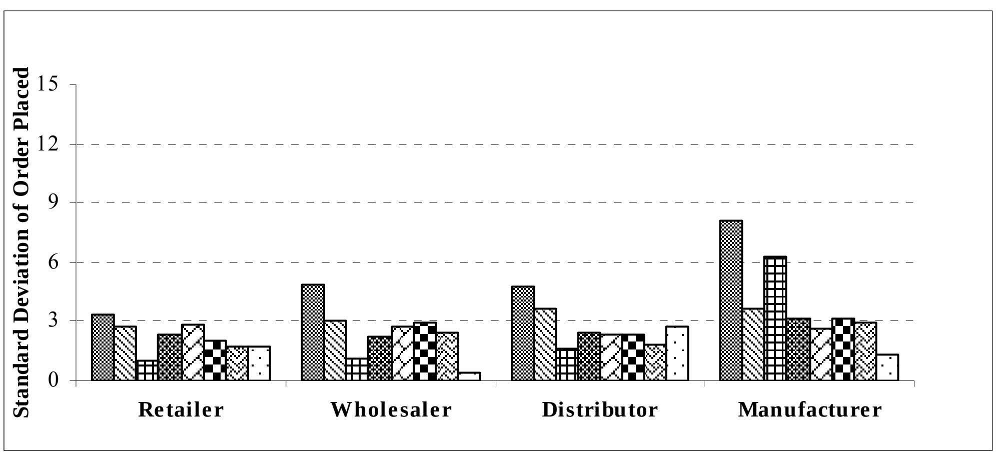


Figure 3.7: Standard Deviation of Order Placed by Role in Treatment NC

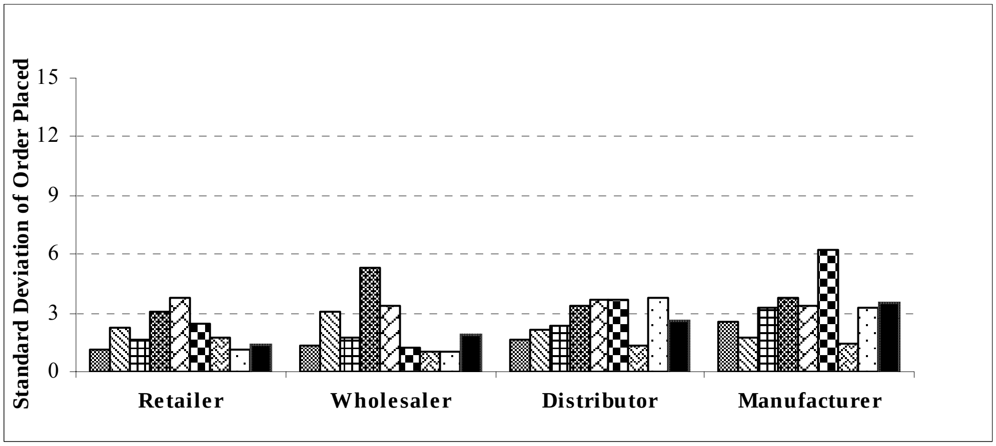


Figure 3.8: Standard Deviation of Order Placed by Role in Treatment RC

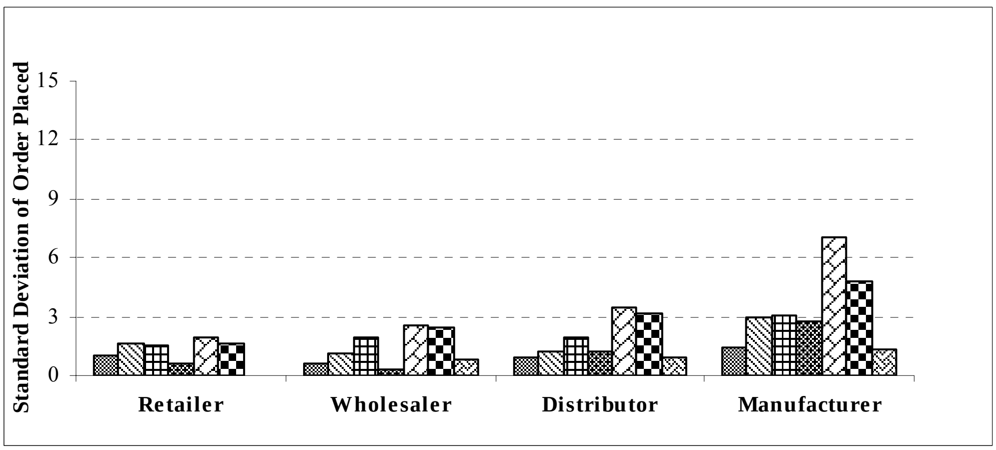


Figure 3.9: Standard Deviation of Order Placed by Role in Treatment SC

Summary Statistics: (Standard Deviation of Orders)	Treatment NN *(I)*	Treatment RN *(II*	Treatment SN *(III)*	Treatment NC *(IV)*	Treatment RC *(V)*	Treatment SC *(VI)*	Treatment Comparison
Median by Role	**Outlier Excluded (outlier included in parenthesis)** (One-sided Wilcoxon Test: ** P < 0.05, *P < 0.10)						
Retailer	2.072	2.221	2.133	2.162	1.744	1.561 (1.468)	1 > 2: 0.323 1 > 3: 0.439
Wholesaler	2.533	2.282	2.499	2.586	1.699	1.142 (1.627)	1 > 4: 0.191 1 > 5: 0.24
Distributor	2.967	4.133	2.852	2.397	2.695	1.271 (2.808)	1 > 6: 0.036** (0.134)
Factory	3.434	4.977	3.566	3.134	3.274	2.917 (4.434)	2 > 3: 0.221 2 > 4: 0.164
Median by Treatment	3.026	3.721	2.766	2.556	2.300	1.745 (2.584)	2 > 5: 0.084* 2 > 6: 0.014** (0.065*)
Average by Treatment	2.859	4.386	2.838	2.820	2.571	1.942 (2.584)	3 > 4: 0.253 3 > 5: 0.212 3 > 6: 0.047** (0.164)
Standard Deviation by Treatment	0.46	2.93	0.85	1.10	0.86	1.11 (2.09)	4 > 5: 0.245 4 > 6: 0.06* (0.191) 5 > 6: 0.071* (0.185)

Table 3.2: Basic Statistics and Oscillation Comparison Summary

Treatment	Overall Success Rate	Overall P-value for the Sign Test	Sign Test Comparison at Adjacent Roles		
			Retailer vs. Wholesaler	Wholesaler vs. Distributor	Distributor vs. Manufacturer
NN	**70.8%**	**0.011****	0.363	0.035	0.004
RN	**79.2%**	**0.001****	0.145	0.035	0.000
SN	**83.3%**	**0.000****	0.004	0.004	0.035
NC	**66.7%**	**0.032****	0.145	0.363	0.004
RC	**66.7%**	**0.076***	0.254	0.020	0.090
SC	**83.3%**	**0.000****	0.363	0.000	0.000
C&D (2004)	**81.8%**	**0.000****	0.000	0.113	0.006

Table 3.3: Sign Test Summary for Amplification Comparison

Our main conclusion from the above analysis is that the effectiveness of communication in reducing the bullwhip behavior depends on the training protocols that participants follow. Communication without training fails to correct misbehaviors, while with training it does help improve behavior, or in other words, it does serve to reduce coordination risk.

To better understand how communication with different training protocols affects participants' behaviors, we explore the strategies teams in the three communication treatments agreed upon. We classify strategies into three types, based on their effectiveness in improving performance, and present the summary in Table 3.4.

Category	Strategy Forms	Number of Teams			Range for STD of Orders
		NC	RC	SC	
1	Strict: constant order of 4s, 0 inventory at W, D and M	0	0	3	0.74 to 1.22
2	Flexible: range of order placed, target inventory levels	6	7	4	1.55 to 3.75
3	No agreement reached, or "cheap talk"	2	2	0	2.23 to 5.27

Table 3.4: Team Strategies in Communication Treatments

From Table 3.4, we see that Type 1 teams reach strategies that clearly and precisely set the inventory targets and ordering decisions per period for all group members, which effectively reduce the order variance. Yet this kind of teams only exists in Treatment SC where both system-wide training and communication are applied. Type 2 teams have relatively flexible policies that require teammates to control the inventory levels and orders placed within certain ranges, therefore order variances still remain but are reduced. The majority of teams, especially in Treatment NC and RC, belong to this type. The last type of teams (that does not occur in Treatment SC at all) fails to reach any effective policy and results in highly variable orders. A closer examination of the teams in SC also shows that 3 out of 7 teams are actually able to develop strategies approaching the centralized policy we discussed in Section 3.3.

Our data seem suggestive of the fact that system wide training might be more effective than role-specific training (although both types of training are clearly more effective than no training at all). To formally establish which type of training is more effective, we compare order variability in SC with that in NC and RC, and find slight improvement in both cases (one-sided $p = 0.06$, and 0.071 respectively).

Our data shows that training alone does not reduce the bullwhip behavior, but it does contribute to the accumulation of critical knowledge for the participants that leads to more effective discussions and team strategies. Thus, it is the *interaction* between training and communication that is important — both components must be present to improve performance. Figure 3.10 summarizes the overall across-treatment effects in our study.

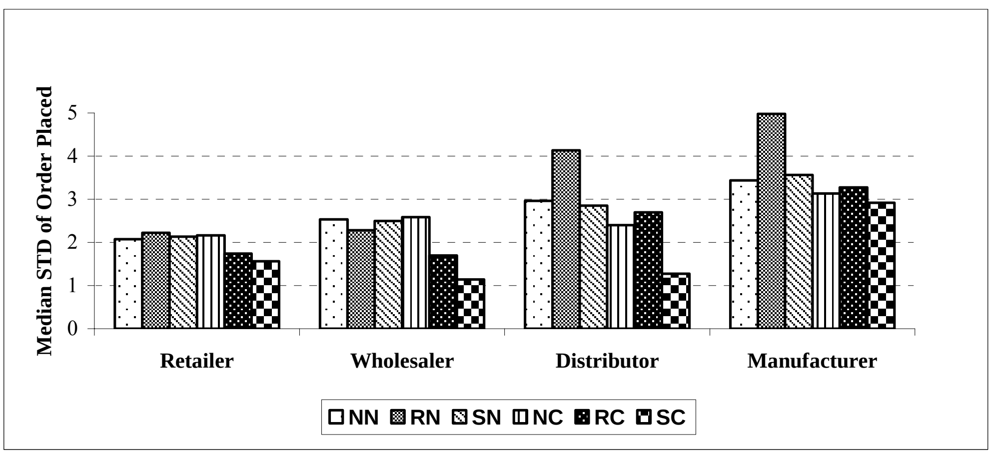


Figure 3.10: Across-Treatment Comparisons by Standard Deviation of Orders by Role

5. Conclusions and Managerial Implications

There are two major contributions that our study offers. The first one is methodological: we implement a set of laboratory controls on the Beer Game that is more rigorous than in previous studies, and we find that results reported in the literature are quite robust. Repeating the game, when not combined with communication, does not improve performance, and displaying the supply line information does decrease order variability, but far from eliminates the bullwhip effect.

The second contribution is managerial, and has potential bearing on the types of training and communication that goes on within individual organizations and within supply chains. We find that giving people hands-on experience with the system improves performance when it is combined with the opportunity to share knowledge and coordinate through communication. System-wide training appears to be marginally more effective than role-specific training, when the effectiveness is measures as the decrease in order variability. It appears to be even more effective, however, when the effectives is measured by examining the nature of the strategies — following system-wide training more teams are able to implement a policy that is close to centralized, and no team fails to implement a policy that is at least marginally effective.

CONCLUSION

So far we have observed three experimental studies that systematically examine several behavioral issues in supply chain management. The first two studies start with observations about the supply chain contracting theory and challenge its traditional assumptions about channel members' behaviors. The experimental evidence we provide shows how actual human behaviors deviate from theoretical predictions, implying the inadequacy of the existing assumptions in the contracting model.

The last study starts with a phenomenon well documented in the field, the bullwhip effect, and uses controlled laboratory settings to test whether human resource activities such as training and communication can help decision makers alleviate the bullwhip behavior. And our results suggest to practitioners that besides the efforts that should be made to reengineer the supply chain, programs should also be designed to improve supply chain managers' decision quality, so as to reduce the order variability and enhance the overall supply chain performance.

This dissertation makes two major contributions to the emerging field of behavioral operations management. First, it demonstrates the relevance and importance of behavioral factors to supply chain management. The first two experimental studies provide an empirical validation of the supply chain contracting theory. Several questionable assumptions are pointed out for future research to improve the practicality of the contracting model. The last study provides a "wind-tunnel" test (Bolton and Kwasnica 2002) of mechanisms to mitigate the bullwhip behavior of supply chain managers.

Second, in terms of methodology, this dissertation demonstrates that behavioral experiments, if properly designed and executed, can provide results and phenomena of interest to both operations researchers and managers. And besides the future work suggested in the conclusion of each chapter that would further examine behavioral issues in supply chain management, other sub-disciplines such as service management where a great deal of human interaction takes place are also worth investigating.

REFERENCES

Azoulay, P. and Shane, S. (2001), Entrepreneurs, Contracts, and the Failure of Young Firms, *Management Science*, 47(3), 337-358.

Balachandran, K. R., and Radhakrishna, S. (2005), Quality Implications of Warranties in a Supply Chain, *Management Science*, 51(8), 1266-1277.

Benzion, U., Cohen, Y, Peled, R. and Shavit, T. (2005), Decision-making and the Newsvendor Problem — An Experimental Study, Working Paper.

Bolton, Gary E. and Katok, E. (2006), Learning-by-doing in the Newsvendor Problem: A Laboratory Investigation, Penn State Working Paper.

Bolton, Gary E. and Kwasnica, Anthony M. eds. (2002), *Special Issue: Experimental Economics in Practice*, *Interfaces*, 32.

Bolton, Gary E. and Ockenfels, A. (2000), ERC, A Theory of Equity, Reciprocity and Competition, *American Economic Review*, 90(1), 166-193.

Bolton, Gary E. and Zwick, Rami (1995), Anonymity versus Punishment in Ultimatum Bargaining, *Games and Economic Behavior*, 10(1), 95-121.

Boudreau, J., Hopp, W., McClain, J, and Joseph Thomas (2003), On the Interface between Operations and Human Resources Management, *Manufacturing and Service Operations Management*, 5, 179-202.

Bendoly, E., Donohue, K. and Schultz, Kenneth L. (2006), Behavior in Operations Management: Assessing Recent Findings and Revisiting Old Assumptions, *Journal of Operations Management* (forthcoming).

Burnetas, A. and Ritchken, P. (2005), Option Pricing with Downward-sloping Demand Curves: The Case of Supply Chain Options, *Management Science*, 51(4), 566-580.

Cachon, G. (2003), Supply Chain Coordination with Contracts, *Handbooks in Operations Research and Management Science: Supply Chain Management*, edited by Steve Graves and Ton de Kok, North Holland.

Cachon, G. and M. Lariviere (2005), Supply Chain Coordination with Revenue Sharing Contracts, *Management* Science, 51(1), 30-44.

Cachon, G. (2004), The Allocation of Inventory Risk in a Supply Chain: Push, Pull, and Advance-Purchase Discount Contracts, *Management Sciences*, 50(2), 222-238.

Camerer, C.F. (2003), *Behavioral Game Theory Experiments in Strategic Interactions*, Princeton University Press, Princeton, NJ.

Checkland, P. (1981), *Systems Thinking, Systems Practice*, Wiley, Chichester.

Chen, F. (1999), Decentralized Supply Chains Subject to Information Delays, *Management Science*, 45, 1076-1090.

Choi, K., Dai, J. G. and Song, J. (2004), On Measuring Supplier Performance under Vendor-Managed-Inventory Programs in Capacitated Supply Chains, *Manufacturing & Service Operations Management*, 6(1), 53-72.

Clark, A. J. and Scarf, H. E. (1958), Optimal Policies for a Multi-Echelon Inventory Problem, *Management Science*, 6, 475-490.

Cohen, J., Cohen, P., West, Stephen G. and Aiken, Leona S. (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, NJ.

Corbett, C. J., Zhou, D., and Tang, C. S. (2004), Designing Supply Contracts: Contract Type and Information Asymmetry, *Management Science*, 50(4), 550-559.

Cox, J., Friedman D. and Gjerstad, S. (2004), A Tractable Model of Reciprocity and Fairness, Economics Working Paper Archive at WUSTL.

Croson, R. and Donohue, K. (2002), Experimental Economics and Supply Chain Management, *Interfaces*, 32, 74-82.

Croson, R. and Donohue, K. (2003), The Impact of POS Data Sharing on Supply Chain Management: An Experimental Study, *Production and Operations Management*, 12, 1.

Croson, R. and Donohue, K. (2004), Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information, OPIM Working Paper.

Croson, R., Donohue K., Katok, E. and Sterman, J. (2004), Order Instability in Supply Chains: Coordination Risk and the Role of Coordination Stock, Penn State Working Paper.

Cui, T. H., Raju, J. S. and Zhang, Z. J. (2004), Fairness and Channel Coordination, Working Paper, the Wharton School of Business, University of Pennsylvania.

Donohue, K. (2000), Efficient Supply Contracts for Fashion Goods with Forecast Updating and Two Production Modes, *Management Science* (2000), 46(11), 1397-1411.

Doyle, J. (1991), Innovations in Training, *Credit Magazine,* 17, 10-14.

Fehr, E. and Schmidt, K. (1999), A Theory of Fairness, Competition and Co-operation, *Quarterly Journal of Economics*, 114, 817-868.

Ferguson, M. E., Guide, Daniel V., and Souza, G. C. (2005), Supply Chain Coordination for False Failure Returns, Working Paper.

Fisher, M. (1994), *National Bicycle*, the Wharton School Case, Philadelphia, PA.

Forrester, J. (1958), Industrial Dynamics: A Major Breakthrough for Decision Makers, *Harvard Business Review*, 36, 37-66.

Frazier, G. (1983), On the Measurement of Interfirm Power in Channels of Distribution, *Journal of Marketing Research*, 20(2), 158-166.

Fuller, Joseph, B., O'Conor J., and Rawlinson, R. (1993), Tailored Logistics: The Next Advantage, *Harvard Business Review*, May-June, 87-98.

Gan, X., Sethi, S. P., and Yan, H. (2004), Coordination of Supply Chains with Risk-averse Agents, *Production and Operations Management*, 13(2), 135-149.

Gopal, A., Sivaramakrishnan, K., Krishnan, M. S., and Mukhopadhyay, T. (2003), Contracts in Offshore Software Development: An Empirical Analysis, *Management Science*, 49(12), 1671-1683.

Grubbs, Frank, E. (1969), Procedures for Detecting Outlying Observations in Samples, *Technometrics*, 11(1), 1-21.

Ho, T. H. and Zhang, J. (2004), Does Format of Pricing Contract Matter, Working Paper No. XL05-002, University of California, Berkeley.

Holt, C. A. and Laury, S. K. (2002), Risk Aversion and Incentive Effects, *American Economic Review*, 92(5), 1644-1655.

Isaac, R. M. and James, D. (2000), Just Who Are You Calling Risk Averse? *Journal of Risk and Uncertainty*, 20(2), 177-187.

Jackson, M.C. (1995), Beyond the Fads: Systems Thinking for Managers, *Systems Research*, **12**, 25-42.

Kagel, John H. and Alvin E. Roth, eds. (1995), *Handbook of Experimental Economics*, Princeton, Princeton University Press.

Kahneman, D., and Tversky, A. (1979), Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47, 263-291.

Kahneman, D., Jack L. Knetsch, and Thaler R. (1986), Fairness and the Assumptions of Economics, *Journal of Business*, 59(4), Part 2: The Behavioral Foundations of Economic Theory, 285-300.

Kamrad, B. and Siddique, A. (2004), Supply Contracts, Profit Sharing, Switching, and Reaction Options, *Management Science*, 50(1), 64-82.

Keser, C. and Paleologo, G. (2004), Experimental Investigation of Retailer-Supplier Contracts: The Wholesale Price Contract, IBM Research Working Paper.

Kim D. (1993), The Link Between Individual and Organizational Learning, *Sloan Management Review*, 35, 37-50.

Kleindorfer, P. R. and Wu, D. J. (2003), Integrating Long- and Short-term Contracting via Business-To-Business Exchanges for Capital-Intensive Industries, *Management Sciences*, 49(11), 1597-1615.

Kumar, N. (1996), The Power of Trust in Manufacturer-Retailer Relationships, *Harvard Business Review*, November-December, 92-106.

Lariviere, M. and Porteus, E. L. (2001), Selling to the Newsvendor: An Analysis of Price-only Contracts, *Manufacturing & Service Operations Management*, 3(4), 293-305.

Lee, H., P. Padmanabhan and Whang, S. (1997), Information Distortion in a Supply Chain: The Bullwhip Effect, *Management Science*, 43, 546-558.

Lee, H., P. Padmanabhan and Whang, S. (2004), Comments on Information Distortion in a Supply Chain: The Bullwhip Effect, *Management Science*, 50, 1887-1893.

Lim, N. (2004), An Experimental Study of Quantity Discount Contracts: Counterfactual Cognition in Multi-block Tariffs, Working Paper, the Wharton School of Business, University of Pennsylvania.

Lurie, N. H. and Swaminathan, J. M. (2005), Is Timely Information Always Better? The Effect of Feedback Frequency on Performance and Knowledge Acquisition, Working Paper.

Mortimer, J. H. (2004), Vertical Contracts in the Video Rental Industry, Harvard Working paper.

Robinson, P. (1994), How Do You Get a Good Deal on Software? It Pays to Pay Attention, *The Seattle Times*, November 06, 1994.

Roth, A. E. (2002), The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, *Econometrica*, 70, 1341-1378.

Schendel, J. D. (1994), Training for Troubleshooting, *Training & Development Journal,* 4**8**, 89-95.

Schweitzer, M. and Cachon, G. (2000), Decision bias in the Newsvendor Problem: Experimental Evidence, *Management Science*, 46(3), 404-420.

Senge, P. M. (1990), *The Fifth Discipline: The Art and Practice of the Learning Organization*, Doubleday/Currency: New York.

Senge, P.M. and Sterman, J. D. (1992), Systems Thinking and Organizational Learning, *European Journal of Operational Research,* 59, 137-150.

Spengler, J. J. (1950), Vertical Integration and Antitrust Policy, *The Journal of Political Economy*, 58(4), 347-352.

Springer, L., Stanne, M. E., and Donovan, S. S. (1999), Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis, *Review of Educational Research*, 69, 21-51.

Stata, Ray (1989), Organizational Learning – The Key to Management Innovation, *Sloan Management Review,* 30, 63-74.

Steckel, J. H., Gupta, S., Banerji, A. (2004), Supply Chain Decision Making: Will Shorter Cycle Times and Shared Point-of-Sale Information Necessarily Help?, *Management Science,* 50, 458-464.

Sterman, J. (1989a), Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment, *Management Science*, 35, 321-339.

Sterman, J. (1989b), Misperceptions of Feedback in Dynamic Decision Making, *Organizational Behavior and Human Decision Processes*, 43, 419-434.

Sterman, J., (2000) *Business Dynamics: System Thinking and Modeling for a Complex World*, McGraw-Hill, NY.

Tyler, T.R., and Lind, E. A. (1992), A Relational Model of Authority in Groups, in M. P. Zanna (Ed.), Advances in Experimental Social Psychology: 151-191, New York: Academic.

Uzzi, B. (1996), The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: the Network Effect. *American Sociological Review*, 61, 674-698.

Wu, D. J. and Kleindorfer, P. R. (2005), Competitive Options, Supply Contracting, and Electronic markets, *Management Science*, 51(3), 452-466.

Wu, Y. and Loch, C. (2006), The Effect of Social Preferences on Supply Chain Performance, Working Paper, INSEAD.

APPENDICES

Appendix A: Sample Instructions for Chapter I[7]

Retailer Game Instructions

In today's study, you will participate in *two games* where you will earn money based on your own decisions. If you follow the instructions carefully and make good decisions, you could earn a considerable amount of money. The unit of currency for this session is called a franc.

The Game Scenario
You are involved in the management of a supply chain that produces and sells widgets over multiple rounds. There are two members in each chain, a *Retailer* and a *Supplier*.

You are the retailer. The supplier is automated.

The Retailer buys widgets from the Supplier and sells widgets to the customer at **12** francs per unit. The supplier produces retailer's orders at **3** francs per unit. In each round, widgets are ordered and produced before you find out the actual customer demand.

Customer Demand

The customer demand per round consists of two parts:

1. There is a guaranteed demand of 50 units
2. There is an additional demand we will call **D**, which is an integer from 1 to 100, each equally likely. That is, there is a 1/100 chance that additional demand will be any one of the integers from 1 to 100

The total demand for the round is the sum of the guaranteed 50 units and the additional demand for that round. The additional demand drawn for any one round is independent of the additional demand for the earlier rounds. So a small or large additional demand in a round has no influence on whether additional demand is small or large in any other rounds.

Your task
Your task is to determine how many widgets to order, in addition to the guaranteed 50 units, so as to maximize your own total profits for the session. Your order always has to be an integer from 1 to 100. The number of additional units you order is called **Q.**

Feedback Information
In each round, after placing your order, you will be reminded of the order you just placed, the additional demand realized, the total customer demand, your own profit, and your automated

[7] Sample instructions below are for the Same-Frame treatments. Instructions for DLOW and DHIGH treatments are analogous, but the demand distributions and order decisions are described directly, and profit formulas do not include the fixed component. The instructions for the Wholesale Price contract are also presented analogously, but without reference to the additional parameters.

supplier's profit for that round. The computer will also display the history of the above information for all pervious rounds.

How you will be paid
You will participate in two separate games, each lasting 100 rounds. The games will differ in the terms of the contract supplier offers. Your total earnings from both games will be converted to US dollars at the rate of 2000 francs per dollar, added to your participation fee of $5 and paid to you in cash at the end of the session.

Buyback Contract Game

Contract terms
The supplier offered you a contract that requires you to pay a wholesale price of **9** francs per unit you order, and agrees to buy back any units unsold at the end of the period at a rebate of **8** francs per unit. Remember that you sell each unit for **12** francs per unit.

Calculating Your Profit
The guaranteed demand is 50 units. For those units you will earn a profit of $50\times(12-9)=150$ francs. You will earn additional profit based on the number of additional units you order and sell.

When your order Q turns out to be the same or lower than the additional customer demand D, your total profit for the round is:

$$\begin{aligned}\text{Your Profit} &= 150 + (12 - 9) \times Q \\ &= 150 + \quad 3 \quad \times Q\end{aligned}$$

For example, if the additional demand is 80 and you order 60, then your total profit for the round is $150 + 3 \times 60 = 330$ francs. Note that when the number of widgets ordered is less than demand, you lose opportunities for sales.

When your order Q turns out to be higher than the additional customer demand D, your total profit for the round is:

$$\begin{aligned}\text{Your Profit} &= 150 + (12 - 9) \times D - (9 - 8) \times (Q - D) \\ &= 150 + \quad 3 \quad \times D - \quad 1 \quad \times (Q - D)\end{aligned}$$

For example, if the additional demand is 40 and you order 60, then your total profit for the round is $150 + 3 \times 40 - 1 \times (60 - 40) = 250$ francs. Note that when the number of widgets ordered exceeds demand, you must dispose of the unsold units (since extra widgets go stale after a round, and cannot be carried as inventory into future rounds), and thus incur cost for excess widgets. Since the supplier will refund you 8 francs for each unsold unit, it costs you $9 - 8 = 1$ per unit.

Calculating Supplier's Profit

The supplier will earn $(9 – 3) \times 50 = 300$ francs per round for the guaranteed demand, $(9 – 3 = 6)$ for each additional unit you order and sell, and $9 - 8 – 3 = -2$ franks for each unit you order and not sell.

For example, if you order 60 and the additional demand is 80, the supplier's profit for the round is $300 + 6 \times 60 = 660$ francs; if you order 60 and the additional demand is 40, then the supplier's profit for the round is $300 + 6 \times 40 – 2 \times (60 – 40) = 500$ francs.

Revenue-Sharing Contract Game

Contract terms

The supplier offered you a contract that requires you to pay a wholesale price of **1** franc per unit you order, but you agree to pay an additional **8** francs for each unit you sell. Remember that you sell each unit for **12** francs per unit.

Calculating Your Profit

The guaranteed demand is 50 units. For those units you will earn a profit of $50 \times (12 - 1 - 8) = 150$ francs. You will earn additional profit based on the number of additional units you order and sell.

When your order Q turns out to be the same or lower than the additional customer demand D, your total profit for the round is:

$$\begin{aligned}\text{Your Profit} &= 150 + (12 – 1 - 8) \times Q \\ &= 150 + 3 \times Q\end{aligned}$$

For example, if the additional demand is 80 and you order 60, then your total profit for the round is $150 + 3 \times 60 = 330$ francs. Note that when the number of widgets ordered is less than demand, you lose opportunities for sales.

When your order Q turns out to be higher than the additional customer demand D, your total profit for the round is:

$$\begin{aligned}\text{Your Profit} &= 150 + (12 – 1 - 8) \times D – 1 \times (Q – D) \\ &= 150 + 3 \times D – 1 \times (Q – D)\end{aligned}$$

For example, if the additional demand is 40 and you order 60, then your total profit for the round is $150 + 3 \times 40 – 1 \times (60 – 40) = 250$ francs. Note that when the number of widgets ordered is greater than demand, you must dispose of the unsold units (since extra widgets go stale after a round, and cannot be carried as inventory into future rounds), and thus incur cost of 1 franc per unit for excess widgets.

Calculating Supplier's Profit

The supplier will earn (1 + 8 – 3) × 50 = 300 francs per round for the guaranteed demand, (1 + 8 – 3 = 6) for each additional unit you order and sell, and 1 – 3 = -2 franks for each unit you order and not sell.

For example, if you order 60 and the additional demand is 80, the supplier's profit for the round is 300 + (8 + 1 – 3) × 60 = 660 francs; if you order 60 and the additional demand is 40, then the supplier's profit for the round is 300 + 6 × 40 – 2 × (60 – 40) = 500 francs.

Supplier Game Instructions

In today's study, you will participate in *two games* where you will earn money based on your own decisions. If you follow the instructions carefully and make good decisions, you could earn a considerable amount of money. The unit of currency for this session is called a franc.

The Game Scenario
You are involved in the management of a supply chain that produces and sells widgets over multiple rounds. There are two members in each chain, a *Retailer* and a *Supplier*.

You are the supplier. The retailer is automated.

The retailer decides how many widgets to buy from the supplier and sells widgets to the customer at **12** francs per unit. The supplier decides how to charge the retailer by specifying some contract terms and produces retailer's orders at **3** francs per unit. In each round, widgets are ordered and produced before you find out the actual customer demand.

Customer Demand

The customer demand per round consists of two parts:

1. There is a guaranteed demand of 50 units.
2. There is an additional demand we will call **D**, which is an integer from 1 to 100, each equally likely. That is, there is a 1/100 chance that additional demand will be any one of the integers from 1 to 100.

The total demand for the round is the sum of the guaranteed 50 units and the additional demand for that round. The additional demand drawn for any one round is independent of the additional demand for the earlier rounds. So a small or large additional demand in a round has no influence on whether the additional demand is small or large in any other rounds.

Your task
Your task is to determine the contract terms that specify the transfer payment between you and the retailer so as to maximize your own total profits in this session. The automated retailer has been programmed to place orders so as to maximize his own expected profits given the contract terms you offered.

Feedback Information
In each round, after you set the contract terms, the computer will display the optimal order quantity that the automated retailer places and the corresponding expected profit of the retailer.

After confirming your decisions, you will receive information on the retailer's order, the additional and total customer demand realized, the retailer's and your own profits for that round. The computer will also display the history of the above information for all pervious rounds.

How you will be paid
You will participate in two separate games, each lasting 100 rounds. The games will differ in the type of the contract that you are asked to use. Your total earnings from both games will be converted to US dollars at the rate of 5000 francs per dollar, added to your participation fee of $5 and paid to you in cash at the end of the session.

Buyback Contract Game

Contract terms
You are asked to set a contract that specifies a wholesale price (**W**) at which you charge for each unit the retailer orders, and a rebate (**B**) at which you agree to buy back any units unsold from the retailer at the end of the period.

Note the wholesale price per unit you offer should be less than 12 francs (the retail price) and greater than 3 francs (your production cost). The rebate per unit you offer cannot be greater then the wholesale price you specify and cannot be less than zero. You are allowed to use at most two decimal places in setting W and B.

Calculating Your Profit
The guaranteed demand is 50 units. For those units you will earn a profit of $(W - 3) \times 50$ francs per round. You will earn additional profit based on the number of units (Q) in addition to the guaranteed demand that your automated retailer orders and sells.

When your retailer's additional order Q turns out to be the same or lower than the additional customer demand D, your total profit for the round is:

$$\text{Your Profit} \quad = [(W - 3) \times 50] + [(W - 3) \times Q]$$

When your retailer's additional order Q turns out to be higher than the additional customer demand D, your total profit for the round is:

$$\text{Your Profit} \quad = [(W - 3) \times 50] + [(W - 3) \times D] + [(W\text{-}B) \times (Q - D)]$$

For example, if you set the wholesale price at 9 francs/unit and the rebate at 7 francs/unit, the automated retailer will order 60 widgets in addition to the 50 units of guaranteed demand for the round. If the additional demand is 80, then your profit is $(9 - 3) \times (50 + 60) = 660$ francs for the round. If the additional demand is 40, then your profit is $(9 - 3) \times 110 - 7 \times (60 - 40) = 520$ francs for the round.

Note that for a fixed rebate (B), the higher the wholesale price you offer, the lower the quantity the retailer orders, and vice versa. While for a fixed wholesale price (W), the higher the rebate you offer, the higher the quantity the retailer orders, and vice versa.

Calculating Retailer's Profit
The retailer will earn (12 – W) × 50 francs per round for the guaranteed demand, (12 – W) for each additional unit he orders and sells, and (B – W) francs for each unit he orders and not sell. For the above example, given W = 9 and B = 7, the automated retailer orders an additional of 60 units for the round. If the additional demand is 80 for the round, the retailer's profit is (12 – 9) × 50 + (12 – 9) × 60 = 330 francs for the round. If the additional demand is 40, the retailer's profit is (12 – 9) × 50 + (12 – 9) × 40 + (7 – 9) × (60 – 40) = 230 francs for the round.

Revenue-Sharing Contract Game

Contract terms
You are asked to set a contract that specifies a wholesale price (**W**) at which you charge for each unit the retailer orders, and a revenue share (**R**) (or commission) that you charge for each unit the retailer sells.

Note the wholesale price per unit per unit you offer should be less than 12 francs (retail price) and no less than zero. The revenue share per unit you acquire should be less than the retailer's total revenue per unit, which is the retailer price minus the wholesale price you offer, and should be no less than zero. You are allowed to use at most two decimal places in setting W and R.

Calculating Your Profit:
The guaranteed demand is 50 units. For those units you will earn a profit of (W – 3 + R) × 50 francs per round. You will earn additional profit based on the number of units (Q) in addition to the guaranteed demand that your automated retailer orders and sells.

When your retailer's additional order Q turns out to be the same or lower than the additional customer demand D, your total profit for the round is:

Your Profit $= [(W - 3 + R) \times 50] + [(W - 3 + R) \times Q]$

When your retailer's additional order Q turns out to be higher than the additional customer demand D, your total profit for the round is:

Your Profit $= [(W - 3 + R) \times 50] + [(W - 3 + R) \times D] + [(W - 3) \times (Q - D)]$

For example, if you set the wholesale price at 2 francs/unit and the revenue share at 7 francs/unit, the automated retailer will order 60 widgets in addition to the 50 units of guaranteed demand for the round. If the additional demand is 80, then your profit is (2 – 3 + 7) × (50 + 60) = 660 francs for the round. If the additional demand is 40, then your profit is (2 – 3 + 7) × 50 + (2 – 3) × 60 + 7 × 40 = 520 francs for the round.

Note that, for a fixed revenue share, the higher the wholesale price you offer, the lower the quantity the retailer orders, and vice versa. And for a fixed wholesale price, the higher the revenue share you charge, the lower the quantity the retailer orders, and vice versa.

Calculating Retailer's Profit

The retailer will earn (12 – W – R) × 50 francs per round for the guaranteed demand, (12 – W – R) for each additional unit he orders and sells, and (– W) francs for each unit he orders and not sell.

For the above example, given W = 2 and R = 7, the automated retailer orders an additional 60 units for the round. If the additional demand is 80 for the round, the retailer's profit is (12 – 2 – 7) × 50 + (12 – 7 – 2) × 60 = 330 francs for the round. If the additional demand is 40, retailer's profit is (12 – 2 – 7) × 50 + (12 – 2 – 7) × 40 + (– 2) × (60 – 40) = 230 francs for the round.

Appendix B: Sample Instructions for Chapter II

Two-person Contracting Game – Buyback Contract

You are about to participate in an experiment regarding supply chain decision-making where you will earn money based on your own decisions and decisions of others. Your earnings are yours to keep. The unit of currency for transactions during the experiment is called a franc. The experiment lasts *100* rounds. Your total earnings accumulated during the entire session will be converted to US dollars at the rate of **2000** francs per dollar, which added to your participation fee of $5, will be paid to you in cash at the end of the session.

Upon arrival, you are assigned a player number that will serve as your identification during the experiment. You are NOT allowed to communicate with the other participants.

The Game Scenario
You are involved in the management of a supply chain that consists of two members: a *Retailer* and a *Supplier*. Your will be randomly assigned one of the two roles and will be matched with a participant who takes the other role. You will remain in the same role and play with the same partner for the entire session anonymously.

The supplier produces some fictional widgets at **3** francs per unit and sells the product to the retailer. The retailer then sells the product to the customer at **12** francs per unit.

The customer demand per round, which we call D, is randomly generated by the computer from a range of **50** to **150** and rounded up to the nearest integer. Each integer in the range is selected *equally likely*. That is, there is a 1/100 chance that demand will be any one of the integers from 51 to 150. The demand drawn for any round is *independent* of the demand from earlier rounds. So a small or large demand in earlier rounds has no influence on whether demand is small or large later on.

In each round, the supplier makes decision *first* to propose a contract offer that specifies how he charges the retailer for providing the product. The retailer then decides whether to accept the contract and how many widgets to order before the customer demand is revealed for the round.

The contract used in this game is called "Buyback" contract, in which the supplier decides a ***wholesale price*** (W) at which to charge each unit the retailer orders, and a **rebate** (B) at which to refund the retailer for each unit unsold. The retailer can ***reject*** the supplier's offer – in that case, both parties will have *zero* profit for that round. If the retailer accepts the offer, he can either order the ***"optimal"*** quantity (Q), an amount computed by the system that maximizes the retailer's *average* profits given the supplier's offer and the random demand, or order the minimum customer demand of ***50***.

Supplier's Decision Interface
In each round, the supplier is asked to input his wholesale price and rebate into some decision boxes. The wholesale price has to be set between 3 francs (the production cost) and 12 francs (the retail price). The rebate has to be no greater then the wholesale price and no less than zero.

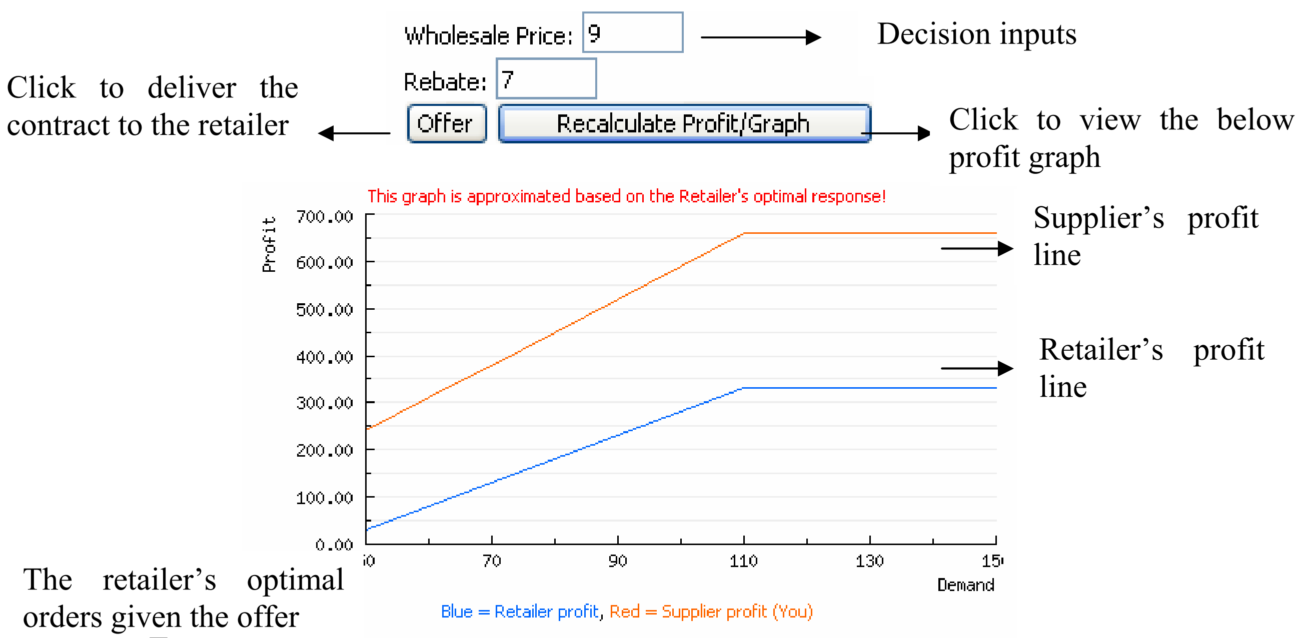
And at most two decimal places are allowed in setting these two parameters. The supplier's profit will depend on the random demand drawn in the round and is calculated as:

$$= W \times Q - 3 \times Q \qquad \text{if } D >= Q$$
$$= W \times Q - 3 \times Q - B \times (Q - D) \qquad \text{if } D < Q$$

The supplier is allowed to view possible profit outcomes for both parties given his offer by clicking a button called *"Recalculate Profit/Graph"*. The supplier will see, given his proposed wholesale price and rebate that:

1. The optimal order quantity computed for the retailer.
2. A "*Profit Graph*" that plots the retailer's and the supplier's profits against different random demand realizations, when the retailer accepts the offer and orders the optimal quantity.

The supplier may try as many different combinations of wholesale prices and rebates as he wants and view the corresponding results by clicking the "*Recalculate Profit/Graph*" button each time. If the supplier is satisfied with his contract offer, by clicking a button called *"Offer"*, he will deliver the final contract to his retailer. The supplier then waits for his retailer to make decisions. The following illustrates the computer interface for the supplier.



This table shows the Retailer's optimal response and cooresponding profits.

Order Quantity	Your Min Profit	Your Max Profit	Your Ave Profit	Their Min Profit	Their Max Profit	Their Ave Profit
110.00	240.00	660.00	533.17	30.00	330.00	239.41

Suppose the supplier inputs a wholesale price of 9 and a rebate of 7. By clicking the "*Recalculate Profit/Graph*" button, the supplier will know the optimal order quantity of the retailer given his contract will be 110. Suppose his retailer accepts his offer and orders 110. If D is 130 and thus the retailer results in no inventory, then supplier's profit is (9 x 110 – 3 x 110) = 660 francs for that round. If D is 90 and thus the retailer results in (110 – 90) = 20 units unsold, the supplier's profit is then (9 x 110 – 3 x 110 – 7 x 20) = 520 francs for that round.

Retailer's Decision Interface
After the supplier makes his decision, the computer will display the proposed contract and compute the optimal order quantity given such offer for the retailer. The retailer is asked to indicate whether to accept the supplier's offer; and if he accepts the offer, whether to order the optimal amount or order 50 for that round.

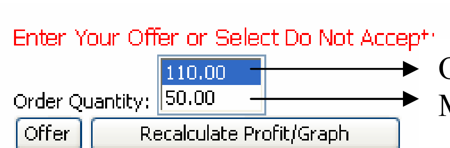
Similarly, the retailer is also provided with the option to view possible profit outcomes for both parties if he accepts the offer. The retailer needs to first choose the order quantity and then click the "*Recalculate Profit/Graph*" button to view the corresponding *"Profit Graph"*. When the retailer has made his final decision, clicking the *"Offer"* button will reveal the results of that round. The following shows the computer interface for the retailer.

Accept	Wholesale Price	Rebate	Accepts Your Offer	Your Minimum Profit	Your Maximum Profit	Your Average Profit	Their Minimum Profit	Their Maximum Profit	Their Average Profit
◉	9	7		30.00	330.00	239.41	240.00	660.00	533.17
○	Do Not Accept Offer								

Supplier's contract offer

Click to reject the offer

Click to deliver the final decision



Optimal order
Minimum demand order

Click to view the Profit Graph

In the above example, the supplier offers a wholesale price of 9 and a rebate of 7. Given such contract, the optimal order quantity for the retailer is 110. The retailer's actual profit depends on the random demand drawn for that period and is calculated as:

$$= 12 \times Q - W \times Q \qquad \text{if } D >= Q$$
$$= 12 \times D - W \times Q + B \times (Q - D) \qquad \text{if } D < Q$$

Suppose the retailer accepts the offer and orders 110. If D is 130, which results in 110 units of sales, then the retailer's profit is (12 x 110 – 9 x 110) = 330 francs for that round. If D is 90 and results in 20 units unsold, then the retailer's profit is (12 x 90 – 9 x 110 + 7 x 20) = 230 francs for that round.

Feedback Information

Information on the decision interface for the retailer is labeled in *blue* and for the supplier is in *red*. After the retailer makes his decision, information on the realized customer demand, decisions and profits made by both parties for that round will be shown. The game will move on to the next period after both the retailer and the supplier click the *"Continue"* button at the bottom of the page. The computer will also display the history of the above information for all previous rounds.

Appendix C: Sample Instructions for Chapter III

Beer Game Instruction

You are about to participate in a study where you will earn money based on your decisions and the decisions of others. All money earned is yours to keep and will be paid to you privately IN CASH at the end of the session. During the session the unit of account will be **tokens**. At the conclusion of the session, all tokens earned will be converted into U.S. dollars at the conversion rate of_***0.015***_dollars per token. Your converted earnings plus a five-dollar-show-up fee will be paid to you.

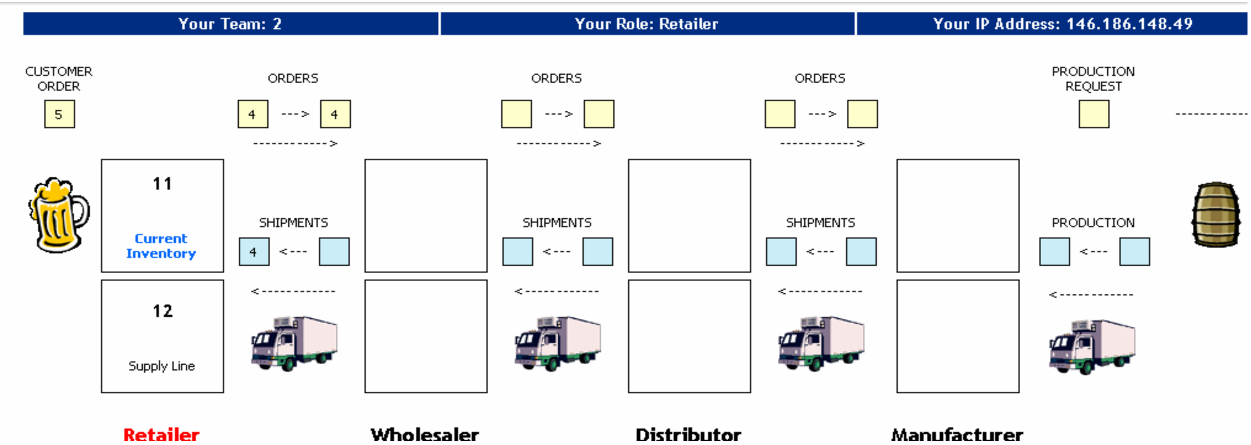
This session will consist of two parts. First you will participate in **20** training periods/weeks to help you better understand the environment. After that you will play with three other participants as a team in the rest of the session. You will have a chance to discuss your strategy with members on your team, but you are only allowed to communicate with them when you are told to. If you have any question, raise your hand, and I will answer it.

The Game

You are about to take the role of a manager in a business of the production and distribution of beer. As in many real companies, consumer's demand is satisfied by bringing the beer from the manufacturer through a *supply chain* instead of by purchasing the beer directly from the brewery. There are four members in the chain, the **Retailer**, **Wholesaler**, **Distributor**, and **Manufacturer**, which work together to meet **Customer** demand.

These four members are *interrelated* through a customer-supplier relationship. Each of them keeps an inventory of beer, which is used to fill the orders they receive from their immediate customer, and also places orders with their immediate supplier. Each member of your team will be assigned to one of these four roles to manage inventory and satisfy the customer demand. Your role is ___________. Please refer to the figure (an example of Retailer) below and the screen in front of you.

- The **retailer** (on the far left on your screen) ships directly kegs of beer to the **customer** of your chain and orders kegs of beer from the **wholesaler,** its immediate supplier.
- The **wholesaler** (to the right of the retailer) ships directly kegs of beer to the **retailer**, its immediate customer, and orders kegs of beer from the **distributor,** its immediate supplier.
- The **distributor** (to the right of the retailer) ships directly kegs of beer to the **wholesaler,** its immediate customer, and orders kegs of beer from the **manufacturer,** its immediate supplier.
- The **manufacturer** (to the right of the retailer) ships directly kegs of beer to the **distributor,** its immediate customer, and produces beer.



Customer Orders
The retailer receives orders from customers, simulated by the computer. In this game, ***<u>the customer demand is UNIFORM distributed from 0 to 8 each week.</u>*** It means that the customer order for each week will be one of the nine integers (0,1,2,3,4,5,6,7,8) with equal chance, i.e., probability of 1/9. And customer demand in any week is *<u>independent</u>* from the demand in any other week. ONLY retailer can see the actual customer order for each week, which is displayed in the yellow box in the top left corner of the screen, labeled *Customer Order*.

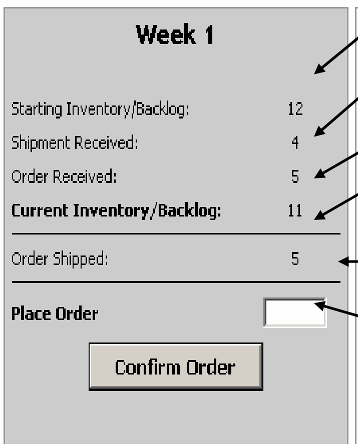
Delays
Yellow boxes and blue boxes between roles, labeled *Orders,* and *Shipments*/*Production* denote two types of **delays** accordingly in the chain: **order delays** and **shipments/production delays.** They reflect the order processing time and shipment/production time experienced in most real supply chains. For the retailer, wholesaler, and distributor, there is a total of four weeks of delay. It takes **two** weeks for orders placed to be received by the supplier and another **two** weeks for shipments to arrive. Thus, if you place an order in the third week of the game, your supplier will receive it in the fifth week, and you will receive the shipments from your supplier in the seventh week. For the manufacturer, it takes only **one** week after an order is placed to authorize production, but **two** weeks for production to complete. Thus, if the manufacturer places an order in the third week of the game, he will receive it in the sixth week after a total of three weeks of delay.

Private Information: Inventory and supply line
The upper big white box labeled *Current Inventory* shows your own ***on-hand inventory*** for each week. The lower big white box labeled *Supply Line* gives information on ***orders that YOU HAVE PLACED to your supplier but have not yet received in shipments.*** It equals the number of all orders you placed during last four weeks (last three weeks for manufacturer) minus the new shipment you have just received in that period. For example, if everyone starts out by placing orders of 4, 4, 4, 4, in the last four weeks and receives a shipment of 4 in the beginning of a period, the supply line will be 4+4+4+4 (4+4+4 for manufacturer) – 4 shipments received =12 (8 for manufacturer). **Note that you will eventually receive those orders.**

You can also see the amounts you shipped in the blue box to the left of your inventory box, the incoming shipments for next week from your immediate supplier in the blue box to the right of your inventory box, and the incoming order for that week from their immediate customer in the yellow box to the left of your inventory box.

*Playing the Game***:** The below table on your screen describes the events happened **within** a week.

Present week activities:



Starting Inventory/Backlog at the beginning of the

Shipment Received from upstream team member this

Order Received from downstream team member this

Current inventory = Starting Inventory/Backlog + Shipment Received - Order Received

Order Shipped to downstream team member this week

Order to be placed to the upstream team

At the start of every week, each player will receive an order (the *Order Received* in the table above), which was placed by his immediate customer two weeks ago. For the retailer, this will be the customers' demand. Players will also receive a shipment from their immediate supplier (*Shipment Received*), which was shipped two weeks ago. For the manufacturer, the Shipment Received will be the amount of beer that was started three weeks ago and has now been produced.

How the Shipment amount is determined

Whenever there is enough beer to send the entire order amount PLUS any backlogged orders from previous weeks, the computer will automatically authorize this shipment amount. Any inventory remaining after subtracting this shipment becomes the starting inventory for the following week. If there is *not* enough beer to ship the entire amount ordered, the computer will ship all available inventory. In this case, some orders will be *backlogged,* and will have to be filled in future weeks. In the case of backlogs, the Ending Inventory/Backlog will be negative, representing the number of kegs the player still owes his customer. Any unfilled orders carry over to the next week, thus **backlog is cumulative**.

For the retailer, wholesaler, and distributor, if your supplier is out of stock, *you might not be able to get all the amounts you want from your immediate supplier* after 4 weeks of delay. For example, if a wholesaler placed an order of 16 in the third week, the order arrives in the fifth week to the distributor, if the distributor has an on hand inventory of 8 in that period and receives no new shipments for that week, then in the seventh week, the wholesaler will only get a shipment of 8, since the distributor could only partially fill the order. The manufacturer is, however, assumed to have infinite production capacity, thus he will always get what he wants after three weeks of delay.

Your task for each week is to decide how many kegs of beer to order from your supplier. You will enter the amount you choose in the box at the bottom of the display (*Place Order*), and click the confirm order button. The program will ask you to confirm your order again. Once you do, the order will be sent to your supplier (**remember the order delay**!) for him to fill.

Note: Your order can be zero or larger, but cannot be a negative or fractional.

In summary, the display describes the order of events in a period.
1. All members begin with their starting inventory/ backlog,
2. Shipments received from upstream immediate supplier, and supply line information updated,
3. Order received from downstream member or customer,
4. Order fulfilled based on the order received, amount available in inventory and any backlog,
5. Current inventory is calculated: starting inventory +shipment received – order received,
6. Ordering decision is made (order placed),
7 Current inventory for this period becomes the starting inventory for the next period.

Cost Parameters and Your own Performance
There are two costs that you might incur in this game, all displayed in the yellow box labeled "*Parameters*":

- <u>Inventory Costs:</u> It costs 0.50 tokens per week to keep a keg of beer in inventory.
- <u>Backlog Costs:</u> It costs 1.00 tokens per week to be in backlog by one keg of beer.

Costs are calculated based on the current inventory/backlog you have at the end of each week. Your own *current period* costs and *cumulative* costs thus far are displayed in the yellow box labeled "*Performance*". The larger your inventory, the higher your costs, but if you run out of beer and are unable to fill all the orders you receive you will incur the larger backlog cost until you can fill the orders in the backlog. It is up to you to decide how much inventory you want to have.

Example 1: Suppose your beginning inventory is 4, the shipment received is 4 and the order received is 2. Using the logic laid out in the display above, your ending Inventory/Backlog is then [4 (Starting Inv/Backlog) + 4 (Shipment Received) – 2 (Orders Received)] = 6 kegs of beer. Note that here Orders Shipped is 2 since there is enough inventory available to fill the entire order. The cost for the week is 6 units of inventory * 0.50 tokens per unit = 3 tokens.

Example 2: Suppose your beginning inventory is 4, the shipment received is 2 and the order received is 8. Now the total amount on hand [4 (in inventory) + 2 (shipment received) = 6] is not enough to entirely fill the order of 8; the most you can ship is 6. Your Ending Inv/Backlog is now [4 (Starting Inv/Backlog) + 2 (Shipment Received) – 8 (Orders Received)] = -2 kegs of beer. The cost for the week is 2 units of backlog * 1 tokens per unit = 2 tokens.

History Plots:
Clicking on the button labeled *View History* shows the graph of your own weekly inventory/backlog, the graph of orders have you placed, and the graph of orders you have received so far. To return to the game, click the *Close* button.

Money Earnings
Each team is given an ***endowment* of *5000* tokens** as a start-up in the beginning. All team members' costs will be added together to calculate the total team costs in order to get your final earnings which is based on the below formula:

Earnings = (Endowment – Total Chain Costs)/4 * Conversion Rate + Show-up fee

Every member of your team will earn the same amount. The lower your *team's* costs are, the more money you will earn in this game. Thus, your *objective is to make ordering decisions that minimize the total costs of your team over the entire game*. However, it is possible for a team to go bankrupt during the game. If your team's endowment minus your chain costs becomes negative before the pre-specified number of weeks end, your earnings will be ONLY the show-up fee.

Training Session and Communication (THIS PART MAY VARY ACROSS TREATMENTS)
Before actually playing the game, you will be given the chance to practice ***twenty*** rounds to get familiar with the system. Results from this part will not affect your final earnings. Specifically, *you will be the only one in your practicing supply chain, thus you will play four roles in this system, manage inventory and make ordering decisions for four locations sequentially in each period.* After the practice, you will have **10 minutes to communicate** with your team members and your conversation will be recorded for research purpose. You will play the real game with your team members, and your role of _____________will always be kept the same.

Note: Both the practice and the actual game start with each participant having **12 kegs in inventory** and **4 kegs in each delay position/box.**

Ending the Game
The number of weeks you will actually play has been determined in advance. However, *you will not know in advance how long the game will last*. Once your team has completed the game, you will see a final screen that shows the total costs for all your team members and your final earnings (including the show-up fee) in this game. After you have completed the game, please fill in the survey by clicking the URL provided. Thanks for your cooperation.

Quiz

Please answer the following questions to check your understanding about the rules of the game. We will go over the answers. If you have questions, raise your hand.

1. Suppose customer demand was 4 for the last 3 periods, is it less likely to be 4 next period? ______. What is the chance it will be 4 in next period? ______, the chance it will be 0? ______, be 7? _____.

2. Suppose your beginning inventory is 6, the shipment received is 2 and the order received is 10. At the end of the week, do you have inventory or are you in backlog and how much? ________. What are your shipments for this period? _____. What is your cost for the week? ________

3. Suppose your beginning inventory is 2, the shipment received is 4 and the order received is 4.At the end of the week, do you have inventory or are you in backlog and how much? ________________ What are your shipments for this period? _____What is your cost for the week? ________

4. Suppose the retailer placed an order of 8 in the fourth week, if wholesaler has -3 on hand inventory (backlogged) in sixth week and receives no shipment for that week, what will be the shipment received by retailer in the eighth week? ________

5. Suppose the orders you have placed during the last 4 periods are 4,4,0,0, and at the start of this week you have just received a shipment of 4, what is your *supply line* at this point? __________

6. Suppose you and your teammates each have costs of 10 tokens every week, what is your total team cost and per week? ________If the game lasts 50 weeks, how much will each member of your team earn from the game (including the participation fee)? ______________________________ (Outline the formula)

Appendix D: Sample Consent Form

INFORMED CONSENT FORM FOR SOCIAL SCIENCE RESEARCH
The Pennsylvania State University

Title of Project: Inventory Ordering Decisions in Individual and Supply Chain Settings
Principal Investigator: Elena Katok, 509H Business Administration Building, University Park, PA 16802 (814) 814-863-2180 ekatok@psu.edu

1. Purpose of the Study: The study in which you will be participating is part of research intended to assess how people make decisions in simple economic situations. By conducting this study, we hope to improve our understanding of how economic institutions do and can work.
2. Procedures to be followed: If you agree to take part in this research, you will participate in a series of economic games. In this session, you may be allowed to discuss with your partners for about ten minutes, and your discussion may be recorded by tape for research purpose only. A more detailed description of the games you will play is included on a separate Instructions sheet. Please examine this sheet now, if you have not already done so.
3. Discomforts and Risks: There are no risks in participating in this research beyond those experienced in everyday life.
4. Benefits:
a. You might learn more about yourself by participating in this study. You might have a better understanding of how to make decisions.
b. This research might provide a better understanding of how people make decisions in economic situations. This information could help plan programs, make student services better. This information might help to draw conclusions about how people exercise their options in economic situations
5. Duration: It will take about 120 minutes to complete the session.
6. Statement of Confidentiality: Only the person in charge, and his/her assistants, will know your identity. If this research is published, no information that would identify you will be written. If you consent to be audiotaped, the tapes will be kept in a locked drawer in the locked office of Dr. Elena Katok and Yan Wu will access to these tapes. The tapes will be destroyed by the year 2008. The Office for Research Protections and the Social Science Institutional Review Board (IRB) may review records related to this project.
7. Right to Ask Questions: You can ask questions about the research. The person in charge will answer your questions. Contact Elena Katok at 814-863-2180 with questions. If you have questions about your rights as a research participant, contact Penn State's Office for Research Protections at (814) 865-1775.
8. Compensation: In return for you participation, you will receive $5 plus any earning from the games you participate in.
9. Voluntary Participation: You do not have to participate in this research. You can end your participation at any time by telling the person in charge. You do not have to answer any questions you do not want to answer.

You must be 18 years of age or older to consent to participate in this research study. If you consent to participate in this research study and to the terms above, please sign your name and indicate the date below.

You will be given a copy of this consent form to keep for your records.

______________________________ Date ____________________

Participant Signature

The informed consent procedure has been followed.

______________________________ Date ____________________

Investigator Signature

ORP USE ONLY: IRB#01B0811
The Pennsylvania State University
Office for Research Protections
Approval Date: 04/25/05 T. Kahler
Expiration Date: 04/24/06 T. Kahler
Social Science Institutional Review Board

VITA

Yan Wu

Feb. 25, 1979	Born at Beijing, People's Republic of China
1997	Graduated from No. 4 High School, Beijing, China
1997-2001	Majored in Management Information Systems, University of International Trade and Economics, China
2001-2006	Graduate Assistant, Department of Supply Chain and Information Systems, Penn State University
2006-present	Assistant Professor of Decision Sciences, School of Business, University of Kansas

FIELD OF STUDY

Major Field: Supply Chain Management with Dual Title in Operations Research

PUBLICATIONS

Wu, Yan and Katok E. (2006), Learning, Communication and the Bullwhip Effect, *Journal of Operations Management (*forthcoming).

Wu, Yan and Katok E. (2006), Contracting in Supply Chains: A Laboratory Investigation, submitted to *Management Science*.

PROFESSIONAL SOCIETIES

The Institute for Operations Research and the Management Sciences (INFORMS)