EXPLAINING RETAILER'S ORDERING BEHAVIOR IN SUPPLY CHAIN EXPERIMENTS

by

GÜLFİDAN AKOĞLU

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
August 2014

EXPLAINING RETAILER’S ORDERING BEHAVIOR IN SUPPLY CHAIN EXPERIMENTS

APPROVED BY:

Assist. Prof. Dr. Murat Kaya ………………………
(Thesis Supervisor)

Assoc. Prof. Dr. Abdullah Daşcı ………………………..

Assoc. Prof. Dr. Kemal Kılıç ………………………..

DATE OF APPROVAL: ……………………………………

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Assistant Professor Murat Kaya, for his guidance and patience throughout this research. His wide knowledge, encouragement, understanding and kindness have been a great value for me.

I am gratefully thankful to my thesis committee for spending their valuable time on my thesis and giving me insightful comments and suggestions. Special thanks to Fall 2012/2013 ENS 492 project group members. I acknowledge the support of my alumni friends, especially to Özge Arabacı, Uğur Emeç, Mahir Yıldırım, Utku Olgun, Ekin Köker and Halil Şen for keeping me company through the research. I owe my loving thanks to my husband, Kürşat Yılmaz Akoğlu, for his patience, for his infinite support endless understanding and love.

My heartfelt appreciation goes to Ongun Karataş the best friend and the best brother of all time, for his advises and patience.

Finally, I owe my greatest gratitude to my parents and grandparents for their endless love, infinite support and trust throughout my life. I especially like to thank to my mother, Perihan Karataş, for her tolerance during my whole life. I dedicate this research to my life's hero, my father Yasin Karataş.

EXPLAINING RETAILER'S ORDERING BEHAVIOR IN SUPPLY CHAIN EXPERIMENTS

Gülfidan Akoğlu

Industrial Engineering, MSc Thesis, 2014

(Thesis Supervisor): Asst. Prof. Murat Kaya

Keywords: supply chain management, contracting, behavioral operations, behavioral experiments, decision biases, decision heuristics, learning

Abstract

In this thesis, we study the retailer's ordering behavior in a manufacturer-retailer supply chain where the retailer faces the newsvendor problem. Analytical literature predicts that the retailer will use the critical ratio solution when determining her order quantity from the manufacturer. When real human beings play the roles of manufacturer and retailer in controlled experiments, however, the retailer decisions are observed to deviate from these theoretical predictions. The deviations are due to (1) individual biases and heuristics, (2) the strategic interaction between the two players. Literature has studied the effects of individual biases and heuristics using simple newsvendor experiments. However, very few researchers have conducted experiments where both sides are human. This extension is valuable because supply chain relations in practice depend on human-to-human interaction between managers. In this study, using data from the supply chain experiments of Şahin and Kaya (2011), we aim to answer the following questions: (1) Do retailer subjects follow the heuristics observed in simple newsvendor experiments? (2) What are the factors affecting retailer decisions? (3) Do retailer subjects learn to make better decisions over time? We find that retailer behavior is highly heterogeneous. While there is support for the use of decision heuristics at the aggregate level, we have mixed results at individual level. Likewise, the factors that affect retailer order quantity are found to be subject-dependent. The extent of learning is also found to differ from subject to subject.

TEDARİK ZİNCİRİ DENEYLERİNDE PERAKENDECİ SİPARİŞ DAVRANIŞININ AÇIKLANMASI

Gülfidan Akoğlu

Endüstri Mühendisliği, Yüksek Lisans Tezi, 2014

(Tez Danışmanı): Yrd. Doç. Dr. Murat Kaya

Anahtar Kelimeler: tedarik zinciri yönetimi, sözleşmeler, davranışsal operasyon, davranışsal deneyler, kararlarda yanlılık, karar sezgiselleri, öğrenme

Özet

Bu tezde, perakendecinin "gazeteci çocuk" problemi ile karşı karşıya kaldığı bir üretici-perakendeci tedarik zincirinde, perakendecinin sipariş davranışını ele aldık. Analitik literatür, perakendecinin üreticiden sipariş edeceği miktarı belirlerken kritik oran çözümünü kullanacağını öngörür. Kontrollü deneylerde gerçek insanlar üretici ve perakendeci rolü aldıklarında ise perakendecinin sipariş miktarı kararlarının teorik tahminlerden saptığı görülmüştür. Bu sapmalar (1) bireysel önyargılar ve sezgisellerden, (2) iki oyuncu arasındaki stratejik etkileşimden kaynaklanmaktadır. Literatür, basit gazeteci çocuk deneyleri kullanarak bireysel önyargıların ve sezgisellerin etkilerini ele almıştır. Ancak, çok az araştırmacı her iki tarafın da insan olduğu deneyler gerçekleştirmiştir. Bu tezde, Şahin ve Kaya (2011)'in tedarik zinciri deney verileri kullanılarak aşağıdaki soruların cevaplanması hedeflenmiştir. (1) Perakendeciler basit gazeteci çocuk deneylerinde gözlemlenen sezgisel yöntemleri kullanıyor mu? (2) Perakendeci kararlarını etkileyen faktörler nelerdir? (3) Perakendeciler zamanla daha iyi kararlar vermeyi öğreniyorlar mı? Ana bulgumuz, perakendeci davranışlarında gözlemlediğimiz heterojenliktir. Sonuçlarımız toplam düzeydeki sezgisel karar kullanımını desteklerken, bireysel düzeydekileri desteklememiştir. Aynı şekilde, hem perakendeci sipariş miktarını etkileyen faktörlerin, hem de öğrenme derecesinin kişiye bağlı olduğu gözlemlenmiştir.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

1 INTRODUCTION

Most products in today's world reach end customers through supply chains that consist of multiple firms. Specifically, the supply chain encompasses all steps it takes to get a good or service from the supplier to the customer. Supply chain management is important for modern businesses because it synchronizes activities of partner businesses, achieving higher efficiency. However, at the same time, it introduces the need for "coordination" between the chain members.

Every supply chain consists of individual firms the purpose of which is maximizing its own profit. Individual profit maximization causes inefficiency from the supply chain point of view, such as the well-known "double marginalization" problem (Spengler 1950). In order to increase the overall profit of the supply chain, the members of a supply chain must improve their coordination with each other. Supply chain coordination can be improved by using proper contracts between supply chain members. For this reason, the study of contracts between supply chain members has attracted great attention in business as well as in academic literature.

The issue of supply chain coordination has been studied by many academics (See, for example Cachon 2003, and Kaya and Ozer 2010). The focus of these studies is the characterization of contract terms that determine how the profit and risk will be shared between the firms. Well-organized contracts can coordinate supply chains and can align the incentives of the individual firms, leading to higher overall efficiency and higher gains for all parties, including the end-consumers. In fact, it may even be possible to achieve total coordination within the chain, i.e., the single integrated firm performance, by choosing the right contract parameters.

The common objectives of supply chain contracts are increasing the total supply chain profit, and sharing the profits, risks and information among the supply chain partners. To study contracting under demand uncertainty, supply chain researchers have utilized a simple game theoretical manufacturer-retailer supply chain model where the retailer faces the newsvendor problem. The manufacturer determines the contract parameters of the retailer's problem. If the retailer accepts the contract, she needs to determine how much to order from manufacturer. If she does not accept the contract, both parties earn zero profit.

This simple supply chain illustrates the strategic interaction between the two decision makers. The total supply chain expected profit is a function of the retailer's order quantity; whereas it is the manufacturer's contract offer that determines the parameters of the retailer's decision problem. The retailer and the manufacturer's incentives are not aligned with each other, which may lead to suboptimal profits for both firms. In particular, the manufacturer must design a contract that encourages the retailer to order a quantity that would maximize the manufacturer's expected profit. This may require, for example, sharing some of the risk that the retailer faces.

At the heart of all these models is the newsvendor model. This model, similar to all analytical models, depends on a number of behavioral assumptions about how human beings make decisions. Theory assumes that people are rational decision makers that aim to maximize expected profit level. However, most empirical studies have shown that people do not behave according to what theory predicts. To study the difference between theory and reality, researchers have started conducting "experiments" with human decision makers where human subjects make newsvendor decisions facing a computerized simulation. Using data from such experimental studies, researchers have identified a number of "decision biases" to explain deviations from theoretical predictions.

In this thesis, we aim to explain the ordering decision behavior of retailers in such experiments. Contrary to most literature, our experiments involve human subjects that represent two firms in a supply chain: A manufacturer, who offers a contract, and a retailer who faces the newsvendor problem. This allows us to include the decision biases due to the strategic interaction between two human decision makers that have conflicting incentives.

We consider two different contracts between the firms:

- *Wholesale price contract* (*w*): This contract has only one parameter, which is the wholesale price, *w* that the retailer pays to the manufacturer per unit she orders. Theory states that wholesale price contract causes retailer to order less than supply chain optimum order quantity, which leads to inefficiency.

- *Buyback contract (w, b)*: In a buyback contract, the manufacturer specifies a wholesale price *w* along with a buyback price *b* at which the retailer can return any unsold units at the end of the season. According to theory, the buyback contract can achieve supply chain coordination with a proper combination of the two parameters *(w, b).* Buyback contracts, in theory, may encourage retailers to increase the order quantity, potentially benefiting both firms.

Using data from the experiments of Sahin and Kaya (2011), we aim to answer the following research questions:

- *Do the subjects follow "decision heuristics" while making their decisions?* Schweitzer and Cachon (2000) identified several "decision heuristics" to explain the ordering behavior of retailer subjects in standard newsvendor experiments. We would like to understand whether such heuristics are present in our supply chain experiments where both firms are represented with human decision makers.

- *What factors do retailers consider in setting order quantities?* In addition to following certain decision heuristics, retailer subjects' decisions are also known to be affected by certain <u>irrelevant</u> factors, such as the profit level realized in the previous period or the expected profit share of the proposed contract. To identify the most effective factors, we build linear regression models to capture each retailer's ordering behavior. We identify the independent variables of these regression models through "feature selection" methodology.

- *Do subjects learn to make better decisions over time?* We would like to understand if and how the subjects' decisions change over time due to learning-by-doing.

The rest of the thesis is organized as follows: In Chapter 2, we summarize the related literature. In Chapter 3, we discuss our simple manufacturer-retailer supply chain model, and provide information on the analytical background. In Chapter 4, we present our experimental design and procedure. In Chapter 5, we discuss the decision heuristics. Chapter 6 presents our selection of the factors and regression analyses study. In Chapter 7, we discuss the learning effect in the newsvendor setting. In Chapter 8, we conclude with discussions and future research suggestions.

CHAPTER 2

2 LITERATURE SURVEY

In this chapter we summarize the related literature on the newsvendor model and on supply chain contracting and coordination.

2.1 The Newsvendor Model

The newsvendor problem is an example of decision making in the face of uncertainty. It is traditionally motivated through the story of a newsvendor who needs to determine how many copies of a newspaper to order and stock at the beginning of a day to meet stochastic demand during the day. If demand turns out to be higher than her order quantity, the difference between order quantity and realized demand becomes left over units. If demand turns out to be lower than her order quantity, the newsvendor misses the chance of selling more units, and the absolute difference between order quantity and realized demand becomes lost sales. Addressing the trade-off between ordering too much and ordering too little, Arrow et al. (1951) came up with famous “critical ratio” solution to the newsvendor problem.

Schweitzer and Cachon (2000) conducted the first laboratory study of the newsvendor problem. They observe the subjects’ orders to be pulled away from the optimal quantities towards the mean demand value. In particular, when the critical ratio was below 0.5 (the low profit condition), the subjects’ average order quantity is higher than the optimum order quantity. On the other hand, when the critical ratio is higher than 0.5 (the high profit condition), the subjects’ average order quantity is lower than the optimum order quantity. They refer to this phenomenon as the “Pull to Center (PTC) effect” because in both cases the average orders are biased towards the center of the demand distribution. Schweitzer and Cachon argue that these deviations cannot be explained by risk aversion, risk seeking, prospect theory, or a number of other possible explanations. Instead, they offer the following three heuristics that can explain the observed deviations from the theoretical optimal.

- **Mean anchor heuristic** implies anchoring on mean demand and insufficiently adjusting towards the optimum order quantity.
- **Demand chasing heuristic** implies anchoring on previous order quantity and adjusting towards the previous demand realization.
- **Minimizing ex–post inventory error heuristic** implies regretting from not ordering the previous period's demand realization, even though there was no way to predict it.

The first two heuristics are related to the "anchoring and insufficient adjustment" type heuristics (Kahneman et al. 1982) where people anchor their decisions around some available but irrelevant information, and insufficiently adjust around this value over time. One of the research questions we consider in this thesis research is to understand whether the subjects in our more complicated experiments (due to strategic interaction) also follow the decision heuristics of Schweitzer and Cachon (2000).

Bolton and Katok (2008) observe the pull to center effect in their experiments that consists of three different studies. In the first study, they limit the number of ordering options from 100 to 9 and 3 respectively. They find that limiting the number of ordering options does not improve performance for both high and low profit conditions. In the second study, they show that providing information about the foregone options does not help improve performance. In the third study, they show that forcing the subjects to place ten-period standing orders improves performance. With standing orders, the subjects learn over time by taking long term decisions rather than focusing on short term fluctuations.

Bostian et al. (2008) aim to explain the pull to center effect with an adaptive learning model, that consists of memory, reinforcement, and probabilistic choice elements. They conclude that subjects learn the attractiveness of each order quantity over time based on their past period experiences. Lurie and Swaminathan (2009) find that more frequent feedback about the results of newsvendor decisions does not always improve performance.

Benzion et al. (2008) study the newsvendor problem using two different demand distributions (uniform and normal) and two different marginal profit conditions (low and high). They find that in all cases learning occurs and is affected by the mean demand, the order-size of the maximum expected profit, and the demand level of the immediately preceding period. To capture the effect of the pull to center effect, the authors model the participants' order quantity is a weighted average of the optimal order and the demand distribution mean.

Benzion et al. (2010) study a similar setting with unknown demand distribution. According to their findings demand information does not improve the subjects' profits. They investigate learning and in one of their hypothesis they claim that the personal order level deviation would become smaller over time. As a result of their experiments they show that the absolute change in the order quantity between two consecutive periods is reduced over time. They also used blocks of half periods and compared the subjects' behavior in the first half of the periods block and the last half of the periods block. They claim that, this kind of analysis would emphasize the trend over time, if it exists. They show that the average order in the first half of the periods is significantly different from the average order in the last half of the periods. They conclude that subjects who knew the distribution used their knowledge to improve their order.

Recently, Lau et al. (2014) question the existence of the pull to center effect. They show that while the pull to center effect can be observed in "group average" data, it does not exist in most individual subjects' data. In a similar paper, Lau et al. (2012) question the existence of demand chasing. They show that some methods that researchers use to measure the heuristic (such as adjustment scores) may exaggerate the extent of demand chasing present in data. They recommend the use of simple correlation between the order quantity and the previous period demand realization.

In addition to these heuristics, researchers have also studied the effects of certain factors (most irrelevant) to the retailer subjects order quantity decisions. Next, we briefly mention some of the most important ones of these factors:

Risk aversion: A risk-averse decision maker orders less than the optimum order quantity while a risk-seeking decision maker orders more than optimum order quantity (Eeckhoudt et al. 1995). Prospect theory (Kahneman and Tversky 1979) predicts that people act risk averse in the domain of gains, but risk-seeking in the domain of losses (reflection effect).

Loss aversion: A loss averse decision maker prefers avoiding losing rather than obtaining gains. Wang and Webster (2009) show that when shortage cost is low, a loss-averse decision maker orders less than a rational decision maker; whereas, when the shortage cost is high, a loss-averse decision maker's order quantity is more than a rational decision maker. Loss averse people tend to avoid situations where probabilities are unknown (uncertainty about uncertainty), and order less than the optimum order quantity, because losses result larger disutility than the value derived from the same size of gains (Camerer and Weber 1992).

Framing: Framing describes how the subjects behave when the emphasis on loses and gains change. Shultz et al. (2007) compare a positive newsvendor frame where the gain is emphasized with a negative frame where the loss is emphasized. No difference was detected between the frames. Kremer et al. (2010) compare the results of newsvendor experiments under two frames: In the "operations frame", the subjects simply make the standard newsvendor decisions using a standard newsvendor story. In the "neutral frame", the decisions are the same but the story is not given in the newsvendor context. Rather, it is given in a generic frame. The authors conclude that the neutral frame is closer to the optimal in both low and high profit conditions.

Bounded rationality: Standard economic theory assumes that people rationally choose the "best response" among alternatives. However, in practice, people make noisy decisions. They may make calculation or recording errors due to limited cognitive ability, limited memory and attention span. When faced with complex decision situations, people may resort to decision heuristics as shortcuts. Su (2008) indicates that the pull to center effect observed in newsvendor experiments can be explained by bounded rationality with a quantal response equilibrium framework. The author concludes that subjects do not always make the best decision, but the good decisions are more likely to occur rather than the worse ones. Gavirneni and Isen (2010) use a verbal protocol analysis to understand the logic behind the decision makers' decisions in newsvendor game. They conclude that most subjects were successful in calculating underage and overage costs but failed to transform them into optimum order quantity. This finding suggests that the newsvendor problem may not be as intuitive as thought by researchers.

Irrational Behavior: Becker-Peth et al. (2013) show that orders can be predicted accurately even human subjects is irrational. They derive response functions for mean orders, variance of orders and expected profit to predict actual human behavior, and use the models to design supply chain contracts instead of newsvendor model. The authors show that the order quantity not only depends on the critical ratio but also on the wholesale price and buyback price. They conclude that the model they derived is quite better than the newsvendor model.

Overconfidence: Croson et al. (2013) find that overconfident decision makers make suboptimal decisions in the newsvendor problem. Bolton et al. (2012) compare the performance of undergraduate students, master students and managers in the newsvendor game. The authors conclude that managers do not perform better than two student groups and students, especially graduates, are better in using the given information that helps find the optimum solution.

Cultural differences: Feng et al. (2010) conducted experiments in order to analyze the cross–national differences between Chinese and American subjects. The results show that Chinese subjects' decisions are more anchored to mean demand than American subjects. The authors also re-examine "thinning set of orders" (Bolton and Katok 2008). They show that when the optimum order is one of the middle options not the extreme one, supply chain efficiency increases and the percentage of choosing the optimum order quantity increases.

Gender Differences: Vericourt et al. (2013) investigate the effect of gender differences in newsvendor game. They measure whether there are significant gender differences in ordering behavior in the newsvendor problem. They conclude that in low profit condition, there is no significant difference between males and females, but in high profit condition, men tend to have greater risk appetite and tend to order more than women.

2.2 Supply Chain Contracting and Coordination

In Section 2.1, we have discussed the newsvendor problem which is related to the order quantity decision of a single decision maker. Here we discuss the literature on supply chain coordination, which deals with the decisions of at least two firms (i.e., decision makers) that are in a strategic relationship with each other. Each member of the chain aims to maximize its own profit. This decentralized decision structure leads to suboptimal total chain profit, as in the case of the famous double marginalization problem (Spengler 1950).

Supply chain contracting literature mainly focuses on how different contract types can be used to align the incentives of the different chain members, which are referred to as the "coordination" of the chain. Coordination allows the total expected supply chain profit to be maximized, and be equal to that of an integrated firm. The contract also determines how the total profit and risk due to uncertain demand will be shared between the chain members. Most popular contract types in the literature include the buyback (Pasternack 1985), revenue sharing (Cachon and Lariviere 2005), and quantity flexibility (Tsay 1999) contracts.

Pasternack (1985) shows that it is possible for a manufacturer to determine a returns policy (buyback contract) that achieves channel coordination. If the manufacturer allows only partial returns, selling price and return policy becomes a function of the retailer's order quantity; whereas, if the manufacturer can buy back all unsold units (an unlimited return policy) then the return policy is independent from retailer's order quantity decision. Emmons and Gilbert (1998) also analyze return policies and find what combination of wholesale price and return policy maximizes manufacturer's expected profit. They conclude that retailer price increases with increased uncertainty and manufacturer gains more profit with buying back unsold units from the retailer.

Kandel (1996) studies different types of contracts that try to allocate the risk between manufacturer and retailer for the unsold inventory. The author shows that manufacturers prefer consignment contracts, where retailers prefer the no return contract.

Next, we outline the experimental/behavioral works on supply chain contracting:

Keser and Paleologo (2004) conducted a laboratory experiment that investigates the simple wholesale price contract. The average wholesale price is observed to be lower than the optimum. Retailers order lower than the optimum order quantity to a given wholesale price. No evidence is found to support Schweitzer and Cachon’s pull to center effect and chasing demand heuristic. Supplier’s realized profit is lower but retailer’s realized profit is higher than then theoretical prediction, which implies a more balanced profit distribution.

In this thesis we aim to understand the reason why the retailers deviate from the optimal newsvendor solution, by using Keser and Paleologo’s parameter setting as our base model. In addition to aggregate-level analysis, we also analyze each retailer’s decision individually. To understand what factors the retailer subjects consider in their order quantity decisions, we apply feature selection to each individual decision maker’s quantity decision (Guyon and Elisseeff 2003).

Katok and Wu (2009) conducted a laboratory experiment that compares buyback contract, wholesale contract and revenue sharing contract to each other. In order to eliminate human decision maker’s biases, human retailers play the game with a computerized supplier, and human suppliers play the game with a computerized retailer. The authors find revenue sharing and buyback contracts to perform better than the wholesale price contract but fail to achieve channel coordination. Retailers’ decisions are more likely show minimizing ex post inventory error than anchoring and insufficiently adjustment heuristic. The difference between buyback and revenue sharing contracts stems from framing of contract types diminishes over time.

Wu (2013) studies the impact of repeated interactions on supply chain contracts by comparing the wholesale price, buyback price and revenue sharing contracts. The author observes that buyback contracts behave differently from revenue sharing contracts by inducing higher order quantities over time and also finds that the behaviors of both the retailer and the supplier deviate from the predictions of the traditional contracting model. The results of the study imply that various contracts can perform differently based upon how the bargaining is distributed within a channel.

Hyndman et al. (2012) consider a two-firm supply chain where the sales are constrained by the capacity choice that each firm makes simultaneously before demand realization. The authors analyze the difference between fixed and random matching on coordination between players. Fixed matching setting is similar to our long-run experiments, and random matching is similar to our short run experiments. The efficiency of fixed match is found to be higher in initial periods, but the situation gets reversed at the last five periods of the game, which is counterintuitive. This is explained by the first impression bias. Learning is also found to be slower under fixed matching.

There are also papers in literature where the retailer faces deterministic demand, hence, is not a newsvendor. These papers are related to our work in that they also study behavioral issues between supply chain members. Lim and Ho (2007) study the effect of the number of blocks in a contract. A two-block tariff contract is found to increase supply chain efficiency more than a linear price contract; however, the increase in efficiency is lower than expected. If the numbers of blocks increase to three, supply chain efficiency improves further, and the manufacturer's profit share increases. The authors propose a Quantal-Response Equilibrium (QRE) model to explain the counterintuitive results, and to better understand the retailer's sensitivity to counterfactual profits.

Loch and Wu (2008) study the effect of social preferences on supply chain coordination. Social preferences refer to concerns about the other firm's welfare, reciprocity stem from positive relationship, and desire of a higher relative payoff compared to the other firm when the status is salient. Lock and Wu's experiments compare the "control condition" in which players are given simple incentives only, "relationship condition" in which the players are assumed to have a friendship, and "status seeking condition" in which players are assumed to compete with each other. In the relationship condition, both parties are found to set prices lower than optimum, and in status seeking condition, both parties set selling prices higher than optimum. Hence, there is evidence that individuals' preference for social relationships may lead to higher than expected cooperation leading to higher profits; whereas, preference for status may lead to destructive actions, leading to inefficiency.

Cui et al. (2007) discuss how fairness concerns may help achieve channel coordination. Using analytical model, the authors show that supply chain coordination can be achieved even with a simple wholesale price contract when the parties are sufficiently concerned about fairness.

Haruvy et al. (2012) compare coordinating contracts such as two part tariff (TPT) and minimum order quantity (MOQ) to wholesale price contract under two different bargaining structures: In Ultimatum Bargaining (UB) the least possible bargaining power is given to the retailer, whereas in Structured Bargaining (SB), retailer has a bargaining power. Results show that under UB, only TPT contract is more efficient than wholesale price contract but under SB, both TPT and MOQ contracts are more efficient than the wholesale price contract. Structured bargaining achieves nearly full channel efficiency.

CHAPTER 3

3 ANALYTICAL BACKGROUND

3.1 The Supply Chain Scenario

We consider a typical supply chain scenario, where a manufacturer produces a certain product at a ***unit production cost of c***, and a retailer buys this product from the manufacturer and sells it to its customers (consumers) at a ***sales price of p***. Consumer demand is probabilistic with cumulative distribution function F(.).

The sequence of events is as follows: At the beginning of the relation, the manufacturer sets the parameters of the contract and offers the contract to the retailer. One of the contract parameters is the ***wholesale price, w***. Manufacturer sells his product to the retailer at this price. Depending on the contract type, the contract may include other parameters. If the retailer's expected profit is positive, she accepts the contract. In this case, she chooses her ***order (stock) quantity q***, and orders this quantity from the manufacturer. The manufacturer produces and delivers these units to the retailer before the selling season. If the retailer's expected profit with her optimal order quantity is negative, the retailer rejects the contract.

During the sales season, ***random consumer demand*** is realized as ***"D"***. The retailer tries to satisfy this demand by using her stock of product. The ***sales quantity*** of the retailer is the minimum of her order quantity *q* and the realized demand. Two cases are possible:

- If the realized consumer demand turns out to be less than the retailer's order quantity (i.e., $D<q$), the retailer will sell *D* units, and *(q-D)* products will be unsold (***leftover products***). These products have zero salvage value.

- If the realized consumer demand turns out to be higher than the retailer's order quantity (i.e., If $D>q$), the retailer will sell all q units, and $(D-q)$ units of demand will be unsatisfied (***unsatisfied demand***). There is no extra penalty for unsatisfied demand to either firm; however, the firms lose the opportunity to make more profit.

Each firm tries to maximize its own *expected profit* in the game. Note that, the expected profit of each firm depends not only on its own decision, but also on the other firm's decision and also on the random demand. Thus, there exists a strategic interaction between the two firms.

The sequence of events can be summarized as follows:

1. The manufacturer offers the contract, by specifying its parameters.
2. If the retailer's expected profit level is non-negative, the retailer accepts the contract and determines her order quantity q.
3. The manufacturer produces q units at a cost of c each, and sends these to retailer.
4. Sales period arrives and the random consumer demand D realizes at the retailer.

3.2 Supply-Chain Optimal Solution

We first determine the supply chain optimal solution before discussing the solutions under different contract types. In this scenario, a single decision maker makes all decisions with the objective of maximizing the total supply chain (manufacturer + retailer) expected profit. The supply chain's problem is formulated as

$$\underset{q}{maximize}\ \ \pi_{total}^{sc}(q) = pE[min(q,D)] - \ cq\ .$$

This is also a newsvendor problem, but this time it is faced by the whole supply chain. Note that the contract parameters are irrelevant for the supply chain's problem because contract decisions are between the firms of the supply chain. The order quantity that maximizes the supply chain's expected profit is calculated as:

$$q^{sc} = F^{-1}\left(\frac{c_u}{c_u+c_o}\right) = F^{-1}\left(\frac{p-c}{p}\right). \tag{1}$$

The supply chain's expected profit with order quantity q^{sc} is equal to

$$\pi_{total}^{sc}(q^{sc}) = pE[min(q^{sc}, D)] - cq^{sc}. \tag{2}$$

In this thesis, we study *decentralized* supply chains where the manufacturer and the retailer are independent firms. Each independent firm considers only its own profit margin in making decisions, not the total supply chain profit margin. Such decentralized decision making results in inefficiencies, known as the "double marginalization' problem (Spengler 1950).

From the equation above, we observe that the supply chain expected profit is a function of the retailer's order quantity decision q. The supply chain achieves its theoretical maximum expected profit when the retailer chooses q^{sc}. This maximum profit level is known as the *integrated firm profit,* or the *centralized solution.*

The ratio of the total (manufacturer + retailer) expected profit level under a contract to the integrated firm profit is known as *contract efficiency*. A contract that achieves 100% efficiency is said to be *coordinating* the supply chain. In this case, the incentives of the firms are aligned, and inefficiencies due to double marginalization are eliminated. Coordination requires the retailer to choose the integrated firm order quantity q^{sc}. Any other order quantity choice will cause suboptimal total expected profit level in the supply chain.

While the retailer's order quantity decision determines the total supply chain profit, the manufacturer's contract parameter decision has three functions:

- They affect the retailer's order quantity q, which in fact determine the total expected supply chain profit.
- They determine how the total supply chain profit will be shared (in expectation) between the two firms.
- They determine how the risk due to uncertain consumer demand will be shared (in expectation with respect to random demand) between the two firms.

We present the theoretical solution for a given customer demand distribution with cdf F(.). In our experiments, consumer demand is uniformly distributed between (D_{max}, D_{min}). For this distribution, one can further characterize the optimal order quantity of the retailer as

$$q^{sc}(contract) = \left(\frac{c_u}{c_u + c_o}\right) * (D_{max} - D_{min}) + D_{min}. \tag{3}$$

Next, using q^**(contract)*, one determines the optimal contract parameters of the manufacturer by solving the manufacturer's problem. Similar to standard game-theoretical models, the manufacturer is assumed to foresee the retailer's q^**(contract)* choice for any contract offer. That is, the manufacturer can solve the retailer's problem. Taking the retailer's q^**(contract)* reaction into account, the manufacturer determines the contract parameters that maximize his own expected profit.

The manufacturer's objective function is in general not jointly concave in the contract parameters. Hence, one cannot find a closed form solution for the manufacturer's problem. Instead, one can use a numeric procedure to determine the manufacturer's optimal contract parameters through a grid search over possible parameter combinations. Using these contract parameters, one can then calculate the retailer's order quantity, expected sales quantity, and the expected profits of the two firms. These values characterize the outcome of the game for the given values of model parameter.

Next, we derive the solution of the game separately under the wholesale price and buyback contracts.

3.3 Wholesale Price Contract (WSP) Model

This contract only has one parameter, the wholesale price, *w*. Theory states that the wholesale price contract causes the retailer to order less than supply chain optimum order quantity, which leads to inefficiency. Given the contract (*w*), the retailer's problem is

$$\underset{q}{maximize}\ \pi_r^w(q) = pE[min(q, D)] - wq.$$

From the standard critical fractile solution, the retailer's optimal order quantity satisfies

$$q^w(w) = F^{-1}\left(\frac{p-w}{p}\right). \tag{4}$$

Comparing Equations (1) and (4) we observe that the wholesale price contract cannot coordinate the supply chain unless the manufacturer sets $w=c$. Such a choice is unlikely because it yields zero expected profit to the manufacturer. Having only one parameter, this contract type fails to align the incentives of the two firms.

For uniformly distributed demand, the retailer's unique order quantity solution becomes

$$q^w(w) = \begin{cases} D_{max} - \frac{w(D_{max}-D_{min})}{p} & if w < p \\ 0 & if w \geq p \end{cases}.$$

Substituting $q^w(w)$, the manufacturer's problem becomes

$$\underset{w}{maximize}\ \pi_m^w = (w-c)\, q^w .$$

The objective function of the manufacturer is quadratic and concave in the interval $[c, p]$ and is equal to zero if $w>p$. Manufacturer's optimal wholesale price is found as

$$w^w = \min\left\{p,\ \frac{c}{2} + \frac{p}{2}\frac{D_{max}}{D_{max}-D_{min}}\right\}.$$

Alternatively, one may use a numeric procedure to determine the manufacturer's optimal wholesale price, w^w, through a grid search over possible w values. Using this wholesale price, one can then calculate the retailer's order quantity, expected sales quantity, and the expected profits of the two firms.

3.4 Buyback (BB) Contract Model

Under a buyback contract, the manufacturer specifies a wholesale price w along with a buyback price b at which the retailer can return any unsold units at the end of the season. By this contract, the manufacturer reduces the retailer's cost of overage, encouraging the retailer to set a higher order quantity. The retailer's problem becomes,

$$\underset{q}{maximize}\ \pi_r^b = pE[\min(q,D)] + bE[q - \min(q,D)] - wq$$
$$= (p-b)E[\min(q,D)] - (w-b)q.$$

The retailer's cost of overage becomes w-b while the cost of underage is p-w. Accordingly, the retailer's optimal order quantity is found as:

$$q^b(w,b) = F^{-1}\left(\frac{c_u}{c_u+c_o}\right) = F^{-1}\left(\frac{p-w}{p-b}\right). \tag{5}$$

Comparing Equations (1) and (5), one can show that the supply chain will be coordinated if the buyback contract parameters satisfy $b = \frac{p(w+c)}{p-c}$.

Substituting $q^b(w,b)$, the manufacturer's problem becomes

$$\underset{w,\,b}{maximize}\ \pi_m^b = (w-c)q^b - bE[q^b - min(q^b,D)].$$

This function is not jointly concave in w and b. One can determine the optimal contract parameters through a grid search over the w and b values.

3.5 Our Parameter Setting and Solution

We consider the following model parameter values:

- Unit production cost, $c = 50$.
- Retail price, $p = 250$.
- Zero salvage value.

- Demand uniformly distributed between 40 and 230, and can take only integer values.
- The decision variables (w, b, q) are expected to take only integer values.

This parameter setting is the same as the one used by Keser and Paleologo (2004). Given these parameters, the manufacturer's wholesale price satisfies $0 \leq w < p = 250$. For a chosen w, the buyback price satisfies $0 \leq b \leq w$. The subgame-perfect equilibrium solutions under the two contracts are summarized in table below.

Table 3.5.1: Comparison of Manufacturer's Optimal Solution under Two Contracts

Contract Type	**Total Profit**	**Contract Efficiency**	**Mfg. Profit**	**Retailer Profit**	**w**	**b**	**q**
Buyback	23,123	98.50%	22,784	333	247	246	183
Wholesale Price	17,137	74.00%	12,126	5,011	176	--	96

We observe that the manufacturer's optimal solution under the buyback contract dominates the solution under the wholesale price contract in terms of total profit. This is primarily due to differences between the retailer's order quantities. In fact, the efficiency of the buyback contract is close to 100%, which is good news from the supply chain point of view. However, the profit distribution under this contract is quite unbalanced. Almost all profit is going to the manufacturer. The wholesale price contract, on the other hand, while inefficient, offers the retailer a decent profit level.

Note that this is only a theoretical comparison which assumes that (1) the retailer will accept any contract that provides her nonzero expected profit; (2) the retailer will determine her order quantity according to the newsvendor formula; (3) the manufacturer will be able to foresee the retailer's reaction to any contract offer. As we will discuss in this thesis, these assumptions are questionable when real human beings make decisions.

CHAPTER 4

4 EXPERIMENTAL DESIGN AND PROCEDURE

In this chapter we present the experimental design and experimental procedure. We use the data of experiments that were conducted by, and reported in Sahin and Kaya (2011).

4.1 Experimental Design

The experimental design is illustrated in Table 4.1.1, where *n* denotes the number of subjects. Two different contract types (wholesale price and buyback contracts) and two relationship length types (long run and short run) were studied[1]. In the long run experiments, the same manufacturer-retailer pair interacts throughout all 30 periods, whereas in the short run experiments, the pairs are re-determined in each period.

Table 4.1.1: Experimental Design and Number of Subjects

		Contract Type	
		Buyback	**Wholesale price**
Relationship Length	**Long run**	Experiment b1a, n=12 Experiment b1b, n=16	Experiment w1a, n=16 Experiment w1b, n=14 Experiment w1c, n=16
	Short run	Experiment b2a, n=16 Experiment b2b, n=16	Experiment w2a, n=14 Experiment w2b, n=16 Experiment w2c, n=14

[1] WL will refer to the wholesale price contract, long run experiments, WS to the wholesale price contract, short run experiments. BL and BS denote the counterparts for the buyback contract.

4.2 Experimental Procedure

Sahin and Kaya conducted their computer-based experiments at the CAFE (Center for Applied Finance Education) computer laboratory of Sabancı University, Faculty of Management. They coded[2] and implemented the experimental model using HP MUMS Software.

Subjects were selected from Sabancı University MS 401 course spring semester 2010/2011 students. These students had already studied the basic newsvendor problem. To provide incentive, each subject's total profit at the end of the experimental session was converted into a bonus grade for the course MS 401. The bonus ranged between 1% and 2.5%, and it is applied to the final grade of the subject in that course.

Instructions were delivered to subjects before they arrive at the laboratory. Sample instructions are provided in Appendix C. At the beginning of each session, instructors explained the experiment once again to ensure that the instructions are clearly understood, and they answered any remaining questions. Before starting the actual experiment, they let the subjects play three pilot (training) periods. During the actual experiments, they did not allow the subjects to communicate with each other. Each experimental session took around two hours.

Each experimental session contained one experiment (treatment) composed of 30 independent periods (rounds). Throughout a given experiment, a particular subject played the role of either manufacturer or retailer. The role was randomly assigned at the beginning of the experiment and remained unchanged in all of the 30 periods.

The term "game" denotes the interaction in a manufacturer-retailer pair in a given period. The sequence of events in the game reflects the three stage interaction in the analytical model. At stage I of the game, the manufacturer sets the contract parameters wholesale price and buyback price (in buyback contract experiments). At stage II, these contract parameters are displayed on the retailer's screen and the retailer determines her order quantity. At stage III, random consumer demand is realized. The results of the

---

[2] Appendix A provides the main script code that is used to define the number of subjects, and to call other functional scripts, as an example. Appendix B illustrates another important part of the code where the parameters, stages and the allocation strategy of subjects to the roles are defined.

game are then reported to the subjects. Each subject is given around 30 seconds to make his decision.

Appendices D and E provide sample screenshots of the manufacturer and the retailer's screens respectively in the buyback contract experiments. The large table in the middle of the screen is the "decision support tool". By using this tool, the subjects could run what-if analysis before submitting their decisions. A retailer subject can enter an order quantity to this tool and obtain the outcome for eight different realizations of the stochastic consumer demand (For $D = 40, 70, 100, 130, 160, 190, 220, 230$). The manufacturer also has a decision support tool. However, he needs to enter contract parameters (w, b), as well as a value for the retailer's order quantity decision to the tool. More detailed explanation about the decision support tool can be found in Appendix C.

Subjects enter their decisions into the cells at the bottom of the screens. At the end of each period, a results screen (as seen in Appendix F) provides the subjects with the results of their game. The game results include the consumer demand realization, the decisions of both firms, number of units sold, and number of units unsold, demand unsatisfied, the period profit and cumulative profit of both firms. These results are provided for all periods up to and including the last period.

After each experiment, a post-experiment survey is conducted where they asked the subjects how they made their decisions, whether they were motivated by the bonus grade and their suggestions. These surveys indicated that the subjects were highly motivated for their decisions, and their responses yield clues about their decision heuristics.

4.3 Experimental Data Analysis

Recall that the outcome of a game is shaped by first the manufacturer's decision, second the retailer's decision and third the realization of random consumer demand. We use the following terms to differentiate the predictions at different levels:

1) Manufacturer's optimal outcome: The subgame-perfect equilibrium of the model corresponds to the manufacturer's optimal outcome (in each period). This is because the manufacturer is the first-mover in the game. In this outcome, the manufacturer offers the contract ($w^*=247$, $b^*=246$), and the retailer stocks the corresponding newsvendor quantity $q^*(w^*, b^*) = 183$. Manufacturer's expected profit is 22,790 and retailer's expected profit is 333. This is what the theory predicts as the outcome of the overall interaction between the two firms in a given period.

2) Newsvendor's predicted outcome: In experiments, manufacturer subjects do not necessarily offer their optimal contract (w^*, b^*). We define the "predicted outcome" as the expected outcome of the interaction given any contract (w, b), assuming that the retailer chooses the newsvendor order quantity $q^*(w, b)$. The difference between the "predicted outcome" and real experiment data is due to the retailer's deviation from the newsvendor model, and due to the realization of random demand.

3) Expected outcome: Retailer subjects also often deviate from the newsvendor order quantity decision. For any contract (w, b) and retailer's response $q\ (w, b)$, the "expected outcome" denotes the expected result with respect to consumer demand distribution.

Next, we present our results. We aim to answer the following research questions:

- Do retailer subjects follow certain decision heuristics while making their decisions?
- What factors do retailers consider in setting their order quantities?
- Do subjects learn to make better decisions over time?

CHAPTER 5

5 RETAILER’S DECISION HEURISTICS

Human decision makers are known to employ decision heuristics. These heuristics can have considerable effect in shaping managerial behavior (Bazerman 2008). Retailers in standard newsvendor experiments are known to use two such heuristics: Mean anchoring heuristic and demand chasing heuristic. Schweitzer and Cachon (2000) showed that the well known “pull to center effect” can be explained by either of these heuristics. In both cases, experimental order decisions are “pulled” towards the mean of the demand distribution, away from the optimal newsvendor quantities.

In this study, we aim to understand whether the retailer subjects in our experiments followed these two heuristics, and whether they exhibit the pull to center effect.

In standard newsvendor experiments, in all periods the retailer faces the same contract offered by the computer. Our experimental setting differs in two respects:

- The optimal order quantity (q^*) for the retailer’s problem changes from one period to the other based on the offered contract.
- The strategic relationship between the manufacturer and retailer players affects the retailer’s quantity choice. The retailer, for example may set a substantially low order quantity to “warn” the manufacturer for offering a bad contract. She may even order the minimum possible quantity or reject the contract.

Due to these differences, measuring the effects of the decision heuristics on retailer’s order quantity in our experimental setting is a difficult task.

5.1 The Pull to Center Effect

Extensive research has demonstrated the existence of the pull to center effect in empirical newsvendor behavior (Schweitzer and Cachon 2000, Benzion et al. 2008, Bolton and Katok 2008, Bostian et al. 2008, Lurie and Swaminathan 2009, Kremer et al. 2010). This implies that instead of ordering the optimal order quantity (q^*), subjects order a quantity between q^* and the mean demand value $\bar{d}$. We will refer to this region as the “pull to center zone” (PTC zone). Note that the PTC zone in our experiments will be re-defined at each period based on the q^* value that is implied by the offered contract.

Figure 5.1.1 illustrates the data of a retailer subject from our experiments who exhibits significant pull to center behavior. Note how the subjects’ order quantities are pulled towards the mean demand and away from the optimal order quantity in most periods.

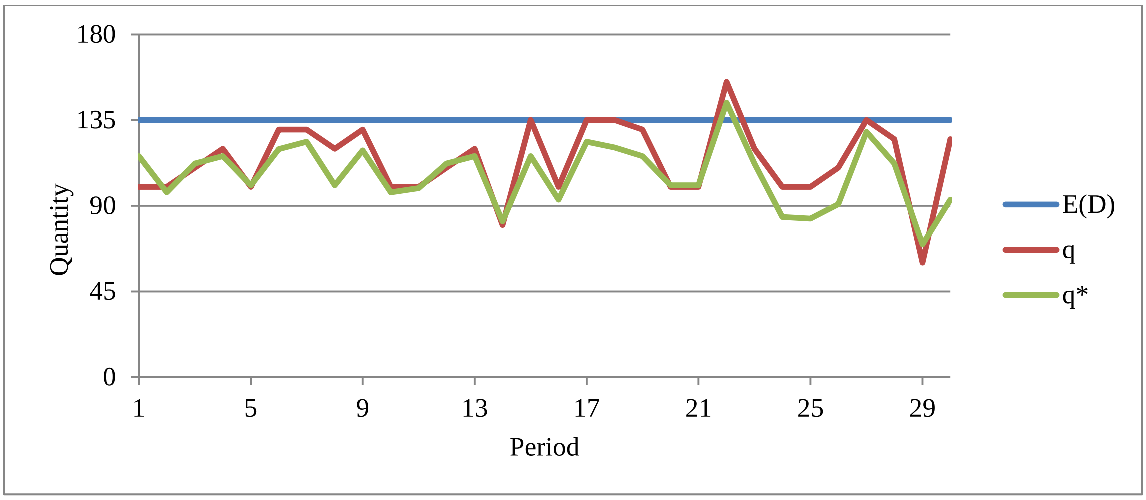


Figure 5.1.1: A Subject that Illustrates Pull to Center Effect

5.1.1 Percentage of Orders in PTC Zone

To understand whether our retailer subjects exhibited pull to center behavior, we check the percentage of orders that fall into the PTC zone (Similar to Lau et al. 2014).We ignore the data of periods in which $q=0$ or $q^*= \bar{d}$. We calculated this percentage for each retailer in each experiment type separately. The resulting histograms are presented in Figure 5.1.2. We observe that in all experiment types, the percentage of orders that fall into the PTC zone is quite small for most retailers. From Figure 5.1.3 and Figure 5.1.4 we also observe that in the long run experiments the percentage of orders in PTC zone are generally higher than the short run experiments, and in buyback experiments the percentage of orders in PTC zone is smaller than the wholesale price experiments.

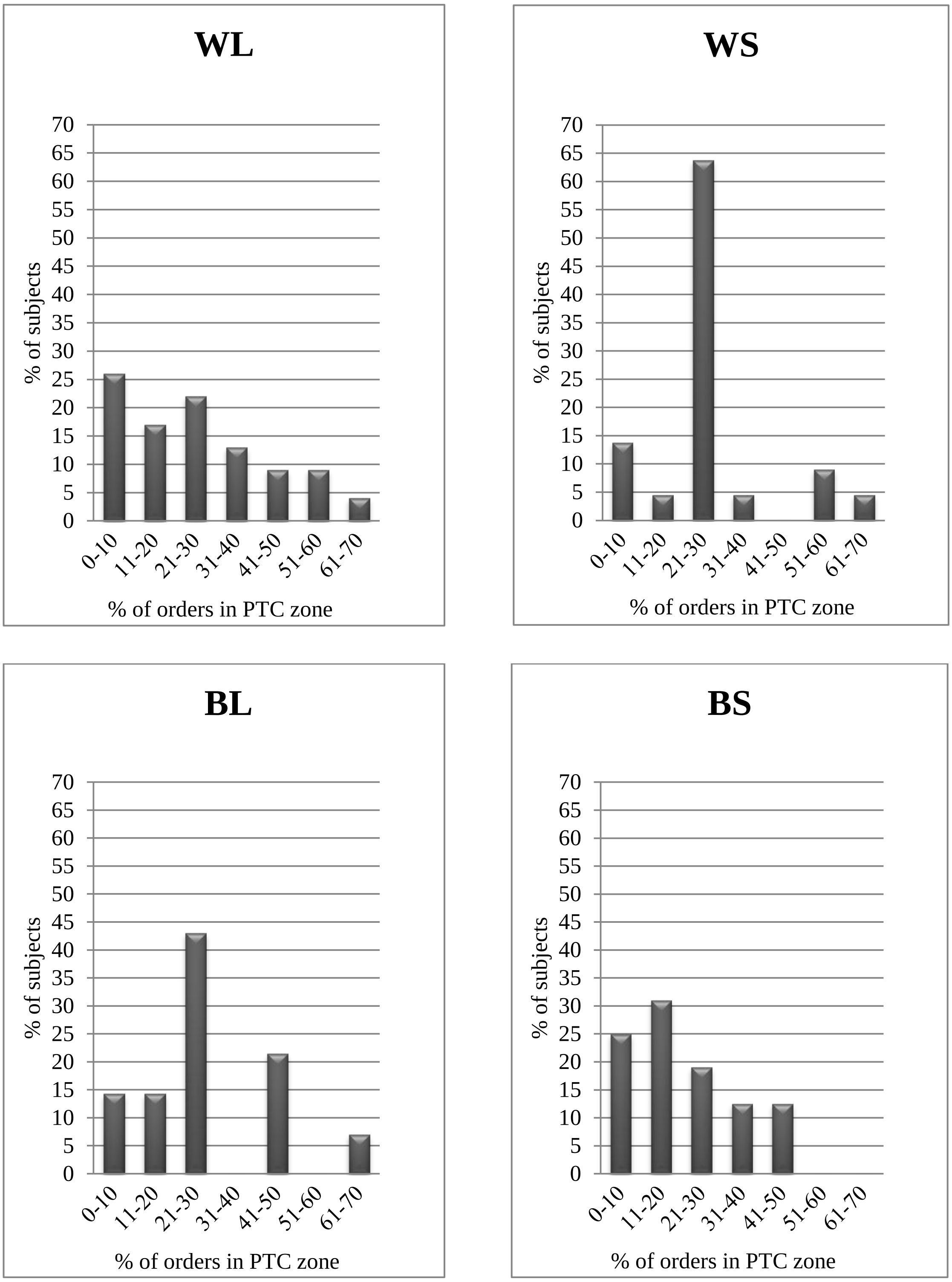


Figure 5.1.2: Percentage of Orders in the PTC Zone

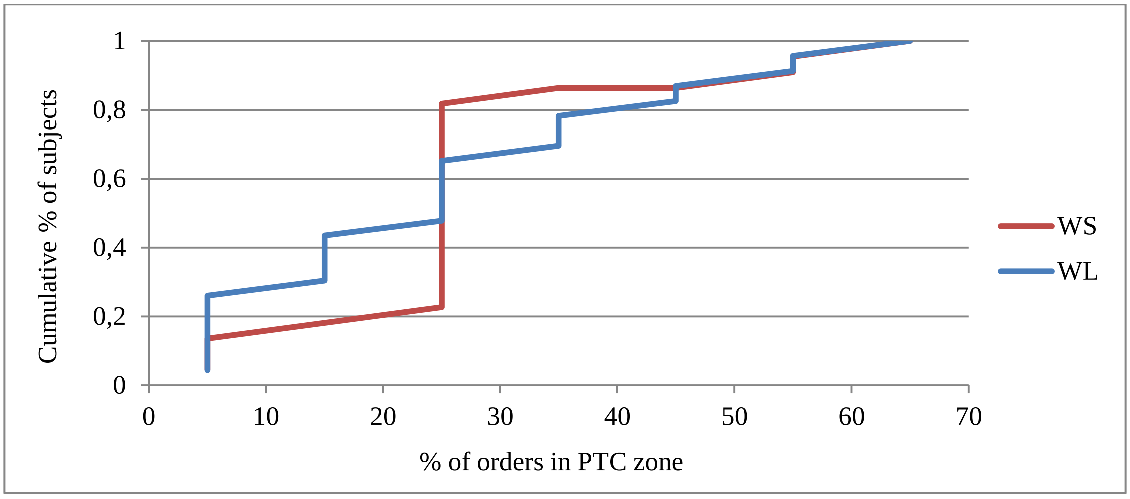


Figure 5.1.3: Cumulative Distribution of Percentage of Orders in PTC Zone in Wholesale Price Experiments

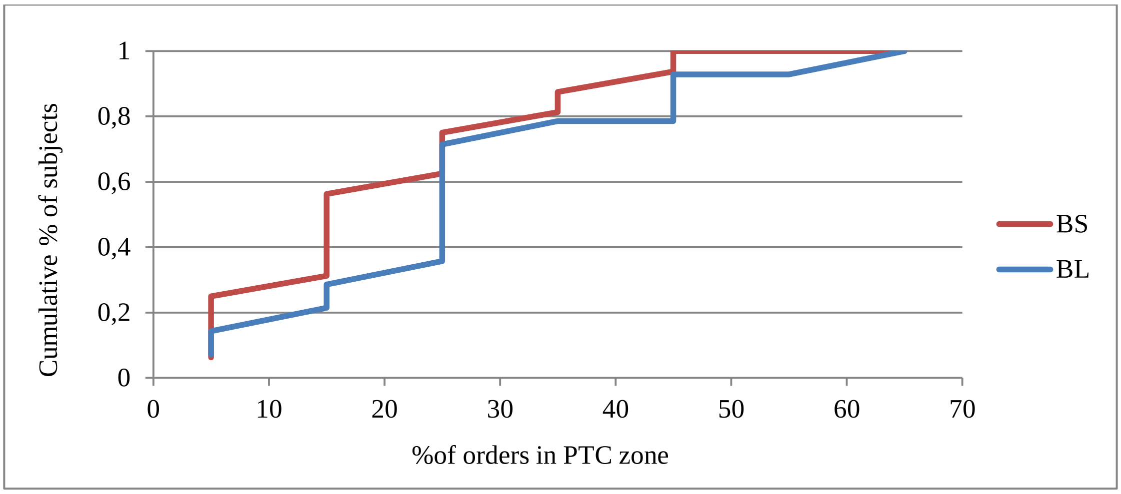


Figure 5.1.4: Cumulative Distribution of Percentage of Orders in PTC Zone in Buyback Experiments

5.1.2 Regression-based Analysis

Another way to test whether subjects exhibit the PTC effect is through the following regression equation (similar to Bostian et.al 2008).

$$q_t = \bar{d} + \alpha(q^* - \bar{d}) + \varepsilon_t \tag{6}$$

Here, the parameter "α" reflects the extent that the subjects deviate from the mean demand toward the optimal order quantity. To be consistent with the pull to center effect, this parameter should fall into the interval (0,1). We determined the "α" values for each individual retailer subject separately by regressing ($q_t - \bar{d}$) on ($q^* - \bar{d}$). Figure 5.1.5 shows the cumulative distribution of α-coefficient obtained from individual regressions.

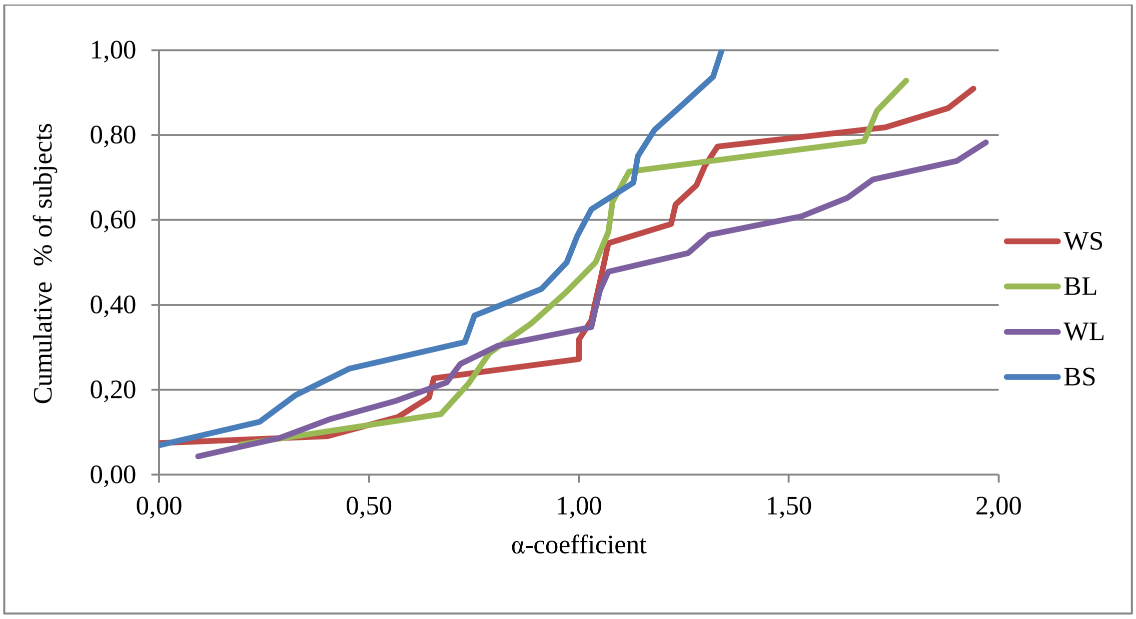


Figure 5.1.5: Cumulative Distribution of α-coefficient

We find 70% of subjects in WL, 82% in WS, 57% in BL, and 50% in BS experiments to have α-coefficients outside the interval (0,1) which is not consistent with the pull to center effect. In addition, by checking the p-values of the coefficients, only 4% of subjects in WL, 14% in WS, 29% in BL, and 31% in BS with α-coefficients inside (0,1) are found to be significant at 0.05 level.

We also conduct an individual-effects regression study based on aggregate data at experiment level. Table 5.1.1 presents the results. Contrary to the individual level analysis, the aggregate data of each experiment type except the WS experiments, is found to have α-coefficients inside (0,1) and is significant at the 0.05 level. Table 5.1.1 indicates the aggregate α-coefficients.

Table 5.1.1: Aggregate Regression Study Results

Experiment Type	α-coefficient	p-value	Adj. R^2	Number of Data Points	Proportion of Retailers with Significant Regression Equation
WL	0.90	0.00	38.75	637	0,70
WS	1.02	0.00	49.12	643	0,55
BL	0.99	0.00	46.78	380	0,79
BS	0.79	0.00	37.65	450	0,50

These findings imply that one should be careful in discussing pull to center results reported in literature. While PTC can be claimed to exist based on aggregate level data, the individual level analysis can tell a different story. Aggregate averages can be misleading; therefore the pull to center effect does not accurately describe individual behavior. This finding parallels the results of Lau et al. (2014) that were obtained for a standard newsvendor experiment.

5.1.3 The Difficulty of Observing the Pull to Center Effect

Next, we discuss why it is difficult to observe the PTC effect at individual level in our experimental setting. Recall that the PTC effect requires the order quantity to fall between the mean demand (which is 135 in our study), and the optimal order quantity which is contract-dependent. This PTC zone is a quite wide interval in most standard newsvendor experiments. For example, in Schweitzer and Cachon (2000), the region is from 150 to 225, whereas in Bolton and Katok (2008), it is from 50 to 75 in high profit experiments.

The difficulty in our study is that our variable size of PTC zone is quite narrow. This is because the optimal q^* value turns out to be close to the mean demand value of 135 for most contract offers. This is shown in Figure 5.1.6 which plots the distribution of optimal order quantities in all offered contracts (2250 data points).

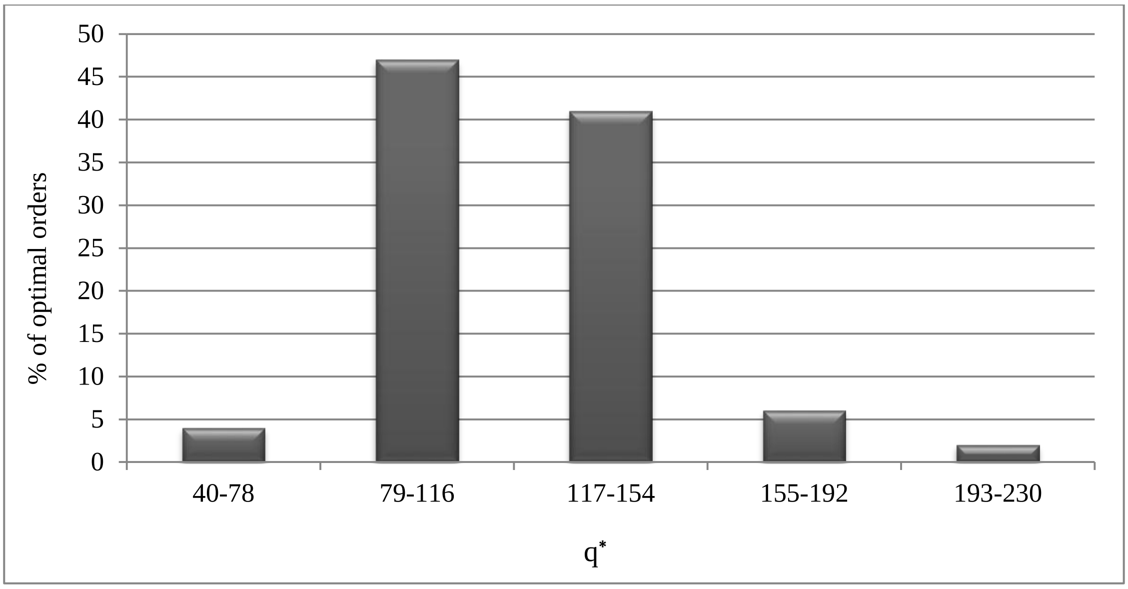


Figure 5.1.6: Distribution of Optimal Order Quantities

Next, we introduce two heuristics that can lead to the pull to center effect, and discuss if we can find evidence for these heuristics in our data.

5.2 The Mean Anchor Heuristic

Under the mean anchor heuristic, the retailer first sets an order quantity (i.e., anchors) around the mean demand value ($\bar{d}$), and over time adjust towards the optimal order quantity (q^*) both in high and low profit conditions.

5.2.1 Counting Changes Anchoring and Adjustments

We count the number of periods, where the retailer's order quantity is between the optimal order quantity (q^*) and the mean demand ($\bar{d}$). We ignore the cases in which $q_t=0$ or $q^*=\bar{d}$. From Table 5.2.1, we observe that the proportion of anchoring periods at the aggregate level is well below the 50% level for all experiments. The subject level results are given in Figure 5.1.2.

Table 5.2.1: The Anchoring Results at Aggregate Level

Experiment Type	**The Aggregate Proportion of Anchoring**
WL	27%
WS	27%
BL	27%
BS	21%

5.2.2 Adjustment Scores

Here we present a method that can measure the magnitude of the heuristic in the periods where anchoring is observed. Following Schweitzer and Cachon (2000), we define the adjustment scores as $(q-\bar{d})/(q^*-\bar{d})$ in the high margin condition, and as $(\bar{d}-q)/(\bar{d}-q^*)$ in the low profit cases.

We ignore the data of periods in which $q_t=0$ or $q^*=\bar{d}$. Figure 5.3.1 shows the percentage distribution of adjustment scores values in periods. We observe that the average adjustment scores of anchoring periods to be around 0.5.

5.3 Demand Chasing Heuristic

Under the demand chasing heuristic, the retailer adjusts his order quantity towards the demand realization in the previous period. Newsvendor behavior literature has suggested that individuals engage in demand chasing (Schweitzer and Cachon 2000, Benzion et al. 2008, Bostian et al. 2008, Lurie and Swaminathan 2009, Kremer et al. 2010).

In a standard newsvendor experiment where q^* is fixed, this heuristic predicts that

- $q_t > q_{t-1}$ if $d_{t-1} > q_{t-1}$
- $q_t < q_{t-1}$ if $d_{t-1} < q_{t-1}$

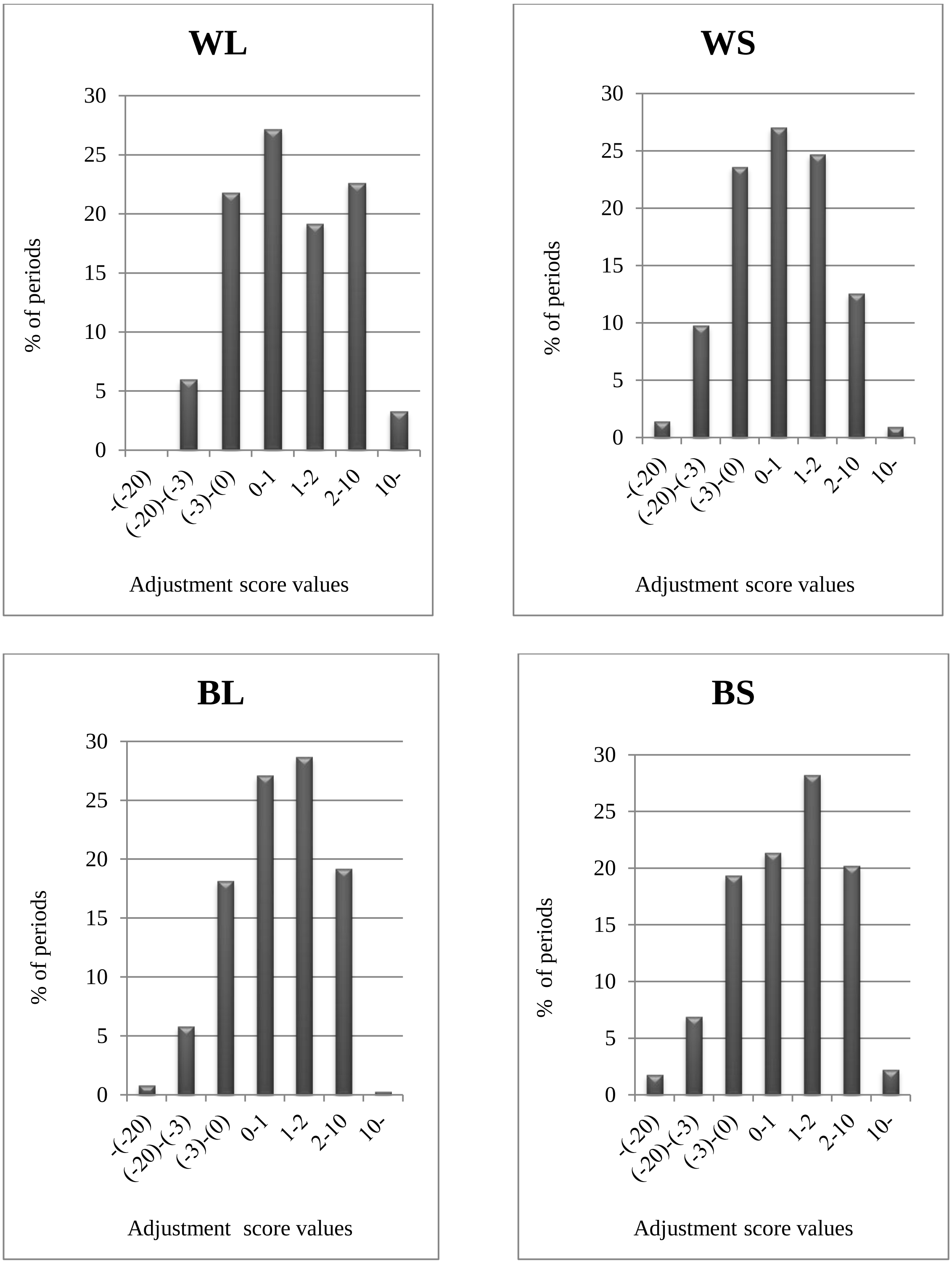


Figure 5.3.1: The Distribution of Adjustment Scores

Note that the demand chasing heuristic is not related to the optimal order quantity q^*. Also, the heuristic does not predict initial choices, but it predicts adjustment patterns across a series of choices (Schweitzer and Cachon 2000).

Researchers have measured the effect of demand chasing heuristic using the following methods:

- Counting changes towards versus away from prior demand
- Adjustment scores
- Regression
- Correlation

Lau et al. (2012) criticize the use of the first three measures. They show these methods to overestimate the extent of demand chasing (i.e., a high false positive rate), and recommend the fourth one, correlation.

Next, we present the analyses of our experimental data using each of these approaches. We also explain how we modified the approaches to address the changing q^* at each period.

5.3.1 Counting Changes Towards vs. Away From Prior Demand

This method is based on comparing the number of adjustments in the order quantity towards and away from the prior demand realization. For standard newsvendor decisions, $q_t – q_{t-1}$ and $d_{t-1} – q_{t-1}$ being of the same sign is counted as an adjustment towards the prior demand; whereas being opposite sign is an adjustment away from it. Note that this “standard metric” ignores the changes in q_t^* from period to period.

To adopt this metric to our data, we came up with a “new metric” that tracks the changes in overage or underage percentages rather than the changes in order quantity itself. The idea is that if the retailer is chasing demand, her overage or underage percentage should be changing in the direction of prior demand realization. To this end, we define as the $qp_t = (q_t - q_t^*)/q_t^*$ overage percentage in period t with respect to the optimal quantity q_t^*. Positive values of qp_t indicate overage percentage and the negatives ones indicate underage percentage.

Using this metric, we count the numbers of adjustments towards and away from prior demand value. In particular, we count ($qp_t – qp_{t-1}$) and ($d_{t-1} – q_{t-1}$) being of the same sign as an adjustment towards the prior demand; and being opposite signs as an adjustment away. A proportion of changes towards greater than 0.5 suggests demand chasing.

We ignore the data of periods in which $q_t = 0$ *or* $d_{t-1} = q_{t-1}$ or $qp_{t-1} = qp_t$. For each experiment type, we report the results both at the subject and aggregate level. Figure 5.3.2 shows the percentage of order adjustments toward prior demand at subject level.

The results say that in all experiment types most of the retailers have an order adjustment proportion greater than 0.5. However; out of these adjustments, 52% in WL, 41% in WS, 36% in BL, and 31% in BS of retailers are significant. A retailer's use of the heuristic is said to be significant if the proportion of *qp* adjustments towards prior demand is significant by the binomial test at 0.10 level.

Next, we present the aggregate results where we pool the data of all retailers together for each experiment type. From Table 5.3.1, we observe that in all four experiment types, the proportion of *qp* adjustments towards prior demand is significantly higher than the changes away. We also observe that results under our new metric and the standard metric are not different from each other. This is probably due to the pooling effects of aggregation.

Table 5.3.1: The Proportion of Adjustments Toward and Away From Prior Demand at Aggregate Level

	Under Standard Metric		**Under New Metric**		
Experiment Type	**Toward**	**Away**	**Toward**	**Away**	**p-value of the Binomial Test**
WL	65%	35%	**64%**	**36%**	***0.00***
WS	63%	37%	**65%**	**35%**	***0.00***
BL	64%	36%	**65%**	**35%**	***0.00***
BS	65%	35%	**63%**	**37%**	***0.00***

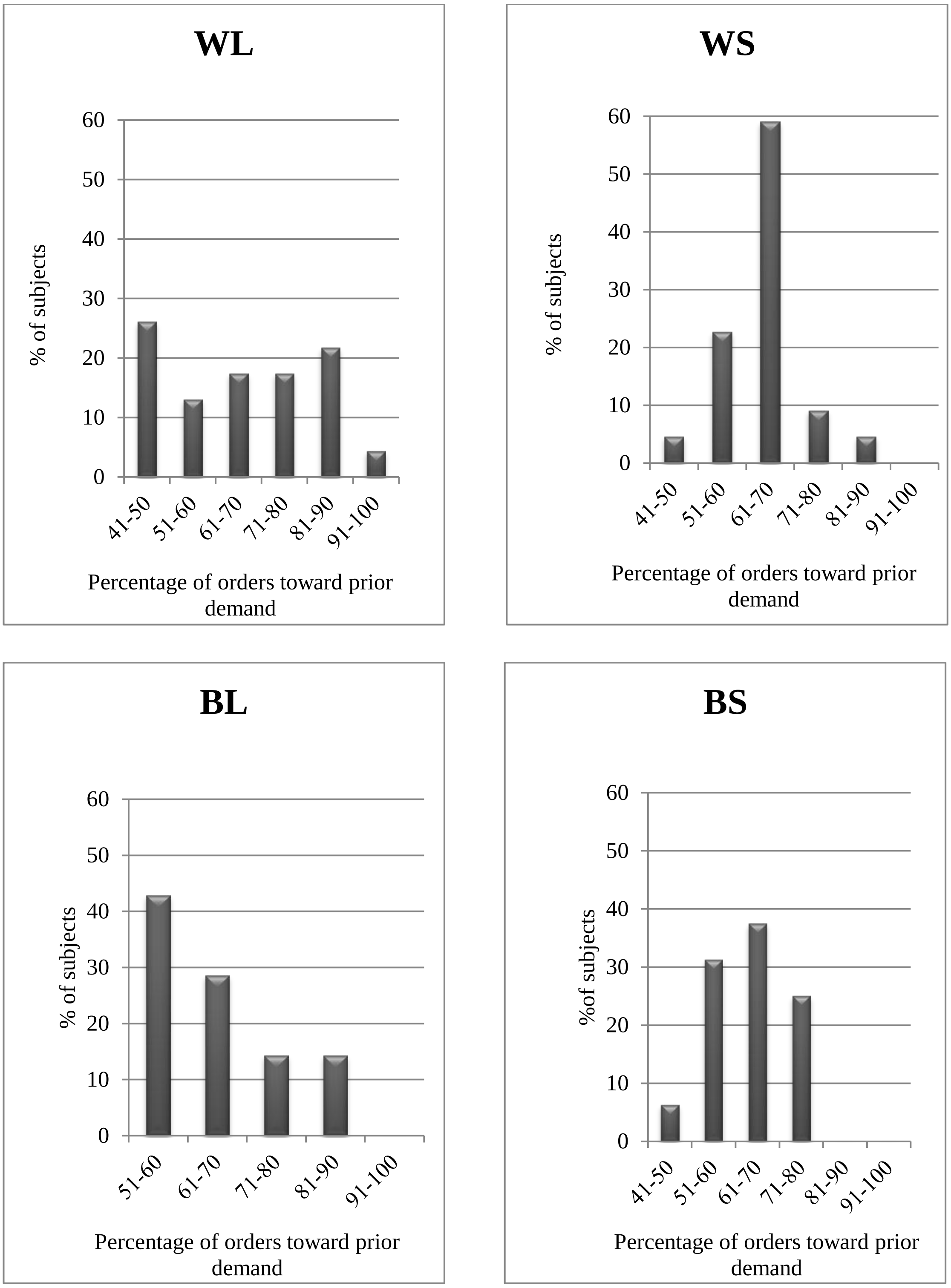


Figure 5.3.2: The Percentage of Order Adjustments toward Prior Demand

Overall, we can observe the effect of demand chasing heuristic in aggregate terms, and for most of the retailers at the subject level.

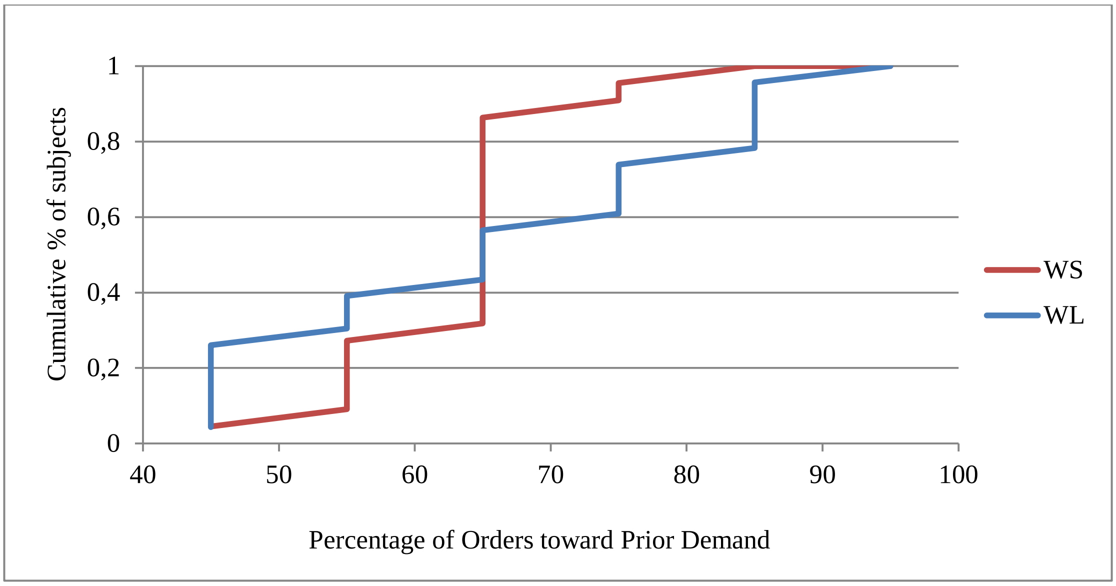


Figure 5.3.3: In Wholesale Price Experiments Cumulative Distribution of Percentage of Orders toward Prior Demand

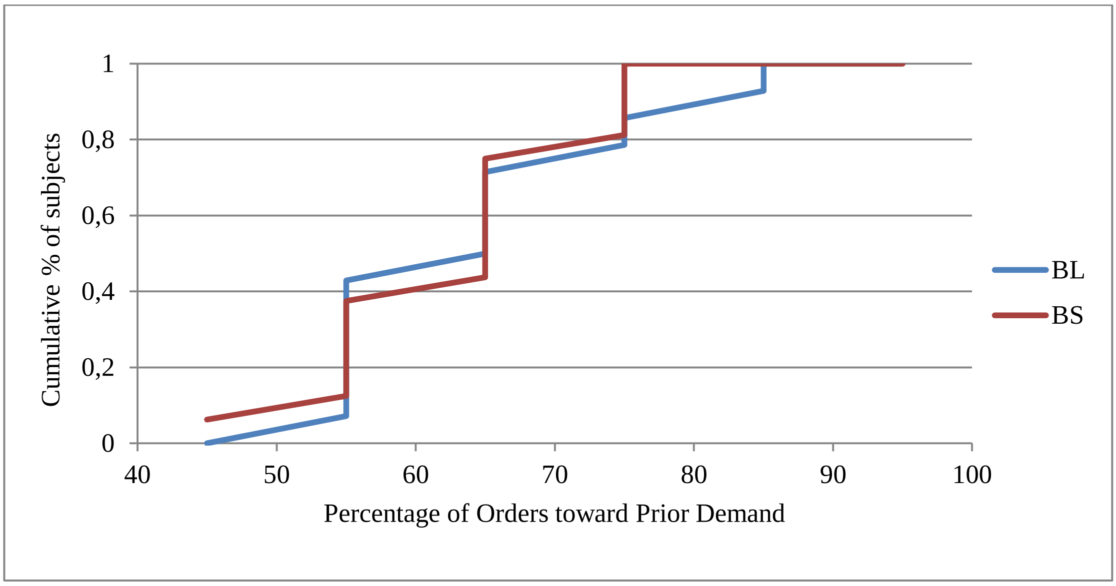


Figure 5.3.4: In Buyback Experiments Cumulative Distribution of Percentage of Orders toward Prior Demand

From Figure 5.3.3 and Figure 5.3.4 we observe that the percentage of orders toward prior demand is higher in short run experiments than in long run experiments. The comparison in the buyback experiments is, though, not that clear.

5.3.2 Adjustment Scores

The first metric that we considered in Section 5.3.1 only counted the number of changes. To capture the "strength" of adjustments, we define the adjustment score as $(qp_t - qp_{t-1}) / (d_{t-1} - q_{t-1})$. This metric is a modified version of the one used in Schweitzer and Cachon (2000). These scores are computed separately for moves toward and away from the previous period's demand. A significant difference (as measured by the Mann-Whitney test) between adjustment scores towards and away was taken as an indicator of demand chasing. We ignore the data of periods in which $q_t = 0$ *or* $d_{t-1} = q_{t-1}$ or $qp_{t-1} = qp_t$.

Table 5.3.2 presents the aggregate results for each experiment type. We compare the median values, because taking average can be misleading due to the wide interval of adjustment scores. We observe only in wholesale price long run experiments (WL) significantly higher toward scores. According to the theory especially in short run experiments, the retailers consider the side effects while they are in decision making situation, such as; previous demand, mean demand, and previous profits. Therefore, we expect significantly higher toward scores in short run experiments. We especially expect in wholesale price short run experiments, due to the more complicated structure of buyback experiments.

Table 5.3.2: Mean and Median Values of Adjustment Scores of Experiments

	The Mean Value	**The Median Value**		
Experiment Type	**Toward**	**Away**	**Toward**	**Away**
WL	0.0122	0.0128	**0.0042***[3]	**0.0035**
WS	0.0099	0.0130	**0.0035**	**0.0039**
BL	0.0103	0.0095	**0.0033**	**0.0033**
BS	0.0097	0.0159	**0.0036**	**0.0034**

[3] *is implements the significance of WL experiments in 0.05 level.

5.3.3 Regression

Bostian et.al (2008) uses the following regression equation to measure the extent of demand chasing.

$$q_t = q_{t-1} + \beta(d_{t-1} - q_{t-1}) + \varepsilon_t \tag{7}$$

where ε is an iid normal error term. Here, the parameter β reflects the "extent of demand chasing". It measures how far the subject moves toward the most recent demand observation relative to their last choice. To be consistent with demand chasing, β must lie in the interval (0, 1] with $\beta=1$ implying full demand chasing. We adopt this method by replacing the q values with the qp values we defined earlier as follows:

$$qp_t = qp_{t-1} + \beta(d_{t-1} - q_{t-1}) + \varepsilon_t \tag{8}$$

For each individual subject and for pooled data we regress (qp_t - qp_{t-1}) on (d_{t-1} - q_{t-1}) and record the β values. Figure 5.3.5 shows the cumulative distribution of β-coefficients obtained from individual regressions.

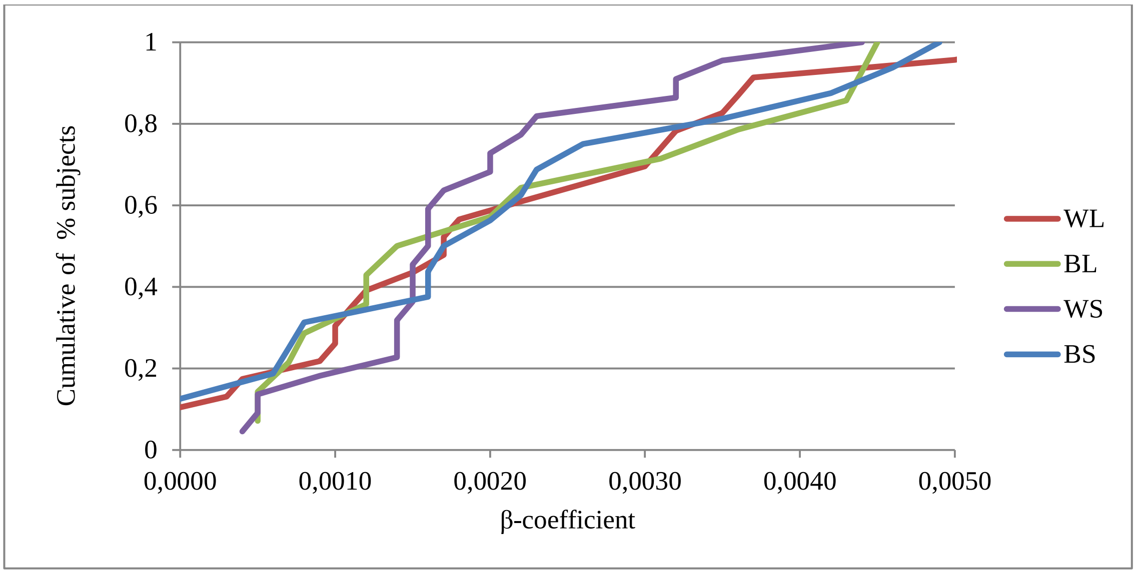


Figure 5.3.5: Cumulative Distribution of β-coefficient

In all experiment types, for every subject β-coefficients are almost positive. However the results say that 61% of subjects in WL, 59% in WS, 57% in BL, and 56% in BS have positive β-coefficients that are significant at the 0.10 level.

The aggregate data of each experiment type has positive β-coefficients and are significant at the 0.10 level. Table 5.3.3 indicates the aggregate β-coefficients. The regression equations are significant for all experiments; however adjusted R^2 values are very low. Hence, although the regression equations are significant and meaningful, the variation in the order quantities cannot be explained by the chosen independent variables. This is mostly due to the heterogeneity in subjects. As a conclusion, we cannot observe the demand chasing by doing regression analyses at aggregate level in our experiments.

Table 5.3.3: The Aggregate β-coefficients

Experiment Type	**β-coefficient**	**p-value**	**Adj. R^2**	**Number of Data Points**	**Proportion of Retailers with Significant Regression Equation**
WL	0.0022	0.00	11.59	604	0.00
WS	0.0019	0.00	9.66	607	0.00
BL	0.0022	0.00	13.59	375	0.07
BS	0.0020	0.00	8.60	439	0.00

5.3.4 Correlation

An alternative way to measure demand chasing in standard newsvendor experiments is to calculate the correlation between q_t and d_{t-1} series for every subject (Bolton and Katok 2008, Lau et al. 2012). A positive correlation between these two variables would support demand chasing. We adopt this approach to our setting by replacing q_t with qp_t.

We find that 57% of subjects in WL, 68% in WS, 50% in BL, and 56% in BS have positive correlation between qp_t and d_{t-1}. Figure 5.3.6 shows the cumulative distribution of correlation values obtained from individual correlations. The results say that the most of the correlation values are around 0.30 in each experiment type.

Hence, according to the first method (counting changes towards vs. away from prior demand), demand chasing can be observed both at subject and aggregate level. According to the third (regression) and fourth method (correlation), demand chasing can be observed only at subject level but not at aggregate level. However, according to the second method (adjustment score), demand chasing cannot be observed both at subject and aggregate levels except at wholesale price long run experiments.

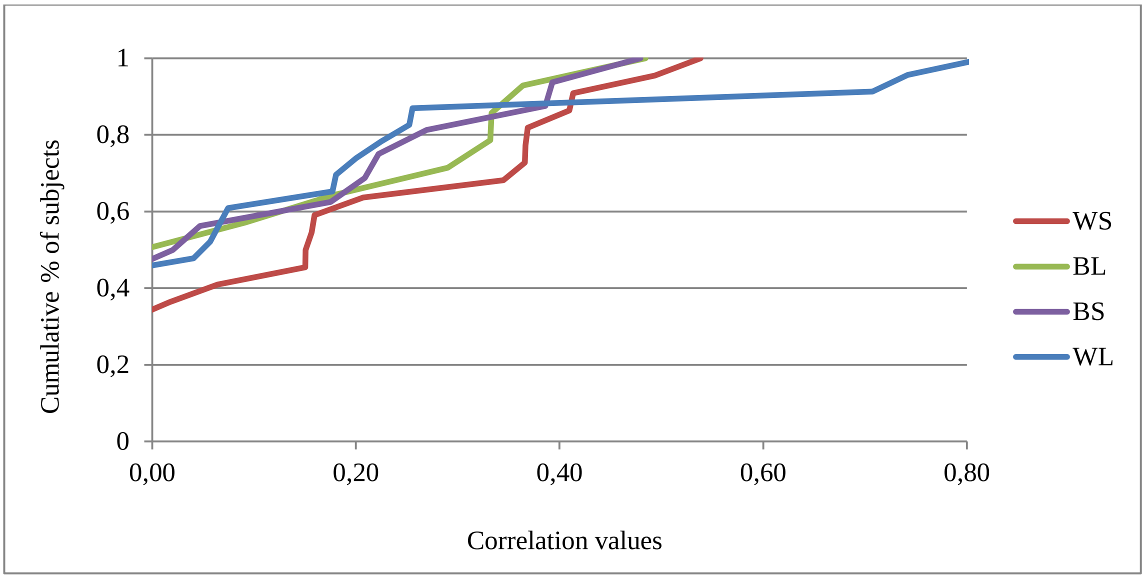


Figure 5.3.6: Cumulative Distribution of Correlation Values

CHAPTER 6

6 DETERMINING THE FACTORS THAT AFFECT ORDER DECISIONS

Until now, we have observed that the retailers deviate from the optimal newsvendor order quantity decision. In this chapter, we try to figure out what factors affect these decisions of the retailers. We use the data of buyback contract short-run experiments to avoid fairness effect that might be present in the long-run experiments. This data consists of 16 retailer and 16 manufacturer's 30 period decisions where different manufacturer-retailer pairs are matched in each period. We analyze each retailer's decision individually, since using the average does not seem to be appropriate.

First, to identify the most important factors, we apply "feature selection" methodology to data. Then, we build regression models to capture the relationship between the order quantity decisions and the selected attributes.

6.1 Selection of the Factors

Feature selection is the process of selecting a subset of relevant features to use in model construction. Feature selection has been an active research area in statistics and data mining. The central assumption when using a feature selection technique is that the data contains many redundant or irrelevant features. The objective of feature selection is three-fold: Improving the prediction performance of the predictors, providing faster and more cost-effective predictors, and providing a better understanding of the underlying process that generated the data (Guyon and Elisseeff 2003).

We apply Feature Selection to each individual retailer decision maker's order quantity decisions, and try to figure out which attributes are effective in each individual's decision making process. To this end, we use the machine learning software Weka, which contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality.

We chose the order quantity as the “output” because of investigating the effects of other attributes on the order quantity decision. Additionally we chose nine “attributes” that can potentially affect the order quantity decision as as shown in Table 6.1.1. These attributes were chosen based on literature including: Schweitzer and Cachon (2000), Bolton et al. (2008), Bolton and Katok (2008), Bostian et al. (2008), and Becker-Peth et al. (2013).

Table 6.1.1: Output Variable and Attributes

Output	**Current Period’s Attributes**	**Abbreviation**	**Previous Period’s Attributes**	**Abbreviation**
Order quantity	Period	period	Past Demand Realization	pdr
	Cost of Underage	cu	Retailer Realized Profit	rr
	Cost Of Overage	co	Manufacturer Realized Profit	mr
	Manufacturer Predicted Profit	mp	Retailer’s Profit Share	profitshare
	Retailer Predicted Profit	rp		

As seen in the table, we used predictor variables that are relevant to both the previous period, and also to the current period.

Current period variables: Period refers to the phases of 30 period decisions. We assign number “1” for the first ten periods’ decision, “2” for the next ten periods’ decision, and “3” for the last ten periods. Cost of underage is the cost that retailer loses if he orders less than demand. Cost of overage is the cost that retailer has to pay if he orders more than demand. These variables are used instead of the contract parameters wholesale price and buyback price, such that the data from different contract types can be compared in future studies. Manufacturer predicted profit and retailer predicted profit are the expected profits of the players in the current period when the retailer sets the newsvendor optimal order quantity.

Previous period variables: Past demand realization, manufacturer realized profit, retailer realized profit, and retailer's profit share refers to the relevant values in the previous period.

For feature selection, the RelieffAttributeEval function of WEKA software was used with the ranker search method. RelieffAttributeEval method evaluates the value of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class. This method operates on both discrete and continuous class data. In ranker method, all attributes are ranked starting from the most important one to the least important one. Cross validation is selected as the Attribute Selection Method.

We applied the same method to all 16 retailer's decisions and we recorded each retailer's five most important features. We assign weights to the attributes such as "5" if the attribute is the most important one and "1" if it is the last one important. Then, we calculated each attribute's weighted sum, and we rank the first five attributes that are most effective in making decisions. The results are shown in Table 6.1.2 and Table 6.1.3.

Table 6.1.2: The Most Important 5 Attributes Selected by Retailers

Retailer	**Selected Att.1**	**Selected Att.2**	**Selected Att.3**	**Selected Att.4**	**Selected Att.5**
1	pdr	co	mr	p	rr
2	co	rp	mp	cu	pdr
3	co	rp	cu	pdr	mp
4	mr	pdr	rr	cu	profitshare
5	co	profitshare	mp	rp	cu
6	mr	co	pdr	profitshare	p
7	co	rp	pdr	mp	mr
8	co	rp	mp	cu	rr
9	co	mp	rr	cu	p
10	cu	co	rp	mp	mr
11	co	mp	pdr	rp	p
12	co	rp	cu	mp	mr
13	pdr	co	mp	profitshare	mr
14	rp	mr	cu	co	p
15	co	mp	rp	cu	profitshare
16	co	mp	rp	cu	pdr

Weighted sum of each selected attribute are shown in Table 6.1.3.We observe that cost of overage is the most important attribute that affects the retailer decisions. The other important attributes include retailer's predicted profit, manufacturer's predicted profit, past demand realization and cost of underage.

Table 6.1.3: Weighted Sum of Each Attribute in Experiments

Attribute	**Weighted Sum**
co	**68**
rp	**38**
mp	**35**
pdr	**27**
cu	**27**
mr	21
profitshare	10
rr	8
p	6

The reason why cost of overage is the most important attribute might be related to risk aversion of the retailer. The buyback price is generally far less than the wholesale price in proposed contracts. Because demand is probabilistic, retailers avoid taking high risk and hence, cost of overage becomes the most important factor affecting their order quantity decision. Predicted profits of both sides might be important due to the fairness concerns. If manufacturer's predicted profit is much higher in a given contract, retailer will not be willing to order high quantities.

6.2 Regression Analysis

The next step after feature selection is classification. Classification takes a data set with known output values and uses this data set to build a model. We apply linear regression method to classify the data in Minitab software. The output is retailer's order quantity decision, and the selected five attributes are the independent variables of the regression model. We hope to build more accurate regression models as we exclude the redundant attributes identified in the feature selection phase.

Recall that the selected five attributes are cost of overage, retailer's predicted profit, manufacturer's predicted profit, past demand realization and cost of underage. We expect order quantity to be increasing in the retailer's predicted profit, and past demand realization and cost of underage; and decreasing in the cost of overage, and manufacturer's predicted profit attributes. Minitab results for each individual retailer are shown below in Table 6.2.1.

Table 6.2.1: Individual Regression Equations

Retailer	**R^2**	**Adjusted R^2**	**p value**	**Regression Equation**
1	37.7%	24.1%	0.042	Q = 569 **+ 0,317 pdr** - 0,017 mp - 0,085 rp + 7,93 cu-3,01 co
2	85.2%	81.9%	0.000	Q = -208 - 0,036 pdr + 0,018 mp + 0,044 rp - 3,10 cu + 0,55 co
3	83.8%	80.3%	0.000	Q = -943 + 0,068 pdr **+ 0,050 mp + 0,095 rp - 6,39 cu + 2,54 co**
4	41.4%	27.4%	0.035	Q= -526 **+ 0,248 pdr** + 0,030 mp - 0,002 rp + 2,68 cu + 0,23 co
5	95.0%	93.9%	0.008	Q= -362 - 0,010 pdr **+ 0,0251 mp + 0,057 rp - 4,03 cu + 1,00 co**
6	18.4%	0.64%	0.420	Q= 1073 + 0,213 pdr - 0,044 mp - 0,095 rp + 6,86 cu - 3,04 co
7	57.0%	46.7%	0.002	Q = -1253 + 0,193 pdr + 0,066 mp + 0,122 rp - 8,03 cu + 2,84 co
8	70.0%	63.5%	0.000	Q = 381 + 0,041 pdr - 0,008 mp + 0,014 rp - 2,70 cu - 0,11 co
9	48.4%	36.1%	0.011	Q = 225 - 0,053 pdr - 0,0014 mp + 0,0001 rp + 0,29 cu - 0,829 co
10	69.4%	62.7%	0.000	Q = -347 + 0,027 pdr **+ 0,025 mp** + 0,036 rp - 1,61 cu + 0,535 co
11	49.2%	35.1%	0.022	Q = 158 - 0,019 pdr - 0,0024 mp + 0,035 rp - 2,51 cu - 0,35 co
12	61.6%	52.5%	0.001	Q = -895 - 0,188 pdr **+ 0,050 mp + 0,118 rp - 8,73 cu + 2,98 co**
13	25.2%	8.9%	0.214	Q = -648 - 0,060 pdr + 0,0383 mp + 0,059 rp - 3,84 cu + 1,77 co
14	58.0%	31.7%	0.153	Q = -1300 + 0,259 pdr + 0,064 mp + 0,087 rp -3,94 cu + 2,54 co
15	81.7%	77.7%	0.000	Q = 264 - 0,064 pdr - 0,0045 mp+ 0,0216 rp - 2,25 cu - 0,418 co
16	57.0%	47.7%	0.001	Q = 136 + 0,185 pdr - 0,0036 mp + 0,0511 rp - 5,04 cu + 0,50 co

The regression equations are almost significant and R-squared values seem high enough. However, most of the factors in regression equations were not found to be significant even at 10% level. Significant factors in equations are shown with bold fonts.

The signs of the beta coefficients in regression equations usually do not follow our predictions. For example, cost of overage usually has a positive sign, whereas cost of underage has negative. We expect the order quantity to be increasing in the retailer's predicted profit, and decreasing in the manufacturer's predicted profit. Interestingly, the manufacturer's predicted profit sign is negative for most retailers but the retailer's predicted profit sign is positive for all retailers as expected. The retailers seem to care positively about the manufacturer's profit as well.

Next, we develop regression models on "pooled" data for each experiment. We consider each retailer as an independent variable not to lose the individual effect of each subject. Results are shown in Table 6.2.2.

Table 6.2.2: Regression Equation of Pooled Data

R^2	**Adjusted R^2**	**p value**	**Regression Equation**
40.1%	37.2%	0.000	Q_16 = 170,1**+0,0532 pdr** - 0,00083 mp - 0,00135 rp+0,459 cu **- 0,609 co** – 0,11 r2 + - 10,63 r3 + 10,71 r4 + 1,03 r5 **- 19,73 r6** - 14,34 r7 + 5,73 r8+9,97 r9**+17,74 r10** - 9,64 r11+12,57 r12 **- 15,49 r13** - 0,3 r14 -13,05 r15 **- 22,71 r16**

The regressions itself, as well as the attributes past demand realization and cost of overage are found to be significant. Their signs are also consistent with our predictions. However, while the manufacturer's predicted profit has negative sign, the retailer's predicted profit has also negative sign which is contrary to our predictions. Among the 16 retailers, only four of them are found to be significant (r_6, r_{10}, r_{13}, and r_{16}).

Based on these observations, we conclude that although it is reasonable to apply regression, we could not find strong evidence between the attributes and the order quantity decision.

CHAPTER 7

7 LEARNING BY DOING

Whether and how learning occurs in the newsvendor problem has been a popular question among experimental OM researchers. We would like to answer the question: "*Do retailer subjects in our experiments learn to make better decisions over time?*"

We will first consider the data from our main experiments. We then present another study that considers the effect of gender difference in learning. For that study, we use the data from another experiment, which involves only the standard newsvendor decision (i.e., no manufacturers).

7.1 Is Learning Effect Observed in the Main Experiments?

Recall that in our experiments every period the optimal order quantity changes. In order to measure whether the subjects learn to make better decisions over time, we consider the absolute difference between the order quantity decisions and the optimal order quantity. If learning exists, this difference should decrease over time. In particular, we hypothesize that for a given retailer subject, the average difference in the last 10 periods should be smaller than the average difference in the first 10 periods.

We ignore the data where $q=0$. Table 7.1.1 shows the individual results for all experiment types. The colored cells indicate the retailers whose average absolute difference between the order quantity decisions and the optimal order quantity decreased from first 10 periods to last 10 periods. Overall, we observe 45% of the retailers to exhibit decreasing difference behavior, hence an indication of learning. However, not all decreases are statistically significant. As Table 7.1.2 indicates, less than 50% of the retailers have significant decrease results in each experiment type (the statistical significance is measured by the Mann-Whitney test at 0.10 level).

Table 7.1.1: The Individual Results for All Experiment Types

Individual Subjects (Retailers)	WL Average |q-q*|		WS Average |q-q*|		BL Average |q-q*|		BS Average |q-q*|	
	First 10 period	Last 10 period	First 10 period	Last 10 period	First 10 period	Last 10 period	First 10 period	Last 10 period
r_1	18	21	16	21	20	34	45	22
r_2	21	19	13	9	23	22	12	7
r_3	45	26	18	6	7	10	11	12
r_4	36	25	40	14	44	10	20	45
r_5	29	21	28	10	26	27	0	0
r_6	18	13	35	47	28	37	51	57
r_7	17	35	39	26	40	25	33	38
r_8	42	43	43	20	23	48	14	18
r_9	14	11	35	24	39	80	47	18
r_{10}	12	22	30	69	35	67	23	20
r_{11}	76	47	22	5	24	17	9	48
r_{12}	10	13	26	27	20	10	24	45
r_{13}	25	67	13	34	7	60	32	23
r_{14}	42	64	54	32	33	24	37	26
r_{15}	17	18	31	20			18	4
r_{16}	17	25	22	36			27	26
r_{17}	8	15	14	38				
r_{18}	64	55	5	12				
r_{19}	18	36	13	36				
r_{20}	11	30	18	48				
r_{21}	16	69	29	30				
r_{22}	7	21	18	9				
r_{23}	67	52						

Table 7.1.2: The Proportion of Retailers with Significant Result

Experiment Type	Proportion of Significant Retailers
WL	35%
WS	41%
BL	29%
BS	38%

7.2 The Effect of Learning in terms of Gender

In this study, we aim to understand whether a learning effect exists and differs from gender to gender in the standard newsvendor experiments. In order to evaluate this topic we used the data which is gathered in a former study by Nukte Sahin, a former graduate student of Dr. Kaya.

The experiment was conducted with 156 students (82 male and 74 female) of the course MS 401 in the spring semester 2010/2011. These students had already studied the basic newsvendor problem. To provide incentive, the subjects' total profit at the end of the experimental session was converted into a bonus applied to the course final grade. The bonus ranged between 1% and 2.5%. Experiments were conducted in the CAFE (Center for Applied Finance Education) computer laboratory of Sabanci University.

The subjects faced the standard newsvendor problem. They need to determine how much to order from the manufacturer before the sales season. They know that the demand (*D*) will be uniformly distributed between *50* and *150*. The purchase price of the product is *w=$35*, and the sales price in the market is *p=$90*. The objective of the subjects is to maximize their total profit over the *40* periods.

To investigate the impact of learning in the newsvendor setting, we regressed the data using the following equation.

$$|q - q^*| = \beta_0 + \beta_1 gender + \beta_2 period + \varepsilon \tag{9}$$

The dependent variable is the absolute difference between the optimal order quantity and the order quantity decision. The independent variables are gender (a binary variable), and period. The analysis is carried out in the Minitab software. Comparisons are assessed by the Mann-Whitney test at 0.05 significance level.

We use two different approaches to test for learning. The first approach is based on measuring experience with respect to period, similar to Bolton and Katok (2008). We run the regression for each subject separately, and record the regression coefficient on period (β_2).

Hypothesis: *The median period coefficient of males (μ_m) is not different from the coefficients of females (μ_f). That is, there is no difference between males and females in learning.*

We find the male median coefficient to be lower than the female coefficient (μ_m*=0.05755,* μ_f*=0.09010).* However, the difference is not significant. Hence, we cannot reject the hypotheses: there is no significant difference between males and females in learning ($p=0.5486$).

In the second approach, we compared the order quantities in the first half of the periods with the second half. This is similar to Benzion et al. (2008).

Hypothesis: *The median period coefficient (μ_l) in the last half of the periods is larger than the median periods' coefficient (μ_b) in the first half of the periods. This holds true for both male and female subjects separately.*

According to the test results, the null hypothesis cannot be rejected both for males ($\mu_{fb}=-0.0026$, $\mu_{fl}=-0.0147$, $p=0.4755$) and females ($\mu_{mb}=0.0000$ $\mu_{ml}= 0.0212$, $p=0.2470$). That is, the order quantities of both male and female subjects are not closer to the optimal level in the last 20 periods than in the first 20 periods. We cannot find evidence for learning both for male and female subjects.

CHAPTER 8

8 CONCLUSIONS AND FUTURE RESEARCH

In this research, we consider decision-making experiments that are conducted with human decision makers on a manufacturer-retailer supply chain where the retailer faces a newsvendor problem. In standard newsvendor experiments, in all periods the retailer faces the same contract offered by the computer. However, in our experimental setting the optimal order quantity for the retailer's problem changes every period based on the offered contract. This strategic interaction, as well as the other documented biases in newsvendor decisions affects the retailer subjects' order quantity choice. Our goal is to explain the decision making mechanism of the retailer subjects in this setting. To this end, we conducted three studies.

In our first study, we aim to answer whether the subjects exhibit the pull to center effect, and whether they follow certain decision heuristics while making their order decisions. We observe that most subjects do not exhibit a strong pull to center behavior. One reason is that the optimal order quantities are rather close to the mean demand value, which causes the pull to center region to shrink. We could find weak support for the mean anchoring heuristic at aggregate level, but not at individual level. For the demand chasing heuristic, aggregate level support is high; however, we could find strong support at individual level. When studying the demand chasing heuristic, in addition to the standard metrics used in literature, we defined our own metric. This metric considers the fact that the optimal order quantity is changing from one period to the other, based on the offered contract.

An important observation we have is that subject-level data is highly heterogeneous. Hence, it would be misleading to accept the existence of these heuristics for a given individual just because they are known to exist at aggregate level. Thus, one needs to be careful in using the aggregate (or, average) results in literature because they do not necessarily apply to a given individual.

In our second study, we apply “feature selection” and “classification” techniques to understand the factors that affect the order quantity decisions of retailer subjects. We came up with nine candidate attributes that are related to either the current or the previous period data. Among these, we determine the most important five factors as cost of overage, manufacturer’s predicted profit, retailer’s predicted profit, cost of underage and past demand realization. We then build regression models separately for each individual. The regression equations are significant, R-squared values are high, and the beta coefficients have almost predicted signs. However, most of the attributes are not found to be significant for most individuals. Here again, heterogeneity plays an important role. One cannot come up with a set of, say, five attributes that turn out to be significant for all individuals.

In our third study, we aim to understand if and how the subjects’ decisions change over time due to learning-by-doing. To test whether they learn to make better decisions over time, we consider the average absolute difference between the stock quantity decisions and the optimal order quantity. If learning exists, this difference should decrease over time. We observe less than 50% of the retailers to have a significant decrease in each experiment type.

We also conducted a side study to test whether subject gender affects learning. For this study, we used data from a simple newsvendor experiment (i.e., no manufacturer subjects). We built a regression model where the absolute difference between the stock quantity decisions and the optimal order quantity is the dependent variable, and the gender (a binary variable) and period are the independent variables. We could not observe a significant learning effect, and also any significant difference in learning between genders. The absolute difference between the subjects’ stock quantity decisions and the optimal order quantity is not decreasing over time.

This work can be extended in numerous ways. One extension is to study the manufacturer’s contract choice behavior. Another is to conduct experiments on other supply chain contract types, such as revenue sharing, quantity discount contract and rebate contract, and present a more complete comparison in terms of the factors that affect the retailer’s stock quantity decisions. Yet another possibility is applying feature selection techniques with a larger set of potential factors. We have only tested whether learning exists. One can also develop a model of learning, such as the Experience Weighted Attraction (EWA) learning model.

BIBLIOGRAPHY

Arrow, K. J., T. Harris, J. Marschak. 1951. Optimal inventory policy. *Econometrica* **19**(3) 250-272.

Bazerman, M. H., (2008). *Judgment in Managerial Decision Making*. Wiley Seventh Edition.

Becker-Peth, M., E. Katok, U. W. Thonemann. 2013. Designing contracts for irrational but predictable newsvendors. *Management Science* **59**(8) 1800-1816.

Benzion, U., Y. Cohen, R. Peled, T. Shavit. 2008. Decision-making and the newsvendor problem: an experimental study. *Journal of the Operational Research Society* **59** 1281-1287.

Benzion, U., Cohen, Y., Shavit, T. 2010. The newsvendor problem with unknown distribution. *Journal of the Operational Research Society* **61** 1022-1031.

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

Bolton, G. E., A. Ockenfels, U. Thonemann. 2012. Managers and students as newsvendors: How out-of-task experience matters. *Management Science* **58**(12) 2225–2233.

Bostian, AJ A., C. A. Holt, A. M. Smith. 2008. Newsvendor “Pull-to-center” effect: Adaptive learning in a laboratory experiment. *Manufacturing & Service Operations Management* **10**(4) 590-608.

Cachon, G. 2003. Supply chain coordination with contracts. A.G. de Kok, S.C. Graves, eds. Chapter 6 in *Handbooks in Operations Research and Management Science, Vol. 11* Elsevier, Amsterdam.

Cachon, G., M. A. Lariviere. 2005. Supply chain coordination with revenue-sharing contracts: Strengths and limitations. *Management Science* **51**(1) 31-44.

Camerer, C., M. Weber. 1992. Recent developments in modeling preferences:uncertainty and ambiguity. *Journal of Risk & Uncertainty* **5** 325–370.

Croson, R., Y. Ren. 2013. Overconfidence in newsvendor orders: An experimental study. *Management Science* **59**(11) 2502-2517.

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

Eeckhoudt, L., C. Gollier, H. Schlesinger. 1995. The risk-averse (and prudent) newsboy. *Management Science* **41**(5) 786-794.

Emmons, H., S. M. Gilbert. 1998. Returns policies in pricing and inventory decisions for catalogue goods. *Management Science* **44**(2) 276-283.

Feng, T., L. R. Keller, X. Zheng. 2010. Decision making in the newsvendor problem: A cross-national laboratory study. *Omega, International Journal of Management Science* **39**(1) 41-50.

Gavirneni*, S., A. M. Isen*. 2010. Anatomy of a newsvendor decision: Observations from a verbal protocol analysis. *Production and Operations Management* **19**(4) 453–462.

Guyon, I., A. Elisseeff. 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* **3** 1157–1182.

Haruvy, E., E. Katok, V. Pavlov. 2012. Can coordinating contracts improve channel efficiency? Working paper.

Hyndman, K., S. Kraiselburd, N. Watson. 2012. Coordination in games with strategic complementories : An experimental on fixed vs. random matching. *Production and Operations Management* **0**(0) 1-18.

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

Kahneman, D., P. Slovic, A. Tversky. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

Kandel, E., 1996. The right to return. *Journal of Law and Economics* **39** 329-356.

Katok, E., D. Y. Wu. 2009. Contracting in supply chains: A laboratory investigation. *Management Science* **55**(12) 1953-1968.

Kaya, M., Ö. Özer. 2010. Risk and information sharing in supply chains through pricing contracts. Ö. Özer and R. Phillips eds. to appear in *Handbook of Pricing Management*. Oxford University Press.

Keser, C., G. A. Paleologo. 2004. Experimental investigation of supplier-retailer contracts: the wholesale price contract. *Scientific Series* C*irano* **57**.

Kremer M., S. Minner, L. N. V. Wassenhove. 2010. Do random errors explain newsvendor behavior? *Manufacturing & Service Operations Management* **12**(4) 673-681.

Lau, N., J. N. Bearden. 2012. Newsvendor demand chasing revisited. *Management Science* **59**(5) 1245-1249.

Lau, N., J. N. Bearden, S. Hasija 2014. Newsvendor pull-to-center reconsidered. *Decision Support System* **58** 68-73.

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

Loch, C. H., Y. Wu. 2008. Social preferences and supply chain performance: An experimental study. *Management Science* **54**(11) 1835-1849.

Lurie, N. H., J. M. Swaminathan. 2009. Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes* **108**(2) 315-329.

Pasternack, B. 1985. Optimal pricing and returns policies for perishable commodities. *Marketing Science* **4** 166-176.

Sahin, N., 2011. Experiments on supply chain contracting: Effects of contract type and relationship lenght. Master thesis, Sabancı University.

Schultz, K. L., L. J. Thomas, J. O. McClain, L. W. Robinson. 2007. The use of framing in inventory decisions. *Johnson School Research Paper Series* **02-07**.

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

Spengler, J.,J. 1950. Vertical integration and antitrust policy. *Journal of Political Economy* **58**(4) 347 – 352.

Su, X. 2008. Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management* **10**(4) 566-589.

Tsay, A. A. 1999. The quantity flexibility contract and supplier-customer incentives. *Marketing Science* **45**(10) 1339 –1358.

Vericourt, F., K. Jain, J. N. Bearden, A. Fillipowicz. 2013. Sex, risk, and newsvendor. *Journal of Operations Management* **31(**1) 86 – 92.

Wang, C. X., S. Webster. 2009. The loss-averse newsvendor problem. *Omega* **37**(1) 93-105.

Wu, D. Y. 2013. The impact of repeated interactions on supply chain contracts: A laboratory study. *Int. J. Production Economics* **142** 3-15.

APPENDICES

Appendix A Main Script Code in Buyback Experiments

```
// Define Player List
        Players p1,p2,p3,p4,p5,p6,p7,p8,p9,p10,p11,p12,p13,p14,p15,p16;
        Integer nplayer = 16;   //number of players
// Declare variables
        Script("c:\program files\hp mums\Scripts\buyback\var-model.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\var-dummy.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\var-state.cfg");
// Set parameter value
        Script("c:\program files\hp mums\Scripts\buyback\dat-parameter.dat");
// Define inputs
        Script("c:\program files\hp mums\Scripts\buyback\def-input.cfg");
//Stage logon
        Script("c:\program files\hp mums\Scripts\buyback\stage-logon.cfg");
// Game stages
        Script("c:\program files\hp mums\Scripts\buyback\stage-start.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-setgrid.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-predisplay.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-fetchdata.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-exchange.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-results.cfg");
        Script("c:\program files\hp mums\Scripts\buyback\stage-periodend.cfg");
Stage writedb {                         // no db write statements in debug
                Script("c:\program files\hp mums\Scripts\buyback\stage-dblog-
period.cfg");
                if (stage=1)
                { End;}
                else
                {Goto start;}
}
```

Appendix B The Script of dat-parameter.dat in Buyback Contract Experiments

```
stage setparameter
{
	if (period=1 & stage=1)
	{	 // parameters start here
		wholesalegiven = 0;
		buybackgiven = 0;

		price = 250;
		unitcost = 50;
		wholesale = 0;
		buyback = 0;
		mindemand = 40;
		maxdemand = 230;		//parameters end here

		// manufacturer's stage description
		stagedesc[0,1] = "Wholesale and buyback price selection";
		stagedesc[0,2] = "Waiting for the retailer";
		stagedesc[0,3] = "Period results";

		// retailer's stage description
		stagedesc[1,1] = "Waiting for manufacturer";
		stagedesc[1,2] = "Stock quantity decision";
		stagedesc[1,3] = "Period results";

		numman = int(nplayer/2);
		numret = nplayer - numman;
	}

	//allocation of fixed roles and variable partners
	if (stage=1)
	{
	matched=0;
	pos1=0;
	pos2=0;
		 for (i=0; i<nplayer/2; i=i+1)
		{
			allocation1[i] = -1;
		}
		for (i=0; i<nplayer/2; i=i+1)
		{
			allocation2[i] = -1;
		}
		for (i=0; i<nplayer/2; i=i+1)
		{
			pos1 = int(nplayer/2*random);
			if (pos1 = nplayer/2)
			{
				pos1 = nplayer/2-1;
```

```
		}


		if (allocation1[pos1] = -1)
		{
			allocation1[pos1] = i;
		}
		else
		{
			while (allocation1[pos1] <> -1)
			{
				pos1 = (pos1 + 1) % (nplayer/2);
			}
			allocation1[pos1] = i;
		}
	}

	for (i=0; i<nplayer/2; i=i+1)
	{
		pos2 = int(nplayer/2*random);
		if (pos2 = nplayer/2)
		{
			pos2 = nplayer/2-1;
		}

		if (allocation2[pos2] = -1)
		{
			allocation2[pos2] = i+nplayer/2;
		}
		else
		{
			while (allocation2[pos2] <> -1)
			{
				pos2 = (pos2 + 1) % (nplayer/2);
			}
			allocation2[pos2] = i+nplayer/2;
		}
	}


	for (i=0; i<nplayer/2; i=i+1)
	{

match[allocation1[i]]=allocation2[i];
match[allocation2[i]]=allocation1[i];

role[allocation1[i]] = 0;  //manufacturer
role[allocation2[i]] = 1;  //retailer
```

```
        demand[allocation1[i]] = mindemand + int((maxdemand -
mindemand)*random);
        demand[allocation2[i]] = 0;
                }


                if (wholesalegiven = 1  &  buybackgiven = 1)
                {
                        for (i=0; i<nplayer; i=i+1)
                        {
                                if (role[i] = 0)
                                {
                                        wholesaleset[i] = wholesale;
                                        buybackset[i] = buyback;
                                }
                                else
                                {
                                        wholesaleset[i] = -1;
                                        buybackset[i] = -1;
                                }
                        }
                }

                if (wholesalegiven = 1  &  buybackgiven = 1)
                {
                        stage = 2;     // advance to stage 2 right away
                }
        }
}
```

Appendix C Instructions for Buyback Contract Experiments with Variable Partners

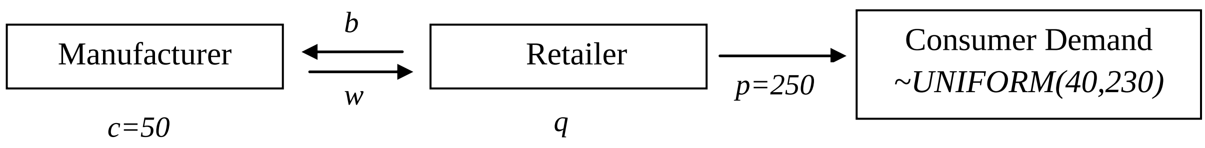
Instructions for Buyback Contract Experiments

March 23th, 2011

Random Match

Scenario

We consider two independent firms: a manufacturer and a retailer. The manufacturer produces a certain product. The retailer buys the product from the manufacturer by paying a ***wholesale price w*** per unit, and sells it to consumers at a ***retail price p=250***. Consumer demand is distributed uniformly *between 40 and 230*. That is, demand is equally likely to be an integer value *between 40 and 230*. After the demand is realized, the manufacturer buys back the products that the retailer cannot sell by paying the retailer ***buyback price b*** per unit.



The game has three stages:

Stage-1: The manufacturer determines the wholesale price, *w* and the buyback price, *b*. The wholesale price cannot be larger than the retail price *p=250*. The buyback price cannot be larger than the wholesale price.

Stage-2: Given the wholesale price and buyback price decisions of the manufacturer, the retailer determines his ***order quantity, q.*** The retailer orders this quantity of products from the manufacturer. The manufacturer produces the products by incurring a ***unit production cost c=50***, and sends them to the retailer. The retailer stocks these products prior to the selling season. The retailer's stock quantity can be either *zero* or lie *between 40 and 230*, the maximum consumer demand value.

Stage-3: Random consumer demand is realized as "*d*". The retailer's ***sales quantity*** is the minimum of his stock quantity and the realized demand: $\min\{q, d\}$. Depending on whether the demand is greater or less than retailer's stock quantity, two cases are possible:

- If $d>q$, then *(d-q)* units of demand will be unsatisfied (***Unsatisfied demand***)
- If $d<q$, then *(q-d)* products will be unsold at the retailer (***leftover products***). The manufacturer will buy back these units from the retailer.

The retailer's payoff is calculated as $p * min\{q, d\} - w * q + b * [q - min\{q, d\}]$.

The manufacturer's payoff is calculated as $(w - c) * q - b * [q - min\{q, d\}]$.

Note that there are three decisions in the game: The manufacturer determines w and b, and then the retailer determines q.

<u>Experiment Preparation</u>

- The experiments will take place at the CAFÉ (Center for Applied Finance Education) computer lab at the G-floor of the FMAN building.
- Please come to the experiments on-time so that we can start and finish on time.
- You will play a pilot experiment to solidify your understanding of the software.
- Please do not open any other program, including other browser windows, during the experiments.
- Please enter "integer values" for all decisions, and pay attention to the data entry rules.

<u>The Experiment</u>

- In the experiments, you will play the role of either a manufacturer or a retailer for a number of "periods". Your role will be fixed in all periods of an experiment. In each period, the server will randomly match each manufacturer with a retailer. That is, you will be (most likely) playing with different opponents at each period. You will not know with whom you are matched.
- The periods are independent of each other. For example, inventory is not carried from one period to the next. Only your payoff will accumulate over periods.

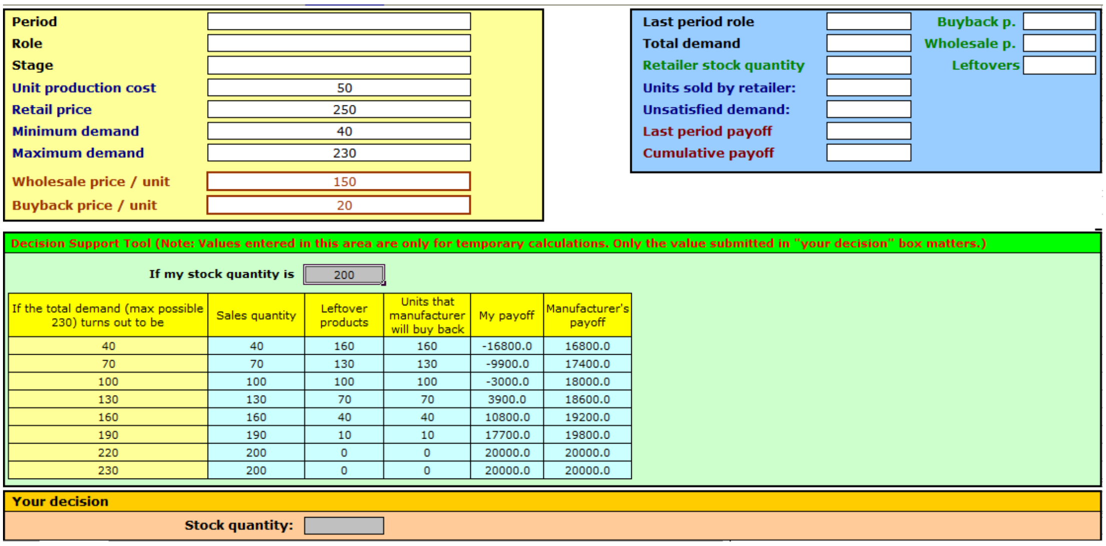
<u>A Sample Screenshot</u>

Figure 1 illustrates how the retailer's screen will look like at stage 2.

- The large table in the middle of the screen is your "**decision support tool**" (to be explained).
- The yellow box on the upper left presents general information including the period number, your current role, the wholesale price, and the buyback price that were set at stage 1. The

box also presents two game parameters that are given and fixed throughout all periods (unit production cost, and retail price).

- The blue box in the upper right presents information on the last period.
- The pink box in the bottom is where you "submit" your decision to the server. You enter your decision value into the related gray box, hit "enter" and then click on the green "Submit" button at the bottom (that will be visible during experiment). The submit button is activated only after you enter a valid decision and hit enter (or, click somewhere in the screen). Invalid entries will cause warnings.
- Note that the cells in which you can enter values are the ones with "gray" background.
- You can check the results of previous periods by clicking the "Historical Results" tab in the bottom. This will open a second worksheet with the titles seen below:



If the total demand (max possible 230) turns out to be	Sales quantity	Leftover products	Units that manufacturer will buy back	My payoff	Manufacturer's payoff
40	40	160	160	-16800.0	16800.0
70	70	130	130	-9900.0	17400.0
100	100	100	100	-3000.0	18000.0
130	130	70	70	3900.0	18600.0
160	160	40	40	10800.0	19200.0
190	190	10	10	17700.0	19800.0
220	200	0	0	20000.0	20000.0
230	200	0	0	20000.0	20000.0

Figure 0.1: Retailer's Screen at Stage 2

Period	Role	Wholesale price	Buyback price	Retailer stock quantity	Customer demand	Sales quantity of the retailer	Leftover products	Unsatisfied demand	Payoff	Cumulative payoff

Figure 0.2: Historical Results Screenshot

<u>The Decision Support Tool</u>

Before you submit a decision, you can use the "what-if" decision-support tool provided to you. This tool allows you to calculate the outcome for certain values of your decision, your opponent’s decision, and for specific realizations of the consumer demand. *Note that the values you enter in this area are only for your temporary calculations.* The only value that goes to the server (i.e., that is recorded) is the one you submit in the “stock quantity” box that you will find at the bottom of the screen.

The retailer’s decision support tool can be seen in Figure 1. You may enter a “stock quantity” value in the top gray cell. To help you visualize the possible outcomes if you really set this stock quantity, the table in the decision support tool summarizes the outcome for different consumer demand realizations (*d=40, 70, …, 230*), each in a row.

In the example above, the retailer’s stock quantity is entered as 200. We observe from the table that if consumer demand turns out to be, for example, 130, you (retailer) will sell 130 units because the demand is smaller than the stock quantity. You leftover inventory will be 200-130=70 units. The manufacturer will buy back these units. Since you satisfied all consumer demand, there will be no unsatisfied consumer demand.

Compare this with the outcome if consumer demand turns out to be 220. In this case, you (the retailer) will sell all of your 200 units, and there will be zero leftover inventory. Unsatisfied demand will be 220-200=20 units. As you sell your entire stock quantity, the manufacturer will not buy back any inventory. The last two columns provide your payoff and the manufacturer’s payoff.

At stage 1, the manufacturer’s decision support tool will look like below:

If my wholesale price is	150
and my buyback price is	20
and retailer's stock quantity is	200

If the total demand (max possible 230) turns out to be	Retailer's sales quantity	Leftover products at the retailer	Units that I should buy back	My payoff	Retailer's payoff
40	40	160	160	16800.0	-16800.0
70	70	130	130	17400.0	-9900.0
100	100	100	100	18000.0	-3000.0
130	130	70	70	18600.0	3900.0
160	160	40	40	19200.0	10800.0
190	190	10	10	19800.0	17700.0
220	200	0	0	20000.0	20000.0
230	200	0	0	20000.0	20000.0

Figure 0.3: Manufacturer's Decision Support Tool at Stage 1

At this stage, you (the manufacturer) will submit your wholesale price and buyback price. However, in order to use the decision support tool, you also need to guess what stock quantity the retailer might determine at stage 2.

Appendix D Manufacturer's Screen at Stage 1

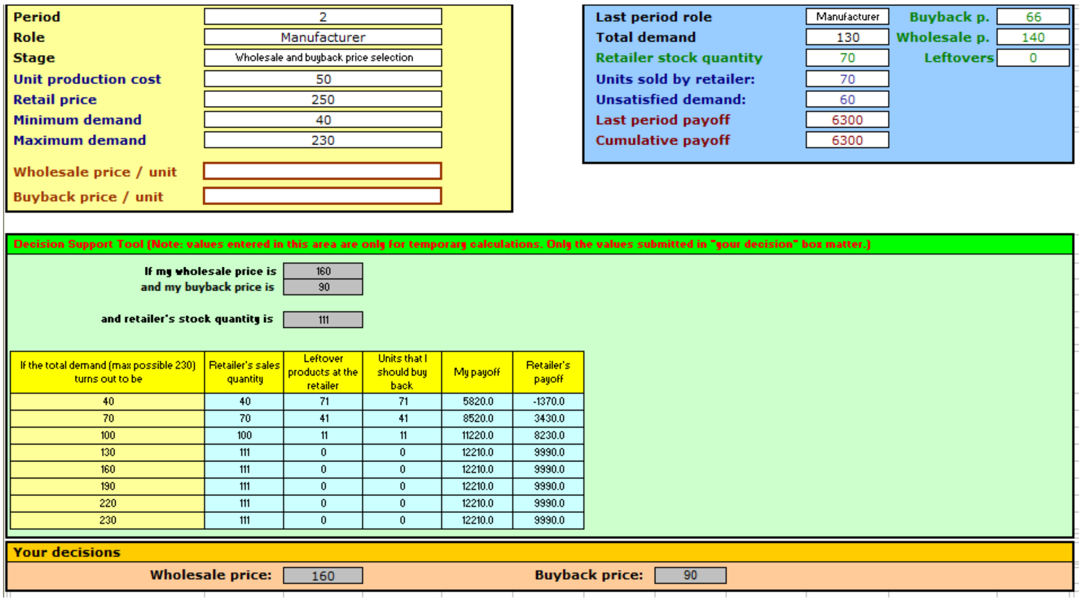
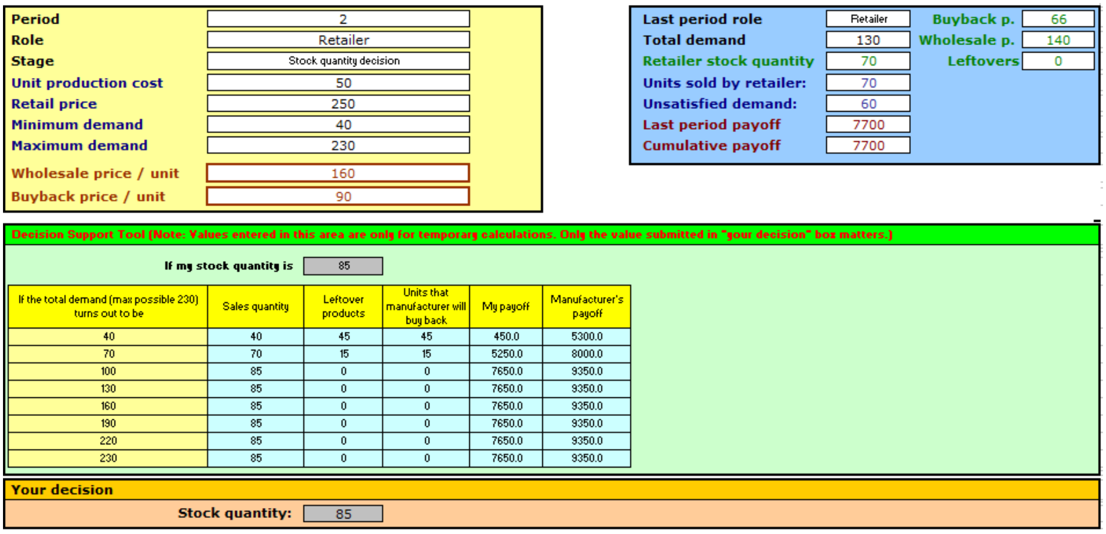


Figure 4: Manufacturer's Screen at Stage 1 Screenshot

Appendix E Retailer's Screen at Stage 2



Period	2
Role	Retailer
Stage	Stock quantity decision
Unit production cost	50
Retail price	250
Minimum demand	40
Maximum demand	230
Wholesale price / unit	160
Buyback price / unit	90

Last period role	Retailer	Buyback p.	66
Total demand	130	Wholesale p.	140
Retailer stock quantity	70	Leftovers	0
Units sold by retailer:	70		
Unsatisfied demand:	60		
Last period payoff	7700		
Cumulative payoff	7700		

Decision Support Tool (Note: Values entered in this area are only for temporary calculations. Only the value submitted in "your decision" box matters.)

If my stock quantity is 85

If the total demand (max possible 230) turns out to be	Sales quantity	Leftover products	Units that manufacturer will buy back	My payoff	Manufacturer's payoff
40	40	45	45	450.0	5300.0
70	70	15	15	5250.0	8000.0
100	85	0	0	7650.0	9350.0
130	85	0	0	7650.0	9350.0
160	85	0	0	7650.0	9350.0
190	85	0	0	7650.0	9350.0
220	85	0	0	7650.0	9350.0
230	85	0	0	7650.0	9350.0

Your decision

Stock quantity: 85

Figure 5: Retailer's Screen at Stage 2 Screenshot

Appendix F Results Screen

Visible to Manufacturer

Period	Role	Wholesale price	Buyback price	Retailer stock quantity	Customer demand	Sales quantity of the retailer	Leftover products	Unsatisfied demand	Payoff	Cumulative payoff
1	Manufacturer	140	66	70	130	70	0	60	6300	6300
2	Manufacturer	160	90	85	110	85	0	25	9350	15650

Figure 6: Manufacturer's Historical Result Sheet Screenshot

Visible to Retailer

Period	Role	Wholesale price	Buyback price	Retailer stock quantity	Customer demand	Sales quantity of the retailer	Leftover products	Unsatisfied demand	Payoff	Cumulative payoff
1	Retailer	140	66	70	130	70	0	60	7700	7700
2	Retailer	160	90	85	110	85	0	25	7650	15350

Figure 7**:** Retailer's Historical Result Sheet Screenshot