BEHAVIORAL EXPERIMENTS ON SUPPLY CHAIN CONTRACTING

by
ÜMMÜHAN AKBAY

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

SABANCI UNIVERSITY
AUGUST, 2016

BEHAVIORAL EXPERIMENTS ON SUPPLY CHAIN CONTRACTING

APPROVED BY:

Asst. Prof. Dr. Murat Kaya ....
(Dissertation Supervisor)

Assoc. Prof. Dr. Abdullah Daşcı ....
(Jury Member)

Assoc. Prof. Dr. Evrim Didem Güneş ....
(Jury Member)

Assoc. Prof. Dr. Kemal Kılıç ....
(Jury Member)

Prof. Dr. Taner Bilgiç ....
(Jury Member)

DATE OF APPROVAL: 10.08.2016

ABSTRACT

BEHAVIORAL EXPERIMENTS ON SUPPLY CHAIN CONTRACTING

Ümmühan Akbay

PhD Dissertation, July 2016

Thesis advisor: Asst. Prof. Dr. Murat Kaya

Each firm within a supply chain aims to maximize its own profit and unless the incentives of these firms are aligned through an efficient contract, individual profit maximization results in suboptimal performance for the supply chain. Researchers have developed a large number contracting models in order to coordinate the decisions of supply chain members through an appropriate allocation of profit and risk. The shortcoming of these models is that they presume decision makers are rational and self-interested profit maximizers. An extensive literature of experimental studies with human subjects reveal that decision makers do not conform to such standard theoretical expectations.

In this dissertation, we conduct and report the findings of decision-making experiments with human subjects in a supply chain setting. We investigate the effects of contract type, power of commitment and fairness priming on contracting decisions in a one-manufacturer-one-retailer scenario, where individual decision biases as well as strategic interaction affect subject behavior. We find significant deviations in subject decisions from theoretical predictions. We develop analytical models to incorporate the behavioral factors that cause these deviations. In another experimental work, we study gender differences in newsvendor decisions. Finally, we investigate the relation between contracting decisions and various personality traits.

Keywords

Supply chain management, supply chain contracts, behavioral experiments, behavioral operations management, newsvendor model, experimental economics, behavioral economics

ÖZET

TEDARİK ZİNCİRİ SÖZLEŞMELERİ ÜZERİNE DAVRANIŞSAL DENEYLER

Ümmühan Akbay

Doktora Tezi, Ağustos 2016

Tez Danışmanı: Yrd. Doç. Dr. Murat Kaya

Tedarik zinciri bünyesindeki her firma kendi kârını eniyilemek amacındadır ve bu firmaların çıkarları verimli bir sözleşme ile paralel hale getirilmediği sürece bağımsız kâr eniyileme çabası tedarik zinciri açısından kâr kaybına yol açar. Araştırmacılar, uygun kâr ve risk paylaşımı çerçevesinde tedarik zinciri firmalarının eşgüdümünü sağlamak için pekçok sözleşme modeli geliştirmiştir. Bu modellerin eksiği karar verenlerin rasyonel, sadece kendi çıkarları ile ilgilenen kâr eniyileyiciler olduğunu varsaymalarıdır. İnsan deneklerle yapılmış deneysel çalışmalardan oluşan geniş bir literatür göstermiştir ki karar verenler bunun gibi standart kuramsal beklentilere göre hareket etmemektedir.

Bu doktora tezinde insan deneklerle yaptığımız tedarik zinciri karar verme deneylerinin sonuçlarını sunuyoruz. Bir üretici ve bir perakendeciden oluşan, bireysel karar verme yargıları kadar bireyler arası stratejik etkileşimin de denek davranışlarını etkilediği basit bir tedarik zinciri senaryosunda sözleşme tipinin, taahhüt gücünün ve hakkaniyet koşullanmasının sözleşme kararları üzerindeki etkilerini inceliyoruz. Denek kararlarının kuramsal beklentilerden anlamlı olarak saptığını gösteriyoruz. Bu sapmalara neden olan faktörleri kapsayan yeni analitik modeller geliştiriyoruz. Başka bir deneysel çalışmada gazeteci çocuk kararlarında kadın ve erkek farklarını çalışıyoruz. Son olarak sözleşme kararları ve değişik kişilik özellikleri arasında bir ilişkiyi araştırıyoruz.

Anahtar Kelimeler:

Tedarik zinciri yönetimi, tedarik zinciri sözleşmeleri, davranışsal deneyler, davranışsal operasyon yönetimi, gazeteci çocuk modeli, deneysel iktisat, davranışsal iktisat

To all those who believed in me

Acknowledgements

A long time ago on a day I was full of doubt, a dearest friend of mine told me a story about Leonardo da Vinci. He said, "When asked how he made one of his beautiful sculptures, da Vinci answered 'I saw the angel in the marble and I carved until I let him free.'" "Likewise," added my friend, "I see a PhD in you."

It took a rather long time, lots of blood, sweat and tears to carve out the PhD in me, and now I would like to express my deepest gratitude:

…To my parents, for the endless love, unconditional support, and for giving me strength when I need it the most.

…To my dear sisters Esra and Ayşegül, for their love and friendship. To Esra for being a great role model despite being the younger one, and to Ayşegül for the much appreciated medical help.

…To Selin Çalışkan, Nursen Aydın, Mahir Umman Yıldırım, Ezgi Aylı, Merve Keskin, and all other residents of FENS 1021 from 2011 to 2015 for their friendship, all the help they have provided and listening to my complaints with patience.

…To Matulya Bansal and Nur Ayvaz Çavdaroğlu for their dear friendship and support.

…To Cornelia Gartner, John Goodwin and the three furry terrors, for their friendship, support and all the fur.

…To Prof. Ülkü Gürler of Bilkent University and Prof. Linda Green of Columbia University, for being excellent academic role models and for laying the foundations of my academic personality.

…To Prof. Özalp Özer of UT Dallas, for recommending me to take graduate level econometrics, which turned out to be an important tool I use frequently during my studies.

…To Prof. Bülent Çatay and Assoc. Prof. Guvenç Şahin for letting us conduct our experiments with their students.

…To all the students who volunteered to participate in our experiments and took the time to answer our surveys.

… To TÜBİTAK for providing us the funding to conduct our experiments through project # 111K454.

…To Prof. Hans Frenk and Assoc. Prof. Tonguç Ünlüyurt for their support.

…To Assoc. Prof. Abdullah Dasci and Assoc. Prof. Kemal Kilic for their support as committee members of my thesis and for their valuable feedback.

…To Prof. Taner Bilgiç and Assoc. Prof. Evrim Didem Güneş for agreeing to be on my thesis jury.

…To my thesis advisor Asst. Prof. Murat Kaya, first and foremost for introducing me to the field of behavioral operations, for generously sharing his knowledge and wisdom, for his guidance and patience over the last five year, for providing me the time and space to conduct independent research, and for teaching me how to conduct experiments, write academic papers and how to be a great thesis advisor.

…Finally, to anyone and everyone who has extended me the tiniest bit of help along the way and whose name is not listed above.

Table of Contents

ABSTRACT....................iv
ÖZET.................... v
Table of Contents....................ix
List of Tables....................xiv
List of Figures....................xviii
1. INTRODUCTION.................... 1
1.1 Methodology.................... 3
2. LITERATURE REVIEW.................... 5
2.1 Introduction.................... 5
2.2 History of Behavioral Operations Management.................... 5
2.3 Newsvendor Decision Making Experiments.................... 7
2.3.1 Behavioral Theories and Decision Making Biases.................... 7
2.3.2 Decision Making Heuristics.................... 8
2.3.3 Learning.................... 9
2.3.4 Feedback.................... 9
2.3.5 Decision Set and Decision Frequency.................... 10
2.3.6 Experiment Subjects.................... 10
2.3.7 Demand Distribution.................... 11
2.3.8 Framing.................... 11
2.3.9 Process Tracing.................... 12
2.3.10 Performance Measures.................... 12
2.4 Supply Chain Contracting Experiments.................... 12
2.4.1 Supply Chain Contracts.................... 13
2.4.2 Social Interaction.................... 14
2.5 Experiments and Analysis.................... 15
2.5.1 Analysis Methods.................... 15
2.5.2 Experimental Design.................... 16
2.6 Conclusion.................... 16
2.7 References.................... 17
Appendix-A: Classification of Behavioral Newsvendor Literature.................... 23
Appendix-B: Classification of behavioral supply chain contracting literature.................... 30

3. EXPERIMENTS ON SUPPLY CHAIN CONTRACTING: EFFECTS OF CONTRACT TYPE .. 37
Abstract 38
3.1 Introduction 39
3.2 Related Literature 40
3.2.1 The Newsvendor Model 40
3.2.2 Supply Chain Contracting and Coordination 42
3.3 Analytical Background and Experimental Procedure 44
3.3.1 The Analytical Model 44
3.3.2 The Theoretical Solution of the Model 45
3.3.3 Experimental Procedure 47
3.4 Contract Comparisons with Theory and with Each Other 49
3.4.1 Comparisons with Theoretical Predictions 49
3.4.2 Comparisons between the Contract Types 50
3.5 Explaining Differences in Contract Performance 51
3.5.1 Effect of Manufacturer’s Contract Parameter Decision 53
3.5.2 Effect of Retailer’s Order Quantity Decision 55
3.5.3 Effect of Retailer’s Under/Overorder Behavior on Manufacturer’s Realized Profit 57
3.6 Contract Rejections and Fairness Considerations 60
3.6.1 Rejected Contracts Analysis 60
3.6.2 Retailer’s Response to Inequality 61
3.7 Conclusions 65
3.8 References 67
3.9 Appendix 70
3.9.1 Sample Experiment Instructions (Buyback Contract) 70
3.9.2 Comparison Results when Rejected Contracts Included 73
4. DIFFERENT UTILITY FUNCTIONS 75
4.1 Introduction 76
4.2 Theoretical Background 77
4.3 Retailer’s Risk Aversion Models 80
4.3.1 Model Estimation 82
4.4 Retailer’s Loss Aversion Models 85
4.4.1 Loss-Aversion Model -1 86

4.4.2 Loss-Aversion Model -2 86
4.4.3 Loss-Aversion Model -3 87
4.4.4 Model Estimation 88
4.5 Retailer’s Inventory Error Models 91
4.5.1 Waste Aversion Model 91
4.5.2 Stock-out Aversion Model 92
4.5.3 Minimizing Ex-post Inventory Error Model 92
4.5.4 Model Estimation 93
4.6 Retailer’s Social Preference Models 96
4.6.1 Fairness Model 96
4.6.2 Status Seeking Model 97
4.6.3 Group Identity Model 98
4.6.4 Model Estimation 98
4.7 Manufacturer’s Social Preference Models 102
4.7.1 Fairness Model 102
4.7.2 Status Seeking Model 102
4.7.3 Group Identity Model 102
4.7.4 Model Estimation 103
4.8 Conclusion 104
4.9 References 105
5. POWER OF COMMITMENT AND IMPACT OF PRIMINT ON SUPPLY CHAIN RELATIONSHIPS 106
Abstract 107
5.1 Introduction 108
5.2 Related Literature 109
5.3 Analytical Background and Experimental Procedure 112
5.3.1 Analytical Background 113
5.3.2 Experimental Procedure 114
5.4 Experimental Study 1: Power of Commitment 115
5.4.1 Comparison with Theoretical Benchmarks (Accepted Contracts) 116
5.4.2 Manufacturer’s Commitment to the Wholesale Price 117
5.4.3 Retailer’s Commitment to Quantity 119
5.4.4 Comparison of Rejected Contracts 120

5.4.5 Modelling Retailer's Behavior 122
5.5 Experimental Study 2: Fairness Priming 124
5.5.1 Comparison with Theoretical Benchmarks (Accepted Contracts).......... 125
5.5.2 Priming Effect 125
5.5.3 Comparison of Rejected Contracts 127
5.5.4 Modelling Retailer's Behavior 129
5.6 Conclusions 130
5.7 References 131
5.8 Appendix 135
5.8.1 Sample Experiment Instructions (Standing Price Scenario) 135
5.8.2 Comparison Results when Rejected Contracts Included 139
6. GENDER DIFFERENCES ON NEWSVENDOR BEHAVIOR 141
6.1 Introduction 142
6.2 Related Literature 143
6.3 Analytical Background and Experimental Design 145
6.3.1 Newsvendor Model and the Theoretical Solution 145
6.3.2 Experimental Design 146
6.4 Results 147
6.4.1 Descriptive Statistics 148
6.4.2 Order Variability 152
6.4.3 Use of Decision Heuristics 153
6.4.4 Aggregate Level Analysis 160
6.5 Conclusions 161
6.6 References 162
6.7 Appendix: Instructions for Experiment (High Profit Margin) 165
7. EFFECTS OF PERSONALITY TRAITS ON SUPPLY CHAIN DECISIONS 168
7.1 Introduction 169
7.2 Related Literature 169
7.3 Experimental Design 170
7.4 Self-Esteem Analysis 171
7.4.1 Newsvendor Experiment Results 174
7.4.2 Contracting Experiment Results 175
7.5 Regret Aversion Analysis 177

7.5.1 Newsvendor Experiment Results 180
7.5.2 Contracting Experiment Results 180
7.6 Risk and Loss Aversion Analysis 183
7.6.1 Newsvendor Experiment Results 186
7.6.2 Contracting Experiment Results 188
7.7 Inequity Aversion Analysis 191
7.7.1 Experiment Results 192
7.8 Conclusion 195
7.9 References 196
8. CONCLUSION 198
8.1 Comparison with a Randomized System 199
8.1.1 Newsvendor Experiment 199
8.1.2 Contracting Experiments 200
9. REFERENCES 204

List of Tables

Table 1: Comparison of theoretical outcomes under different contracts 47

Table 2: Experimental design 47

Table 3: Comparison of the three contract type’s experiment results with theory 50

Table 4: Comparison of contracts with each other 51

Table 5: Comparison of manufacturer's newsvendor-predicted profit 54

Table 6: Comparison of retailer's deviation from Q* 55

Table 7: Comparison of the manufacturer’s ratio of expected profit to predicted profit 56

Table 8: Underorder and overorder occurrences (Pooled data) 58

Table 9: Underorder and overorder magnitudes: Contract comparisons and p-values... 59

Table 10: Comparison of rejected contracts 61

Table 11: Definition of variables for the regression equation 62

Table 12: Random effects ordered logistic regression results for the retailer's decisions 63

Table 13: Sample computation for regression results 65

Table 14: Estimation results of the risk attitude model for WSP retailers 83

Table 15: Estimation results of the risk attitude model for BB retailers 84

Table 16:Estimation results of the risk attitude model for RS retailers 85

Table 17: Estimation results of loss aversion models for WSP retailers 89

Table 18: Estimation results of loss aversion models for BB retailers 89

Table 19: Estimation results of loss aversion models for RS retailers 90

Table 20: Estimation results of inventory error models for WSP retailers 94

Table 21: Estimation results of inventory error models for BB retailers 95

Table 22: Estimation results of inventory error models for RS retailers 95

Table 23: Estimation results of social preference models for WSP retailers 100

Table 24: Estimation results of social preference models for BB retailers 100

Table 25: Estimation results of social preference models for RS retailers 101

Table 26: Estimation results of social preference models for WSP manufacturers...... 104

Table 27: Theoretical solution of wholesale price contract scenario 114

Table 28: Comparison of treatments 116

Table 29: Comparison of experiment results with theory (Base, SP and SPQ treatments) 117

Table 30: Comparison of experiment results between the Base and SP treatments 118
Table 31: Comparison of experiment results for SP and SPQ treatments 119
Table 32: Comparison of rejected contracts (Base, SP, and SPQ treatments).............. 120
Table 33: Rejection rates (Base, SP, SPQ treatments) 121
Table 34: Variable definitions 122
Table 35: Regression results for the retailer's behavior (power of commitment)......... 123
Table 36: Sample estimation of probabilities for results in Table 8 124
Table 37: Comparison of FP treatment results with theory 125
Table 38: Comparison of experiment results for Base treatment and Fairness Priming treatment 126
Table 39: Comparison of rejected contracts (Base and FP treatments)........................ 127
Table 40: Rejection rates (Base and FP treatments) 128
Table 41: Regression results for the retailer's behavior (Fairness Priming) 129
Table 42: Sample estimation of probabilities for results in Table 15 130
Table 43: Experiment parameters and theoretical solution 146
Table 44: Sample sizes of the treatments.. 147
Table 45: Comparison of the order quantity and profit with theory 148
Table 46: Comparison of the expected profit with theory 149
Table 47: Order quantity and risk-aversion regression results 150
Table 48: Gender comparison of the experiment performances 151
Table 49: Gender comparison of order standard deviation... 152
Table 50: Gender comparison of unique order counts.. 152
Table 51: Comparison of subjects whose order quantities changed from first 10 to last 10 periods.. 153
Table 52: Gender comparison of orders in and out of the PTC region......................... 155
Table 53: Gender comparison of estimated mean anchor heuristic α coefficients 155
Table 54: Individual level comparison of mean anchor α coefficients......................... 156
Table 55: Percentage of subjects with modal order of 100 and Q^* 156
Table 56: Percentage of orders changed towards and away from the previous demand 159
Table 57: Gender comparison of demand chasing regression estimations 159
Table 58: Percentage of subjects with orders correlated with the demand................... 160
Table 59: Aggregate regression results... 161

Table 60: Number of subjects participating in the survey study 170
Table 61: Number of pairs participating in the survey study for each treatment 171
Table 62: Rosenberg's self-esteem scale... 173
Table 63: Evaluation of Rosenberg's self-esteem scale 173
Table 64: Self-esteem comparison results of the newsvendor experiment................... 174
Table 65: Self-esteem comparison results of the contracting experiment retailers 175
Table 66: Self-esteem comparison results of the contracting experiment manufacturers 176
Table 67: Schwartz et al.'s regret scale 179
Table 68: Evaluation of Schwarz et al.'s regret scale 179
Table 69: Regret-tendency comparison of the newsvendor experiment....................... 180
Table 70: Regret tendency comparison of the pull-to-center effect for contracting experiment retailers.. 181
Table 71: Regret-tendency comparison of the contracting experiment retailers 182
Table 72: Regret-tendency comparison of the contracting experiment manufacturers 183
Table 73: Risk and loss aversion survey questions.. 185
Table 74: DOSPERT scale questions 186
Table 75: Risk and loss aversion comparison of the newsvendor experiment 187
Table 76 DOSPERT scale risk aversion comparison of newsvendor experiment........ 187
Table 77: Question 2 Risk aversion comparison of the contracting experiment retailers 188
Table 78: Question 4 Risk and loss aversion comparison of the contracting experiment retailers.. 189
Table 79: Risk aversion comparison of the contracting experiment manufacturers – Question 2 189
Table 80: Risk and loss aversion comparison of the contracting experiment manufacturers – Question 4 190
Table 81: Inequity aversion survey questions.. 192
Table 82: Inequity aversion comparison of the contracting experiment retailers – high sensitivity.. 193
Table 83: Inequity aversion comparison of the contracting experiment retailers – moderate sensitivity 194
Table 84: Inequity aversion of contracting experiment manufacturers 194

Table 85: Comparison of the experiment results with randomized system – Newsvendor experiment 200
Table 86: Comparison of the experiment results with randomized system – WSP contract 201
Table 87: Comparison of the experiment results with randomized system – BB contract 201
Table 88: Comparison of the experiment results with randomized system – RS contract 202

List of Figures

Figure 1: The supply chain under buyback contract 44
Figure 2: Sample manufacturer screen (Buyback contract) 48
Figure 3: Inter-level comparison of the manufacturer’s cumulative profit distribution . 52
Figure 4: Inter-contract comparison of the manufacturer’s cumulative profit distribution 53
Figure 5: Comparison of manufacturer's expected and newsvendor-predicted profit values 57
Figure 6: Retailer's cumulative newsvendor-predicted profit share (**A**: Accepted, **R**: Rejected) 61
Figure 7: Order decision vs *Q*(w,b)* under the buyback contract 77
Figure 8: Manufacturer's expected profit under the buyback contract 79
Figure 9: Manufacturer's expected profit under the revenue sharing contract 79
Figure 10: Expected profits of the supply chain under the wholesale price contract 80
Figure 11: Cumulative distribution of newsvendor-predicted profit share of the retailer (Base, SP, SPQ) 121
Figure 12: Post-rejection changes in wholesale price (Base, SP and SPQ treatments) 122
Figure 13: Cumulative distribution of newsvendor-predicted profit share of the retailer (Base and FP treatments) 128
Figure 14: Post-rejection changes in wholesale price (Base and FP treatments) 128
Figure 15: Retailer’s screen at stage 2 137
Figure 16: Expected profit as a function of the order quantity 149
Figure 17: Histogram and cumulative distribution of order decisions (Pooled data) ... 151
Figure 18: Percentage of orders in and out of the PTC region 154
Figure 19: Evolution of the average order quantity over time 158
Figure 20 Screenshot of Excel sheet 167

Chapter 1

1. INTRODUCTION

Supply chains consist of individual firms, each aiming to maximize its own profit. As a result of this decentralized individual decision making, supply chains achieve less than optimal profit. This is why the study of contracts between supply chain members has attracted great attention in business as well as in academic literature. An efficient contract can align the incentives of the individual firms, leading to higher overall efficiency and higher gains for all parties, including the end-consumers. More than simply a pricing agreement, the contract is a tool to share profits, risks and information.

Supply chain contracting and coordination literature has studied many different types of contracts and produced a wealth of analytical models (Cachon, 2003; Kaya and Ozer, 2012). All of these models are based on a number of behavioral assumptions regarding how people make decisions (rational decision makers who aim to maximize their expected utility), and how people strategically interact (game-theoretic equilibrium concepts). While widely used in modeling, experimental economists have been challenging these assumptions through controlled experiments with human decision makers (Kagel and Roth 1995). Such experimental studies have uncovered significant differences between human decisions and the predictions of analytical models. In fact, analytical models' inability to explain and predict human behavior has caused a significant gap between supply chain contracting research and practice. Experiments are particularly valuable for supply chain contracting research where field studies are extremely difficult to conduct. Experiments uncover the gaps between theoretical predictions and human decisions, allowing the development of better analytical models that have higher explanatory and prediction power.

In this dissertation, we conduct and report the results of decision making experiments with human subjects on the simplest supply chain setting that involves inventory risk and contracting. The rest of the dissertation is organized as follows:
Chapter 2 reviews the relevant behavioral operations literature and compares the articles in terms of behavioral variables, behavioral theories, research questions, experimental factors, results and behavioral implications.

Chapter 3 focuses on the effects of contract type on supply chain contracting. On a one-manufacturer-one-retailer supply chain scenario, we compare the performance of the wholesale price, buyback and revenue sharing contracts and present a detailed analysis of the contract differences on the manufacturer and the retailer's decision levels.

Chapter 4 develops several analytical models based on expected utility theory for the retailers and the manufacturers of the study presented in Chapter 3. The goal here is to incorporate behavioral factors that might be affecting subject behavior into the theory.

Chapter 5 studies power of commitment and impact of fairness priming on supply chain relationships. We find that the firm committing to its respective decision obtains a strategic advantage over the other firm and increases its profit. As for fairness priming, contrary to our expectation, the priming resulted in a less equitable allocation of the profits.

Chapter 6 presents the findings of a newsvendor experiment and focuses on gender differences. We find that female subjects place smaller orders than the males do under high profit margin. The difference is not significant under low profit margin. We also show that female subjects are more prone to demand-chasing heuristic than male subjects.

Finally, Chapter 7 investigates the correlation between various personality traits and results of the experiments presented in Chapters 3, 5 and 6. We study the effects of self-esteem, regret-aversion, risk and loss aversion, and inequity aversion. We find meaningful differences in subject decisions.

1.1 Methodology

Here we briefly list the various analysis methods used in this dissertation.

1. **Non-parametric statistics**

   We do not make any distributional assumptions about the decisions of the experiment subjects. Hence, for all comparisons regardless of the sample size, we use nonparametric hypothesis tests. For one sample, we use the Wilcoxon signed rank test, and for two independent samples we use Mann Whitney U test. The exception is the significance tests for linear regression analyses, where we use t-distribution with the appropriate degrees of freedom. (For more information see Siegel 1956.)

2. **Linear Regression Models**

   In *Chapter 5* we conduct linear regression analyses on individual subject data in order to estimate the various behavioral models studied in the chapter. In *Chapter 6* we use linear regression to compute individual coefficients for the mean anchor and the demand chasing decision heuristics. In *Chapter 3 Section 3.5.2*, we use linear regression to briefly compare the retailer's response to the offered contract.

3. **Advanced Regression Models**

   In order to take into account individual differences while drawing an overall conclusion on the experiment treatments, we make use of advanced linear regression models such as fixed-effects and random-effects models.

   3.1. Fixed effects models

   In these models, the explanatory variables are allowed to be correlated with the individual variability. This analysis tool is used in *Chapter 6 Section 6.4.4.*

   3.2. Random-effects models

   In these models, the explanatory variables are assumed to be independent of the individual variability. In this dissertation we don't use a random-effects model as a linear regression analysis, rather we use it in a random-effects ordered-logit model.

3.3. Ordered Logit models

Here the dependent variable is not linear but rather restricted to an ordinal discrete set of categories. The regression model estimates the coefficients to maximize the joint probability of all observations falling onto their respective categories. The value assigned to the category is irrelevant to the computation as long as the correct order is preserved. We use random-effects ordered-logit models in *Chapter 3 Section 3.6.2*, *Chapter 5 Sections 5.4.5* and *5.5.4*. In these models the dependent variable is the retailer's order decision and it is categorized to the set {contract rejection, underorder, near optimal and overorder}.

Chapter 2

2. LITERATURE REVIEW

2.1 Introduction

Newsvendor problem is one of the simplest yet powerful models in inventory control and supply chain management. Its power results from its economic intuitiveness and applicability to a wide variety of real life problems. It has a very well established simple and nice solution which can be easily computed. The model provides the optimal order policy that will maximize the expected profit in a stochastic demand environment where unsold products will lose their value, and unmet demand will be lost. Despite its simplicity, human decision makers have been observed to systematically deviate from the optimal orders. Within the last decade, researchers have conducted experiments and studied this model to uncover the factors that are in effect in this suboptimal behavior.

Here we review the majority of the literature of newsvendor decision making and supply chain contracting experiments. We provide a detailed comparison of aforementioned research papers in Appendices A and B.

2.2 History of Behavioral Operations Management

For centuries experiments have been used in scientific methodology first to answer questions about nature of reality, then to formulate, test and verify and thus establish the validity of theories, ideas and hypotheses. Behavioral experiments with human subjects have been an extensively utilized tool in behavioral economics and psychology fields.

However the history of experiments with human decision makers in operations setting does not extend to far away.
It all started with Schweitzer and Cachon's study in 2000, which was motivated by behavioral economics models. Schweitzer and Cachon conducted two experiments with MBA students and revealed that human decision makers are systematically deviating from the optimal order quantity in a newsvendor setting. Not only the deviations were systematic and robust with respect to different settings of gain or loss domains, but they were inexplicable by the existing behavioral theories. The results of Schweitzer and Cachon's experiments were disclosing a "too low too high" pattern in which the subjects were understocking in high profit margin and overstocking in low profit margin frames. In other words, the subjects facing a high profit margin (which theoretically implies an optimal order quantity greater than the mean of the uniform demand distribution assumed in the study) were ordering less than the optimal, whereas the subjects facing low profit condition were ordering more than the optimal order quantity which is less than the aforementioned demand mean. It was as if the order quantity decisions were pulled from the optimal quantity towards the mean of the demand distribution. Hence, Schweitzer and Cachon named this too low too high pattern as the "pull to center" effect.

These order quantities under the effect of pull to center did not comply with most of the possible theoretical explanations. Risk aversion would explain high profit margin orders being less than optimal but did not explain the overstocking in low profit margin. Similarly none of loss aversion, waste aversion, stock-out aversion, undervaluing opportunity costs and such well-known theories were successful in explaining this pattern. In addition, there was some weak support for decision making heuristics such as mean anchor and demand chasing but it was only enough to explain a small portion of the deviations.

What was really behind this systematic and robust deviation from the optimal? What was determining the direction and magnitude of the deviations? Could the factors behind this sub-optimality be identified and incorporated into the theory? Was the newsvendor model really realistic? And if the simplest inventory model used as a building block in much more complicated models was not realistic, what would happen to the whole supply chain management theory based on that model?

And there it was, an intriguing and promising research question sitting at the core of most important operations management models with all its appeal and waiting for the scientists to solve it. So it began.

2.3 Newsvendor Decision Making Experiments

Newsvendor experiments are concerned with the decision making behavior of the individuals who face a random demand for a perishable product. In these experiments, the individual can only decide on the order quantity as both the selling price and the purchasing cost are determined by the experimenter. Before the experiments the subjects are usually given training on the newsvendor problem, and their level of understanding of the model is tested before proceeding to the experiment.

Next we review focus, subject and method of the newsvendor experiments.

2.3.1 Behavioral Theories and Decision Making Biases

Bias, by definition, means an inclination to present or hold a partial perspective at the expense of possibly equally valid alternatives. Decision biases are the ones that govern the decision making process of the individuals. For instance, an individual biased against risk would prefer less risky outcomes over possibly better ones. These biases which govern the individual decision making play an important role Newsvendor experiments. To name some of these decision biases and articles that study them:

Bounded Rationality: In economic theory, a rational agent refers to a human being with clear preferences who acts to maximize her total utility. In other words a rational agent is easily able to compare two options and identify the one which will provide her with higher utility. When faced with several products with several features, such a rational agent will not suffer from indecisiveness and easily pick the one that offers the best value to her. A rational agent not only knows what is best for her, but acts to attain this best. Any other behavior would be called irrationality.

In operations management, the existing literature has based their studies on the assumption that decision makers are rational agents who will make the decisions that will maximize their utility. The term bounded rationality refers to the case that the decision maker will pick the optimal option not for sure, but with some probability. The decision maker in this case is not irrational, but has cognitive limitations so her rationality is bounded.

Su (2008), studies bounded rationality by means of a quantal choice model and is able to explain the pull-to-center effect by this model.

Prospect Theory: Also known as context-dependent preferences. The individual's behavior is different under different contexts. For instance the individual may be risk averse when she has a higher initial wealth, that is, when she has a lot to lose, while she may be risk seeking when she has nothing much to lose. As explained by Schweitzer and Cachon (2000), prospect theory is able to explain the too-low-too-high pattern.

Loss Aversion: The decision bias of placing more (negative) weight on losses than gains is studied in Wang and Webster (2009). The authors show that if the opportunity cost is not negligible in a newsvendor setting, a loss averse individual will stock more than a risk neutral individual.

2.3.2 Decision Making Heuristics

Heuristic is an experience-based method for problem solving. Decision making heuristics in our context are the ones that individuals make use while placing their order decisions. The most common heuristics found in BOM literature are anchoring and demand chasing heuristics, which is followed by the minimizing ex-post inventory error heuristic.

Anchoring heuristic leads the individual to base her decision at a reference point and then adjust towards another point or towards demand realization. Mean anchor heuristic, for instance, means that the individual has the mean of the demand distribution as the anchor point and adjusts towards the optimal order quantity. Demand chasing heuristic can also be considered as an anchoring heuristic as it anchors at the previous period's order quantity and adjusts towards the most recent realization of the demand.

Gavirneni and Xia (2009) study the anchoring behavior of the newsvendors by providing them with several possible reference points. The authors find that the subjects anchor at the reference point which is closest to the optimal order quantity.

2.3.3 Learning

Probably the first question about the newsvendor decisions that comes into mind is whether there is any learning, in other words, whether the subjects are improving their performance with time, and if so, how fast this learning effect is. Experiments in literature that have only 30 or 40 rounds (such as Schweitzer & Cachon, 2000, Schultz et al., 2007) do not find any trend towards improving the performance.

Yet the absence of any improvement in the decision making performance over time in these studies may be a result of the learning effect not having kicked in yet. As such, in Bolton and Katok (2008) study, where the experiment was replicated up to 100 rounds, we can observe the clear impact of learning over time.

2.3.4 Feedback

What feedback the individuals receive after decision making and how frequently they receive feedback has an important role on the newsvendor performance. Bolton and Katok (2008) and Lurie and Swaminathan (2009) study the effect of feedback frequency on the newsvendor decision making. They find that more frequent feedback does not necessarily improve performance, on the contrary, it might degrade it. However too infrequent feedback also degrades performance as shown by Lurie and Swaminathan (2009). Bolton and Katok (2008) also study the impact of reinforced feedback with additional information on the counterfactual payoffs for the forgone options. They find that such reinforced feedback does not have a significant effect on the newsvendor performance.

2.3.5 Decision Set and Decision Frequency

In most of the experiments the decision options for the order quantity are not restricted, except for the range of possible demand realizations. Bolton and Katok (2008) restrict the decision options to test the hypothesis that decision makers are placing suboptimal orders due to their limited capabilities, hence, when the decision set is restricted newsvendor performance should be improved. The authors find no significant improvement. However in their decision sets the optimal order quantity is one of the extreme options. Thus, extremeness aversion is a possible explanation of newsvendors not picking the optimal quantity.

Feng et al. (2011) follow the same restricting method but making sure that the optimal order quantity is not an extreme option in the set. These authors show that the newsvendor performance is improved.

Decision frequency is the frequency of changes in order quantity. Bolton and Katok (2008), restrict the order decisions to be same for 10 periods and illustrate that standing orders significantly improve newsvendor performance.

2.3.6 Experiment Subjects

The subject pool is the most important component of experimental studies. In Behavioral Operations Management experiments mostly students, undergraduate or MBA, are used as subjects. There have been some concerns about how realistic the results obtained from student experiments will be in the real world. Bolton et al. (2012) answered these concerns by showing that there is no significant difference in the newsvendor performance between managers and students when the demand distribution is given.

De Vericourt et al. (2001) studied the gender differences and found that female subjects place significantly less orders than male subjects. The authors explain this with these subjects' being more risk averse than male subjects rather than being female.

Another study about the subject pools is the one comparing subjects from different cultures by Cui et al. in (2011). They show that Chinese subjects ask more questions before making a decision, which signals that Chinese subjects are more worried about making mistakes. Also the Chinese are more cognizant than American subjects, however they are not able to identify out-of-the box, creative solutions. Feng et al. (2011) also studied cultural differences between Chinese and American subjects in terms of newsvendor decision making. They find that the pull-to-center effect is more prominent in the Chinese than the Americans.

2.3.7 Demand Distribution

In most of these articles the demand distribution is given and is discrete uniform to make sure it is simple enough for the subject pool to understand and work with. There are a few articles that study Normal distribution. One such example is Benzion et al. (2008), however they find the use of Normal distribution not to affect the newsvendor performance.

Rudi and Drake (2011) study the case where the demand is censored, meaning that if the demand is higher than the stocking level, the information about lost demand is censored, which is the case in most real life practices. They show that demand censoring leads to lower order quantities and stronger demand chasing. Benzion et al. (2010), on the other hand study the case where the demand distribution is unknown, which is again a practically realistic assumption. Their results show that the behavior of the subjects who know and don’t know the distribution are different, however knowing the distribution does not necessarily improve the newsvendor performance. Supply surplus has a stronger impact on the ordering decisions possibly due to waste aversion bias.

2.3.8 Framing

Framing refers to the way a problem is presented to the subjects. One very famous example of framing is the tale of a king who once had a dream and wanted to have it interpreted. One interpreter told him that the king would see the deaths of all his loved ones and would be very lonely afterwards. Hearing this, the king was furious and had the interpreter executed. Another interpreter came and the king repeated his dream to this interpreter. However this interpreter, told the king that he would live a long and healthy life. The king was very happy to hear this interpretation and had the man rewarded with gold.

As extreme as the above anecdote might seem, in newsvendor experiments as well as in real life applications, the framing of the problem plays a crucial role. Kremer (2008) study the impact of neutral or operations frame of the newsvendor model, and show that the decision errors are significantly smaller in the neutral frame. Schultz et al. (2007) study the framing effects of cost or profit emphasis and show that framing does not significantly affect the performance.

2.3.9 Process Tracing

Process tracing is the method of tracing the thought process of the individuals while making the ordering decision via verbal protocols or technically keeping track of which pieces of information is acquired. Gavirneni and Isen (2010), and Cui et al (2011) pursue such a method and show how individuals make their ordering decisions.

2.3.10 Performance Measures

In most of these studies the performance measure is the order quantity compared to the optimal quantity. Some papers use the probability of making the optimal decision and the proportion of the maximum expected profit achieved. Brown and Tang (2006) demonstrate that with a performance measure of targeting a certain profit level, the suboptimal ordering behavior can be explained.

2.4 Supply Chain Contracting Experiments

In practice and in theory contracts are used to legalize the agreement between two (or more) parties. In supply chain management context the contracts govern the trade between two firms one of which is supplying the other with either raw material of semi-finished products. The buying firm will sell its products to the end customer or to another firm.

In Behavioral Operations Management literature mostly two-echelon supply chains are considered in experimental studies. The upstream firm is generally referred to as the supplier or the manufacturer and the downstream firm is referred to as the retailer in these studies. Supply chain models with three or more echelons are also considered within BOM studies, to investigate the behavioral properties of the bullwhip effect, but that is beyond the scope of this study.

2.4.1 Supply Chain Contracts

Contracts in supply chain literature can be categorized depending on which firm holds the power of the contract at hand. In Push contracts the supplier is more powerful and therefore the contract terms are determined by the supplier. In Pull contracts the retailer is more powerful and determines the contract terms.

The wholesale price contract is the most famous and most commonly used push contract. It is also the simplest contract with only one term: the wholesale price. In theory, due to double marginalization between the supplier and the retailer the wholesale price contract falls short of coordinating the chain and thus it is not efficient. For that reason several other contracts which are capable of coordinating the chain are designed. Two of these are the buyback and the revenue share contracts.

Although in theory buyback and revenue share contracts are equivalent n experiments they prove otherwise. In a study run by Katok and Wu (2009), buyback and revenue sharing contracts are shown to perform differently, and that they both fail to coordinate the supply chain.

Other examples of Push contracts are quantity discount, sales rebate, quantity flexibility and tariffs. Ho and Zhang (2008) study the empirical differences between wholesale, quantity discount and two part tariff contracts. They show that double marginalization is higher in wholesale price contract, and quantity discount contract generates higher channel efficiency than the two part tariff. In another study, Lim and Ho (2007) compare the impact of the number of blocks in tariffs on the manufacturer's share of profit and the supply chain profit. Kalkanci et al. (2011) show that more complex contracts do not necessarily improve the supply chain performance and in fact a simple wholesale contract is sufficient with asymmetric information.

Davis (2010) studies the behavioral properties of the Pull type contracts which are namely the wholesale price, overstock allowance and service level agreement. They illustrate that while the wholesale price contract is significantly inferior to overstock allowance and service level agreement contracts, these two contracts are not significantly different from each other. Kalkanci et al. (2012) study the supplier's preferences on contract complexity and show that human-to-human interactions strengthen the supplier's preference for simpler contracts.

2.4.2 Social Interaction

When subjects are interacting with each other instead of interacting with a computer, their preferences and decisions are affected. There are several dimension of this interaction.

2.4.2.1 Social Preferences

The literature assumes that human beings are self-interested, utility maximizing, completely rational decision makers who doesn't have any concern about what happens to other human beings. That is in a way assuming that humans act like computers while making their decisions. However humans make their decisions using not only their minds but also with their hearts, and sometimes the heart might overwrite the mind's decisions.

What the standard theory oversees is that human beings derive utility from intangible and noneconomic goods. For instance helping someone makes one feel better and the reward is immediate. As such, some social preferences effect decisions in the supply chain contracting context, namely, fairness concerns or inequity aversion, reciprocity, relationship, status seeking, group decision making.

In Keser and Paleologo (2004)'s human-to-human supplier-retailer experiments, the supplier charges lower wholesale prices, which results in more equitable distribution of the supply chain profits. This in turn signals fairness concerns in supplier decision making. Loch and Wu (2008) show that efficiency is higher when the subjects meet before the experiment and shake hands, signaling the relationship and reciprocity preferences of the individuals.

2.4.2.2 Duration of Interaction

If the interaction between the subjects is not one shot, there will be a long term relationship and if the subjects know about this, subjects' decision making behavior changes substantially. Wu (2012) show that when future opportunities to punish are available, fairness and reciprocity concerns are reinforced and reputation building behavior is motivated to achieve long term economic benefits. Hence the overall supply chain performance is improved. Hyndman et al. (2012) show that long term relationships are better for aligning capacity decisions, while profits are variable and learning is slower. On the other hand, Şahin (2011) find evidence that one shot relationships are better than long term relationships.

2.4.2.3 Information

When the demand forecast is private information, Özer et al. (2011) show that trust significantly affects the outcome of cheap talk forecast communication and improves the overall channel efficiency. In another study the authors compare the trust and trustworthiness properties of Chinese and American subjects and find Chinese subjects to exhibit lower trust and trustworthiness.

2.5 Experiments and Analysis

In this section we discuss the experiment design techniques and analysis methods.

2.5.1 Analysis Methods

Statistics is the major analysis tool of the behavioral operations management experiments. Basic data analysis and hypothesis testing can be found in almost every BOM article. In statistical analysis, if the data set is not large enough and if some certain assumptions are not satisfied, or if the researcher does not want to make any assumptions on the distribution of the observations, non-parametric statistical methods are to be used. Accordingly many of the articles mentioned in this paper use non-parametric statistical methods. However, parametric statistical methods are also used as much as non-parametric ones. For instance, Schweitzer and Cachon (2000) use parametric statics while Bostian et al. (2008) use non-parametric statistics.

A majority of the aforementioned studies use aggregate analysis, meaning that they calculate averages over the individual decisions and draw inferences from this aggregate observation. This approach has given rise to serious concerns about the validity and applicability of these aggregate analysis results to individual decision making analyses. Lau et al. (2011) argue that drawing conclusions on individual behavior from aggregate data might be misleading. The authors show that even though pull-to-center is a plausible simplification of the newsvendor decision making on the aggregate level, it might not be appropriate at individual level. Revisiting the results of previous literature, Lau et al. show that individual observations do not support the existence of the pull-to-center effect. In fact, for instance, the aggregate results in Schultz et al. (2007) support demand chasing heuristic, even though only 16% of the individuals pursue that strategy. Although few, there are also papers that pursue a different path and use individual-based analysis such as Gavirneni and Isen (2010).

2.5.2 **Experimental Design**

Some researchers use manual recording of experiment data, while most others used a software or web applications. Currently, the most popular experimental software is the z-tree, whereas alternatives include MUMS (e.g., Şahin, 2011) or certain other web-based applications (e.g., Bostian et al. 2008).

2.6 Conclusion

Newsvendor model is the building block of many operations management models. Scientists have shown through experiments that individuals are systematically making suboptimal order decisions when faced with a newsvendor problem. Moreover, when humans interact with other humans via supply chain contracts, social preferences and effects of interaction start to impact the decision making processes. Behavioral Operations researchers have been trying to identify the factors behind these behaviors in order to embed them into analytical models. Here we presented a brief review of various studies from behavioral operations literature.

2.7 References

Becker-Peth, M., Katok E., Thonemann U. W. (2013). Designing Buyback Contracts for Irrational but Predictable Newsvendors. Management Science 59(8),1800-1816.

Benzion, U., Cohen Y., Peled R., Shavit T. (2008). Decision-making and the newsvendor problem: an experimental study, Journal of Operations 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., Katok E. (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback, Manufacturing and Service Operations Management, 10, 519-538.

Bolton, G. E., Ockenfels A., Thonemann U. W. (2012). Managers and Students as Newsvendors. Management Science, 58(12), 2225-2233.

Bostian, AJ. A, Holt C. A., Jain S., Ramdas, K. (2012). Is Transshipment a Behaviorally-Robust Risk-Pooling Strategy? Working Paper.

Bostian, AJ. A, Holt C. A., Smith A. M. (2008). Newsvendor "Pull-to-Center" Effect: Adaptive Learning in a Laboratory Experiment, Manufacturing and Service Operations Management, 10(4), 590-608.

Brown, A. O., Tang C. S. (2006). The Impact of Alternative Performance Measures on Single-Period Inventory Policy, Journal of Industrial and Management Optimization, 2, 297-318.

Caliskan-Demirag, O., Chen Y., Li J. (2010). Channel Coordination under Fairness Concerns and Nonlinear Demand, European Journal of Operations Research, 207, 1321-1326.

Chen, L., Kök A. G., Tong J. (2013). The Effect of Payment Schemes on Inventory Decisions: The Role of Mental Accounting. Management Science, 59(2), 436-451.

Corbett, C. J., Fransoo, J. C. (2007). Entrepreneurs and newsvendors: do small businesses follow the newsvendor logic when making inventory decisions? Working Paper.

Cui, T. H., Malluci P. (2016). Bounded Rationality and Fairness Ideals in Distribution Channels. Working Paper, to appear in Journal of Marketing Research.

Cui, Y., Chen L. G., Chen J., Gavirneni S., Wang, Qi. (2013). Chinese Perspectives on Newsvendor Bias: An Exploratory Note. Journal of Operations Management, 31, 93-97.

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

Davis, A. M. (2015). An Experimental Investigation of Pull Contracts. Production and Operations Management, 24(2), 325-340.

Davis, A. M., Katok, E., Santamar, N. (2014). Push, Pull, or Both? A Behavioral Study of How the Allocation of Inventory Risk Affects Channel Efficiency. Management Science, 60(11), 2666-2683.

De Vericourt, F. D., Bearden, J. N., Filipowicz, A. (2013). Sex, Risk and the Newsvendor. Journal of Operations Management, 31, 86-92.

Dimitriou, S., Robinson S., Kotiadis K. (2009). The Impact of Human Decision Makers' Individualities on the wholesale price contract's efficiency, Winter Simulation Conference, Austin, TX, USA, 2009, 2353-2364.

Erhun, F. (2011). Contract Complexity and Performance, Tutorials in Operations Research,

Feng, T., Keller L. R., Zheng X. (2011). Decision making in the newsvendor problem: A cross-national laboratory study, Omega, 39, 41-50.

Gavirneni, S., Isen A. M. (2010). Anatomy of a Newsvendor Decision: Observations from a Verbal Protocol Analysis, Production and Operations Management, 19, 453-462.

Gavirneni, S., Xia Y. (2009). Anchor Selection and Group Dynamics in Newsvendor Decisions-A note, Decision Analysis, 6, 87-97.

Haruvy, E., Katok E., Pavlov, V. (2016). Can Structured Bargaining Improve Channel Efficiency? Working Paper.

Ho, T., Zhang, J (2008). Designing Pricing Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter? Management Science, 54(4), 686-700.

Hoskin, R. E. (1983). Opportunity Cost and Behavior, Journal of Accounting and Research, 21, 78-95.

Hung, K., Ryu S. (2008). Changing risk preferences in supply chain inventory decisions, Production Planning and Control: The Management of Operations, 19, 770-780.

Hyndman, K., Kraiselburd S., Watson N. (2014). Coordination in Games with Strategic Complementarities: An Experiment on Fixed vs. Random Matching. Production and Operations Management, 23, 221–238.

Hyndman, K., Kraiselburd, S., Watson, N. (2013). Aligning Capacity Decisions in Supply Chains When Demand Forecasts Are Private Information: Theory and Experiment. Manufacturing & Service Operations Management, 15(1),102-117.

Kalkanci, B., Chen, K., Erhun, F., (2014). Complexity as a Contract Design Factor: A Human-to-Human Experimental Study. Production and Operations Management, 23(2), 269–284.

Kalkanci, B., K. Chen, F. Erhun. (2011). Contract Complexity and Performance under Asymmetric Demand Information: An Experimental Evaluation. Management Science, 57(4), 689-704.

Katok, E., Olsen T., Pavlov V. (2014). Wholesale Pricing under Mild and Privately Known Concerns for Fairness. Production and Operations Management, 23(2). 285-302

Katok, E., Wu, D. Y. (2009). Contracting in Supply Chains: A Laboratory Investigation, Management Science, 55, 1953-1968.

Keser, C., Paleologo, G. A. 2004. Experimental Investigation of Supplier-Retailer Contracts: The Wholesale Price Contract. Working paper, CIRANO Scientific Series

Kremer, M. (2008). Behavioral Perspectives on Risk Sharing in Supply Chains, (Doctoral Dissertation), University of Mannheim.

Kremer, M., Minner, S., Van Wassenhove, L. N. (2010). Do Random Errors Explain Newsvendor Behavior? Manufacturing and Service Operations Management, 12(4), 673-681.

Lau, N., Bearden, J. N., Hasija, S. (2014). Newsvendor Pull-to-Center Reconsidered. Decision Support Systems, 58, 68-73.

Lim, N., Ho T. (2007). Designing Price Contracts for Boundedly Rational Customers: Does the Number of Blocks Matter? Marketing Science, 26, 312-326.

Loch, C. H., Wu Y. (2008). Social Preferences and Supply Chain Performance: An Experimental Study, Management Science, 54(11), 1835-1849.

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

Moritz, B. (2010). Cognition and Heterogeneity in Supply Chain Planning: A Study of Inventory Decision Making, (Doctoral Dissertation), University of Minnesota.

Olivares, M., Terwiesch C., Cassorla L. (2009). Structural Estimation of the Newsvendor Model: An Application to Reserving Operating Room Time, Management Science, 54, 41-55.

Özalp, Ö. Zheng, Y. (2012). A Tale of Two Countries: Trust and Trustworthiness in China and the US. MSOM Conference Proceedings, Columbia University, New York, NY.

Özer, Ö., Zheng, Y., Chen, K. (2011). Trust in Forecast Information Sharing, Management Science, 57, 1111-1137.

Pavlov, V. (2009). Essays on Supply Chain Coordination, (Doctoral Dissertation), The Pennsylvania State University.

Pavlov, V., Katok E., Fairness and Coordination Failures in Supply Chain Contracts, Working Paper, (2011).

Pranoto, Y. (2005). Effects of Human Decision Bias on Supply Chain Performance, (Doctoral Dissertation), Georgia Institute of Technology.

Ren, Y., Croson, R. (2013). Overconfidence in Newsvendor Orders: An Experimental Study. Management Science, 59(11), 2502-2517.

Rudi, N., Drake, D. (2014). Observation Bias: The Impact of Demand Censoring on Newsvendor Level and Adjustment Behavior. Management Science, 60(5),1334-1345.

Şahin, N. (2011). Behavioral Factors affecting the Newsvendor Problem, Masters Thesis, Sabancı University.

Schultz, K. L., J. O. McClain, L. Robinson, L. J. Thomas. (2007). The Use of Framing in Inventory Decisions. Working paper, Johnson School Research Papers #02, Cornell University.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Shen, S., Wang H. (2011). Supplier's Pricing Decision When Facing Irrational Retailers, 8th International Conference on Service Systems and Service Management (ICSSSM), Tianjin, China, 2011, 1-5.

Su, X. (2008). Bounded Rationality in Newsvendor Models. Manufacturing and Service Operations Management, 10(4), 566-589.

Tokar, T., Aloysius J. A., Waller M. A. (2012). Supply Chain Inventory Replenishment: The Debiasing Effect of Declarative Knowledge, Decision Sciences, 43.

Wachtel, R. E., Dexter, F. (2010). Review of Behavioral Operations Experimental Studies of Newsvendor Problems for Operating Room Management, Economics, Education, and Policy, 110, 1698-1710.

Wang, C. X., Webster, S. (2007). Channel Coordination for a Supply Chain with a Risk-Neutral Manufacturer and a Loss-Averse Retailer, Decision Sciences, 38, 361-389.

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

Wu, D. Y. (2013). The Impact of Repeated Interactions on Supply Chain Contracts: A Laboratory Study. International Journal of Production Economics, 142, 3-15.

Wu, Y. (2006). Social Preferences in Coordination and Delegation Decisions- Models and Empirical Tests, (Doctoral Dissertation), INSEAD.

Zhang, Y., Donohue K., Cui H. T. (2016). Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing. Management Science, 62(6), 1734-1754.

Appendix-A: Classification of Behavioral Newsvendor Literature

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Schweitzer, Cachon, 2000	Selection of the order quantity in a newsvendor setting	Decision bias, anchoring behavior, prospect theory	Empirical evaluation of the newsvendor problem in a lab experiment.	Gain/loss domain of the game, profit margin	Order decisions systematically deviate from the optimal towards the center of the demand distribution. This result is robust with respect to loss/gain domain and profit margin.	Pull-to-center effect, anchoring and insufficient adjustments.
Bolton, Katok, 2008	Selection of the order quantity in a newsvendor setting	Bounded rationality, adaptive and forward looking learning,	Investigating the impact of extended and improved experience and feedback on the newsvendor's performance.	Number of order options, feedback on counterfactual payoffs, upfront information, standing orders	Experience and restricting changes in order quantity improves newsvendor performance. Info about forgone options does not affect while upfront info, though weakly, improves performance.	Anchoring and insufficient adjustments.
Bostian, Holt, Smith, 2008	Selection of the order quantity in a newsvendor setting	Experience weighted attraction learning, logit rationality, reinforcement, and memory	Empirical investigation of the pull-to-center effect in the newsvendor problem through adaptive learning models.	High, low and medium levels of profit margin, doubled payoffs.	Asymmetric existence of learning in the newsvendor performance. Ineffectiveness of doubled payoffs. Fitted model of learning, explains the data fairly well.	Pull-to-center effect asymmetry in the level of profit margin.
Benzion, Cohen, Peled, Shavit, 2008	Selection of the order quantity in a newsvendor setting	Bounded rationality	Investigation of the effect of the demand distribution and realizations on the newsvendor performance, and effect of learning.	Profit margin, demand distribution.	Order decisions are affected by the mean demand, optimal order quantity and the previous period's demand realization. Learning and convergence in the order quantities exist.	Effect of demand or supply surplus.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Lurie, Swaminathan, 2009	Selection of the order quantity in a newsvendor setting	Adaptive decision making	Investigation of the effect of feedback frequency on the newsvendor order decisions. Tracing the process of the newsvendor decision making.	Feedback in every round, every three or six rounds. Cost associated with changing orders.	Frequent feedback leads to demand chasing and deteriorations in the newsvendor performance and changes how the feedback is processed.	More frequent feedback is not necessarily better.
Su, 2008	Selection of the order quantity in a newsvendor setting	Bounded rationality.	Investigating the effectiveness of the quantal choice model built by the authors in estimating the newsvendor order decisions.	Profit margin.	Bounded rationality incorporated in the newsvendor model and theoretically proved effective in supply chain coordination	Pull-to-center explained by bounded rationality.
Kremer, 2008	Selection of the order quantity in a newsvendor setting	Decision heuristics, bounded rationality.	Empirical investigation of the behavioral perspectives on risk sharing in supply chains.	Several studies each with different set of experimental factors.	Framing in the newsvendor problem significantly impacts the ordering decisions.	-
Bolton, Ockenfels, Thonemann, 2012	Selection of the order quantity in a newsvendor setting	-	Investigation of the differences in ordering tendencies of different groups of subjects.	Subject pool, info on demand distribution and level of newsvendor training before the experiment.	Managers perform better when demand distribution is unknown, however when distribution is known there is no significant difference between students and managers.	The use of students in lab experiments is justified.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Becker-Peth, Katok, Thonemann, 2013.	Selection of the order quantity in a newsvendor setting.	Bounded rationality.	Empirical investigation of the rationality of the decision makers assumption and deriving response functions to model the relationship between the contract parameters and order decisions.	Various combinations of wholesale and buyback prices.	The response functions derived for expected orders, variance of orders and expected profit predict the actual behavior quite accurately.	-
Corbett, Fransoo, 2007.	Selection of the order quantity in a newsvendor setting	Prospect theory	Empirical investigation of newsvendor order decisions of entrepreneurs.	-	Entrepreneurs follow a newsvendor logic more for high margin products than bestselling ones. Risk aversion in profit leads to overstocking.	Effects of risk aversion on stocking levels.
Kremer, Minner, Wassenhove, 2010	Selection of the order quantity in a newsvendor setting	Random choice, decision heuristics	Empirical investigation of the hypothesis that newsvendors make random decisions.	Operations or neutral context of the problem, low or high profit condition.	The newsvendor errors are smaller in neutral context than in operations context. Decision heuristics are stronger in operations context.	Context dependent strategies of mean anchoring, demand chasing and regret.
Rudi, Drake, 2014	Selection of the order quantity in a newsvendor setting	Level, adjustment and observation biases.	Investigation of presence and impact of three decision biases in newsvendor decision making.	Censored or uncensored demand, profit margin.	Censored demand information leads to lower order quantities and stronger demand chasing. Learning is stronger in adjustment bias than in level bias.	Interaction between decision making biases and demand censoring.
De Vericourt, Bearden, Filipowicz, 2013	Selection of the order quantity in a newsvendor setting	Gender differences, risk appetite	Empirical investigation of gender differences in newsvendor decision making and factors behind these differences.	Gender of the subjects. Profit margin.	Gender differences in newsvendor ordering are partially driven by differences in risk appetite. Male subjects tend to be more risk taking and place higher orders in high profit margin case.	Gender differences are effected by profit margin.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Chen, Kök, Tong, 2013	Selection of the order quantity in a newsvendor setting	Prospective Accounting Theory	Investigating the impact of payment schemes on the ordering behavior of the newsvendor.	Payment schemes, payment times, profit margin	The newsvendors' behavior complies with a model that underweights order time payments, and framing of the scheme is sufficient to achieve this result.	Supply chain coordination may be possible with use of appropriate payment scheme by the supplier.
Schultz, McClain, Robinson, Thomas, 2007	Selection of the order quantity in a newsvendor setting	Framing, risk reversal	Empirical investigation of effects of framing on the newsvendor decision making.	High, low profit margin, negative or positive framing of the costs and profits.	Framing does not significantly affect the performance. No significant evidence of learning.	Anchoring and insufficient adjustments. Demand chasing. Individual observations deviate from aggregate ones.
Brown, Tang, 2006	Selection of the order quantity in a newsvendor setting	Different objective functions	Empirical investigation of alternative performance measures on the single period newsvendor performance.	Subject pool, students vs. buyers.	Less than optimal order quantities for both subject groups. Alternative performance measures explain less than optimal ordering behavior.	Profit level targeting.
Gavirneni, Isen, 2010	Selection of the order quantity in a newsvendor setting	-	Exploratory analysis of verbal protocols in a think alod newsvendor experiment to obtain deeper insights into the decision making process.	-	Majority of the subjects struggled with the abstractness of the business setting. A large number of subjects correctly identified cost of underage and overage but failed to convert to optimal order decision. Decision biases are significantly affected by the specific type of risk identified closer to the making of the decision.	Recency effect.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Wachtel, Dexter, 2010	Selecting the number of hours of OR time for which the staffing should be planned.	-	Review of existing behavioral operations management literature on newsvendor experiments which are mathematically equivalent to OR staff scheduling problems.	-	-	-
Feng, Keller, Zheng, 2011	Selection of the order quantity in a newsvendor setting	Context dependent preferences, adaptive learning, extremeness aversion, the doctrine of the mean.	Empirical investigation of the cross cultural differences in newsvendor decision making between American and Chinese subjects.	Subject pool. Number of order options, feedback on counterfactual payoffs, upfront information, standing orders	Pull-to-center effect is more prominent in Chinese than Americans. Chinese anchor at the mean and adjust less towards the mean. Flat maximums hypothesis works when coupled with extremeness aversion.	Cultural differences.
Cui, Chen, Chen, Gavirneni, 2013	Selection of the order quantity in a newsvendor setting	-	Empirical investigation of the cross cultural differences in newsvendor decision making between American and Chinese subjects.	Subject pool.	Chinese ask more questions before making the decision signaling that they are more worried about making the wrong decision, are more able to come up with a new decision quantity and more cognizant of the salvage values. But they are not able to identify creative, outside-the-box solutions.	Cultural differences.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Wang, Webster, 2009	Selection of the order quantity in a newsvendor setting	Loss aversion.	Theoretical investigation of the impact of loss aversion on the ordering decisions in a newsvendor setting.	-	If the cost of underage is not negligible loss averse newsvendor is expected to stock more than a risk neutral one.	
Ren, Croson, 2013	Selection of the order quantity in a newsvendor setting	Overconfidence bias.	Proposition of a theoretical behavioral model of overconfident newsvendors and testing it with the existing results in the literature.	-	Overconfident newsvendors are expected to place suboptimal orders and achieve suboptimal profit values.	Salvage values and price adjustments may be used to induce optimal decisions in an overconfident newsvendor.
Olivares, Terwiesch, Cassorla, 2009	Selection of the order quantity in a newsvendor setting	Cost estimation and forecasting biases in hospitals, overconfidence	Structural estimation of the unobserved cost parameters governing the newsvendor decision making and application to OR reservations.	-	Hospitals place more emphasis on the cost of idle time than on schedule overruns and long working hours of the staff.	-
Moritz, 2010	Selection of the order quantity in a newsvendor setting	Individual choice theory, cognitive psychology.	Empirical investigation of the correlation between cognitive reflection, dissonance and system neglect and newsvendor performance.	Several studies each with different set of experimental factors.	Individuals who score higher on the cognitive reflection test, perform better on the high profit condition, and they chase demand less. The correlation is not significant in low profit condition.	-

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Hoskin, 1983	Selection of the order quantity in a newsvendor setting	-	Empirical investigation of the impact of opportunity cost in newsvendor decision making.	Level of opportunity cost.	Subjects underweight the opportunity costs when it is low and overweight it when it is high. Ex-post outcome information improved ordering performance.	Impact of framing and impact of management reports on framing.
Gavirneni, Xia, 2009	Selection of the order quantity in a newsvendor setting	Decision biases and group decision making	Empirical investigation of the anchor selection behavior of the newsvendor decision makers and impact of group decision making on the performance.	Combination of different levels of possible anchor points.	Subjects select the anchor value that is closest to the optimal order quantity and the groups are less prone to errors than individual decision makers.	-
Bostian, Holt, Jain, Ramdas, 2012	Selection of the order quantity in a newsvendor setting	Experience weighted attraction, cognitive hierarchy, heterogeneity, adaptive learning, risk management	Empirical investigation of the efficiency of the transshipment in a newsvendor setting.	Transshipment costs	Order quantities increase as the transshipment costs increase. Transshipment generates higher profits even in the presence of behavioral decision biases.	-
Lau, Bearden, Hasija, 2014	Selection of the order quantity in a newsvendor setting	Pull-to-center	A review of pull to center effect and revisiting the results existing in the literature and reexamining them with an individual based analysis.	-	Drawing conclusions on individual behavior from aggregate data might be misleading, and pull to center might be inappropriate simplification of the newsvendor decision making.	Models based on pull to center effect need to be revisited.
Benzion, Cohen, Shavit, 2010	Selection of the order quantity in a newsvendor setting	Learning, decision biases.	Empirical investigation of the impact of the demand distribution being known or unknown on the newsvendor performance.	Information on demand distribution	Subjects knowing the demand distribution act differently from the subjects not knowing the distribution. However, knowing the distribution does not necessarily lead to higher profits or better performance. Supply surplus has a strong impact on the ordering decisions.	Waste aversion.

Appendix-B: Classification of behavioral supply chain contracting literature

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Keser, Paleologo, 2004	Selection of the wholesale price for the supplier, selection of the order quantity for the retailer	-	Empirical investigation of the wholesale price contracts.	-	Supplier charges lower than equilibrium price, retailer orders less than bet response. Yet, there is no significant efficiency loss compared to theory.	Fairness concerns in supplier's decisions. Cognitive limitations. Risk aversion in retailer's decisions. Quantity-price anchoring.
Katok, Wu, 2009	Selection of the contract parameters for the supplier and selection of the order quantity for the retailer.	Bounded rationality, loss aversion	Empirical investigation and comparison of the performance of three commonly used supply chain contracts.	Contract type, decision maker's role, experience, demand distribution and distribution frame.	Efficiency improvement in the channel profit due to coordinating contracts is experimentally less than the theory. RS and BB contracts are not equivalent.	Bounded rationality, cognitive limitations.
Ho, Zhang, 2008	Selection of the price for both supplier and the retailer. (Linear deterministic demand function)	Bounded rationality, effect of framing	Investigation of the impact of framing of the fixed fee on the retailer behavior.	Contract type and framing of the fixed fee.	Double marginalization is higher in linear price contract. QD generates higher channel efficiency than TPT. Minimal evidence of learning.	Reference dependent utilities. Loss aversion, contract complexity. No implication of fairness concerns.
Lim, Ho, 2007	Selection of the price for both supplier and the retailer. (Linear deterministic demand function)	Bounded rationality.	Investigation of the effects of number of blocks in a price contract on the retailer behavior.	Number of tariff blocks, actual managerial experience, stake size, group decision making.	Increasing number of blocks has different impacts on the channel efficiency and manufacturer's share than in theory.	Quantal response equilibrium model and effect of counterfactual payoffs to model bounded rationality.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Loch, Wu, 2008	Selection of the price for both supplier and the retailer. (Linear deterministic demand function)	Social preferences, relationship, status.	Experimental investigation of social preferences on economic decision making in supply chain transactions.	Relationship among contracting parties, winner of the round (status)	Higher efficiency in relationship, lower efficiency in status treatments.	Reciprocity, status seeking
Davis, 2010	Selection of the contract parameters for the retailer and selection of the production quantity for the supplier.	Risk aversion, anticipated regret	Empirical investigation of the pull type contracts and comparison of performance for wholesale, overstock allowance and service level agreement contracts.	Contract types, decision maker's role and decision maker's opponent. Standing orders, team, expected profit info.	OA and SLA are significantly better than WS, but they are not significantly different. Learning in early rounds	Winner's regret, loser's regret. Risk aversion, task complexity.
Kalkanci, Chen, Erhun, 2011	Selection of the contract type either price only or quantity discount with 2-3 blocks, and contract parameters.	Probabilistic choice bias, reinforcement bias, memory bias	Investigation of contract complexity and asymmetric information on supply chain performance and comparison with theory.	Contract type and number of blocks	QD does not necessarily improve supplier's profit, more equitable distribution of profits. Simpler contracts are sufficient with asymmetric info.	Experience-weighted attraction learning model.
Katok, Olsen, Pavlov, 2014	Selection of the wholesale price for the supplier, decision of accepting or rejecting the contract and selection of the order quantity for the retailer.	Fairness concerns as a private information.	Theoretical and empirical investigation of wholesale pricing when fairness concerns are private information.	Outside option for the retailer.	If fairness concerns are strong enough, coordination is possible. If fairness preferences are mild, retailers face disadvantageous inequality and the coordination fails.	Empirical distribution of preferences.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Pavlov, Katok, 2011	Selection of minimum order quantity and the wholesale price for the supplier, decision of accepting or rejecting the contract and the decision of order quantity for the retailer.	Inequality aversion, bounded rationality, incomplete information	Identification of possible causes of channel inefficiency	Retailer's profit in case of rejection, automated retailer, boundedly rational retailer, incomplete information.	Inequality aversion is the primary factor in retailer's decisions while incomplete information is in the supplier's decisions. Both are affected by bounded rationality but to a lesser extent.	Incomplete information on the retailer's aversion and bounded rationality parameters. Bargaining can help obtain this info.
Cui, Raju, Zhang, 2007	Selection of the wholesale price for the supplier and selection of the selling price for the retailer.	Fairness concerns	Investigation of fairness concerns in a two echelon supply chain with a wholesale contract.			
Wang, Webster, 2007	Selection of the contract parameters and amount of gain/loss information to share.	Gain/loss sharing, loss aversion	Theoretical investigation of the impact of gain/loss information sharing in a supply chain setting with buyback contract.	-	There exists a special class of distribution free gain/loss sharing buyback contracts that coordinate the chain and arbitrarily allocate the profit among the supplier and the retailer.	-
Özer, Zheng, Chen, 2011	Selection of the forecast information to share for the manufacturer, selection of production capacity for the supplier.	Trust, trustworthiness in information sharing	Investigation of trust on forecast information sharing in supply chain environment and how this affects channel coordination.	Capacity cost and market uncertainty	Trust significantly affects the outcome of cheap talk forecast communication and overall channel efficiency in a good way.	Trust, trustworthiness

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Özer, Zheng, 2012	Selection of the forecast information to share for the manufacturer, selection of production capacity for the supplier.	Trust, trustworthiness in information sharing, cultural differences	Investigation of trust and trustworthiness on country level.	Capacity cost and market uncertainty, country of the subject	Pull-to-Center revalidated and China exhibits lower trust and trustworthiness than US.	Trust, trustworthiness, cultural differences
Erhun, 2011	-	-	A tutorial of theoretical and behavioral aspects of contract design with an emphasis on the contract complexity.	-	-	-
Pranoto, 2005	Selection of the order quantity for the retailer and selection of contract parameters for the supplier.	Decision biases. Risk aversion	Empirical investigation of decision biases in newsvendor setting and impacts of initial wealth, salvage value, profit margin and interaction between supplier and retailer on the order decision making.	Initial wealth, profit margin, salvage value, newsvendor training, and the interaction between the supplier and the retailer.	Initial wealth significantly affects ordering decisions and risk appetite under different profit margins. The salvage value or theoretical newsvendor training do not significantly affect the performance. Anchoring at the previous period's order quantity.	-
Caliskan-Demirag, Chen, Li, 2010	-	-	A theoretical investigation of channel coordination under fairness concerns and nonlinear demand.	-	With exponential demand function, less stringent conditions are required to achieve coordination when only the retailer is fairness concerned.	-

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Kalkanci, Chen, Erhun, 2014	Selection of the contract type and parameters for the supplier and the selection of order quantity for the retailer.	Reinforcement bias and bounded rationality	Empirical investigation of the supplier's preferences in contract complexity and impact of contract complexity on the supply chain performance.	Contract type and number of blocks	Human to human interaction strengthen the supplier's preference for simpler contracts. The performance of price-only contract and quantity discount contracts are not significantly different. Supply chain profits are higher than expected by the theory.	Social preferences and equity concerns.
Shen, Wang, 2011	Selection of the order quantity for the retailer.	Bounded rationality, demand heuristics.	Empirical investigation of the ordering behaviors of retailers facing newsvendor problem and the theoretical analysis of optimal reaction of the suppliers to the irrational retailers.	Profit margin.	Pull-to-center effect in retailers' order decisions. A rational supplier should select a higher wholesale price for an irrational retailer. And in this case, the supplier may extract more profit than facing a rational retailer.	Mean anchor heuristic, minimizing ex-post inventory error
Hyndman, Kraiselburd, Watson, 2014	Simultaneous selection of capacity for both the retailer and the supplier.	First impressions bias, reputation building, long-run relationships	Empirical investigation of the impact of fixed and random matching between contracting parties, which also means the impact of duration of the relationship on the coordination performance.	Fixed or randomly matched pairs.	Capacity decisions are better aligned under fixed pairs while profits are more variable. Learning is slower for fixed pairs. Initial decisions affect the remaining decisions.	

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Hyndman, Kraiselburd, Watson, 2013	Simultaneous selection of capacity for both the retailer and the supplier.	Relationship, imperfect information, trust, signal processing	Theoretical and empirical investigation of the impact of lack of common knowledge on demand forecasts.	Noisiness of signals, number of rounds, prior on demand, level of common knowledge and number of signals.	Imperfect demand forecast information does not necessarily lead to worse performance for both firms, however capacities are more misaligned. Pre-play communication improves alignment and profits. Honesty is shown to be the best policy when sharing private information.	Social preferences and relationship.
Wu, 2013	Selection of the order quantity for the retailer and selection of contract parameters for the supplier.	Social preferences, relationship, bounded rationality.	Empirical investigation of the impact of repeated interactions on three different supply chain contracts, namely, wholesale, buyback and revenue sharing contracts.	Contract type, decision maker's role, experience, demand distribution and distribution frame.	When future opportunities to punish are available, social preferences for fairness and reciprocity is reinforced and reputation building behaviors are motivated to achieve long term economic benefits. As a result overall supply chain performance is enhanced. Moreover buyback contracts result in different behavior than revenue sharing contracts.	Fairness concerns, reciprocity, reputation building.
Cui, Malluci, 2016	Selection of the level of investment for both players, then selection of the wholesale price for the manufacturer and selection of the retail price for the retailer.	Bounded rationality, fairness concerns	Analytical and empirical investigation of firms' decision making in a two echelon supply chain where the firms decide on the level of investment first then the price.	Return rates and player roles.	There are significant fairness concerns in distribution channels and players are adverse to inequity and the aversion for disadvantageous inequity is greater than the advantageous inequity.	New fairness ideal is proposed.

Article	Behavioral Variables	Behavioral Theories	Research Question	Experimental Factors	Results	Behavioral Implication
Wu, 2006	Selection of wholesale price for the supplier and selection of the retail price for the retailer.	Bounded rationality, extended rationality, reciprocity, status, relationship, group identity.	Theoretical and empirical investigation of the impact of social preferences in coordination and delegation decisions.	Duration of interaction, manipulations of social preferences	Social preferences systematically affect economic decision making in business transactions. Social preferences can improve supply chain performance.	-
Zhang, Donohue, Cui, 2016	Selection of the contract type for an individual playing the supplier.	Loss aversion	Theoretical and empirical investigation of the impact of loss aversion on the supplier's contract selection.	Contract type and profit margin	Loss averse supplier chooses revenue sharing contracts in high profit margin environments and buyback contracts in low profit margin environments. This result complies with the theoretical prediction.	Heterogeneous time considerations in the decisions of individuals.
Pavlov, 2009	Study 1: Selection of the wholesale price and announced production cost for the supplier. Selection of the order quantity and the announced retail price for the retailer.	Information asymmetry, fairness concerns, bounded rationality, private information sharing	Theoretical and empirical investigation of supply chain coordination.	Several studies each with different set of experimental factors.	Fairness concerns play a significant role in supply chain coordination and failure of coordination.	-

Chapter 3

3. EXPERIMENTS ON SUPPLY CHAIN CONTRACTING: EFFECTS OF CONTRACT TYPE

Experiments on Supply Chain Contracting:
Effects of Contract Type

Abstract

We conduct decision experiments on a simple manufacturer-retailer supply chain where the retailer faces the newsvendor problem. Both firms in our experiments are represented by human subjects that engage in repeated interaction with each other. Different from the literature, we aim to understand the combined effects of individual decision biases and the strategic interaction between the subjects. We show the simple wholesale price contract to perform as good as the more sophisticated buyback and revenue sharing contracts. In addition, the buyback contract performed better than the mathematically-equivalent revenue sharing contract by certain measures. We separately explain the effects of the manufacturer's and the retailer's decisions on these outcomes. In particular, a detailed analysis on the retailer's underorder and overorder behavior is provided. We also develop a logistic regression model to predict the retailer's contract rejection, underorder and overorder tendencies as a function of contract fairness.

Keywords: Newsvendor experiments; Supply chain contracts; Fairness, Behavioral operations

3.1 Introduction

Achieving coordination between supply chain members in the face of uncertain consumer demand is one of the key challenges in supply chain management. Matching supply with demand, which is already a tough problem, becomes even more difficult when independent supply chain members are not coordinated. The resulting inefficiency can be mitigated by developing efficient supply chain contracts that align the incentives of members through proper profit and risk sharing. Due to the substance of the issue, a large number of researchers have worked on developing analytical supply chain contracting models, leading to an extensive research literature.

Aforementioned analytical models depend on a number of key assumptions. First, they assume rational decision makers who aim to maximize their individual expected profit. Second, they use game-theoretical approaches to model the strategic interaction between the decision makers. These assumptions have been challenged by a large number of experimental studies with human subjects, which display systematic deviations between model predictions and experiment data. In fact, analytic models' inability to predict real human behavior, even in controlled laboratory experiments, is likely to a reason behind the well-known gap between the research and practice of supply chain contracting. Experimental studies with human decision makers are particularly valuable for contracting research, because firms hardly ever share contractual data that might support field studies.

In this paper, we aim to understand how different contract types affect human decisions in a supply chain setting. To this end, we conduct laboratory experiments on a simple manufacturer-retailer supply chain where both firms are represented by human subjects. In the experiments, the same manufacturer-retailer pair interacts for 40 periods, capturing a long-run relationship. In each period, first the manufacturer offers a contract to the retailer. If the retailer accepts the contract, she determines her order quantity. Finally, random demand is realized, and the profits of both firms are calculated.

This research has three goals. First, we aim to compare the experimental performances of the three supply chain contracts: namely, the wholesale price, buyback and revenue sharing contracts. Different from literature, the effects of both individual decision biases and strategic considerations are present in our data, because both firms are represented by human subjects. Second, we aim to understand the separate roles of the manufacturer and the retailer's decisions in the outcome. To this end, we analyze the data at three different levels, In particular, we present a detailed analysis on the retailer's underorder and overorder behavior. Third, we want to characterize the role of fairness in retailer decisions. For this purpose, we develop a logistic regression model to predict the retailer's contract rejection, underorder and overorder tendencies.

The paper is organized as follows. Related literature is presented in Section 3.2. Next, in Section 3.3 we discuss the analytical background for our study and outline the experimental procedure. In Section 3.4, we compare each contract's experiment data with theoretical model predictions and among each other. In Section 3.5, we provide detailed studies about the effect of the manufacturer and retailer's decisions on contract performance. Section 3.6 provides further analysis on contract rejections and presents our fairness-model. We summarize the results and conclude in Section 3.7.

3.2 Related Literature

This work is related to the analytical and behavioral literature on the newsvendor model, and on supply chain coordination.

3.2.1 The Newsvendor Model

The analytical model we consider focuses on the retailer's newsvendor problem, and how the manufacturer can manipulate this problem through the choice of the contract parameters. The newsvendor model, thanks to its simple and elegant nature, has been used extensively as a building block in the development of more complicated stochastic inventory models. However, empirical studies indicate that managers do not necessarily follow the newsvendor solution in relevant problem settings (See Fisher and Raman 1996, Corbett and Fransoo 2007).

The newsvendor model, similar to any analytical model of human decision making, depends on the assumptions that human beings are rational decision makers that aim to maximize expected profit. However, a number of experimental studies involving human decision makers consistently identified biases (i.e., observed systematic deviations in decision making) between theoretical predictions and subject decisions. The first laboratory study of the newsvendor problem, conducted by Schweitzer and Cachon (2000), revealed that newsvendors (retailers) overstock for a low-profit-margin product, but understock for a high-profit-margin product. The authors show that this "pull to center effect" cannot be explained by risk preferences, prospect theory preferences, loss aversion, waste aversion, stockout aversion or an underestimation of opportunity costs. This pioneering work of Schweitzer and Cachon was followed by a large number of studies (For a summary see Bendoly et al. 2006, Loch and Wu 2008, Katok 2011, and Croson et al. 2013).

As discussed by Tversky and Kahneman (1981), people's decisions are affected by the "framing" of a problem, which refers to the way the problem is presented. Schultz et al. (2007) compare the newsvendor results under a positive frame that highlights profit, and a negative one that highlights costs. To their surprise, experiments indicate no significant difference. Ho and Zhang (2008) study the effect of framing of the fixed fee in supply chain contracts. Contrary to their theoretical equivalence, due to the differences in the framing of the fixed fee, the quantity discount frame and the two-part tariff frame are found to differ significantly.

Standard economic theory assumes decision makers to rationally choose the best response among alternatives. However, in practice, people are observed to act "boundedly rational" and make noisy decisions (see Gigerenzer et al. 2002). They may make calculation or recording errors due to limited cognitive ability, limited memory and attention span. When faced with complex decision situations, people often resort to decision heuristics as shortcuts. Su (2008) generalizes the newsvendor model to account for bounded rationality using a quantal response equilibrium (QRE) framework. According to this framework, people do not always make the best decision, however, better decisions have a higher probability of being made. Gavirneni and Isen (2010) record and analyze the thought process of newsvendor subjects in experiments. They find that most subjects correctly identified the overage and underage costs, but failed to convert this into the optimal order quantity. To sum up, as all the aforementioned results suggest, the newsvendor problem may not be as intuitive as what researcher think.

3.2.2 Supply Chain Contracting and Coordination

In a typical supply chain each firm aims to maximize its own profit, and this decentralized decision making process reduces total supply chain profits (Spengler 1950). Supply chain contracts can be used to align the incentives of the firms with that of the supply chain, leading to coordination. As summarized in Cachon (2003), researchers have studied different contract types to achieve coordination. Similar to our setting, these studies generally involve one manufacturer and one retailer, where the retailer faces the newsvendor problem. Pasternack (1985) is the first to show that a buyback contract can coordinate such a supply chain. Cachon and Lariviere (2005) show the buyback and revenue sharing contracts to be equivalent in the sense that they generate the same profit values for the same demand realization.

Our work is related to experimental papers on supply chain contracting where the retailer faces a newsvendor problem. This setting was first studied by Keser and Paleologo (2004). Considering only the wholesale price contract, these authors find the manufacturers to charge lower wholesale prices than predicted, and the retailers to understock (contrary to Schweitzer and Cachon's observation) relative to the newsvendor quantity. As a result, while the total profit in experiments is close to theoretical predicted values, it is more equitably shared between the two firms. The authors find support for a decision heuristic where the retailers anchor on some price-quantity combination in the first period, and adjust around this point based on the changes in the offered wholesale prices in the subsequent periods.

Our work is most related to Katok and Wu (2009), and Wu (2013). In fact, we use the same parameter setting as these two papers. Katok and Wu (2009) study the buyback and revenue sharing contracts, focusing on their coordination capabilities. These authors, however, conduct experiments where only the manufacturer or the retailer is human, whereas the other firm is computer simulated. Ignoring the strategic interaction between two human players, these authors only consider the effects of personal decision biases and bounded rationality. Katok and Wu find that the buyback and revenue sharing contracts' experimental performance are below theoretical predictions, and that these contracts fail to coordinate the supply chain. In addition, these two theoretically-equivalent contracts' experimental performance is shown to be different from each other. The initial performance difference between the two contracts is explained with loss aversion and framing, where the effect of framing is shown to vanish over periods.

Supply chain contracting involves relations between at least two independent decision makers (firms). This requires one to consider strategic/social factors in addition to individual decision biases. For example, rather than being purely self-interested, as assumed by standard economic theory, people may also care about the well-being of the others and the fairness of the relationship. In an analytical study, Cui et al. (2007) show that a simple wholesale price contract can achieve coordination when firms are concerned about fairness. Katok and Pavlov (2013) develop a model to explain contract rejections and the more equitable sharing of profits between the firms when the manufacturer has incomplete information regarding the retailer's preference for fairness. Loch and Wu (2008) study the effect of social considerations in a wholesale price contracting setting where the manufacturer and the retailer interact repeatedly. Relationship and status seeking considerations are shown to shift the subjects' equilibrium behavior significantly.

Wu (2013), using a supply chain scenario similar to Katok and Wu (2009), focuses on the strategic interaction between the two decision-makers, yet ignores (to a large extent) the individual decision biases and bounded rationality associated with the retailer's newsvendor problem. To this end, Wu conducts human-to-human experiments where the newsvendor retailer's choice set is constrained to only three choices: (1) Contract rejection, (2) Minimum possible order, (3) The newsvendor-optimal order quantity which is given to the subject. The author finds the long-run relationship arising from the repeated interaction between the human pairs to reinforce the social preferences for fairness and reciprocity, which improve supply chain performance. Similar to Katok and Wu (2009), Wu (2013) also observes the inequality between the theoretically-equivalent buyback and revenue sharing contracts.

Zhang et al. (2016) study the manufacturer's preference over the buyback and revenue sharing contracts in an experimental study where manufacturer subjects are matched with computerized retailers. The authors show that due to loss aversion, as the critical ratio increases, contract preferences of the manufacturers switch from buyback contract to revenue sharing contract. Additionally, contract parameter selection is shown to be affected by the framing of the contract.

In this paper we compare a single-parameter (wholesale price) contract with more complicated two-parameter (buyback and revenue sharing) contracts. Kalkanci et al. (2011)'s experiments demonstrate that simple contracts can outperform more sophisticated ones under certain conditions. In a related work, Kalkanci (2014) finds the manufacturer's preference for simpler contracts to increase with human-to-human interaction.

3.3 Analytical Background and Experimental Procedure

In this section, we first present our analytical model under the buyback (BB) contract scenario as an example. Next we solve this model, and outline the solutions under the two other contracts. Then, we explain our experimental procedure.

3.3.1 The Analytical Model

We consider a simple manufacturer-retailer supply chain as illustrated in Figure 1. The retailer purchases and stocks a certain product from the manufacturer prior to the selling season. The manufacturer produces the retailer's order by incurring a unit production cost of $c=\$3$. The consumer demand that the retailer faces during the sales season is discrete uniform distributed between 51 and 150.

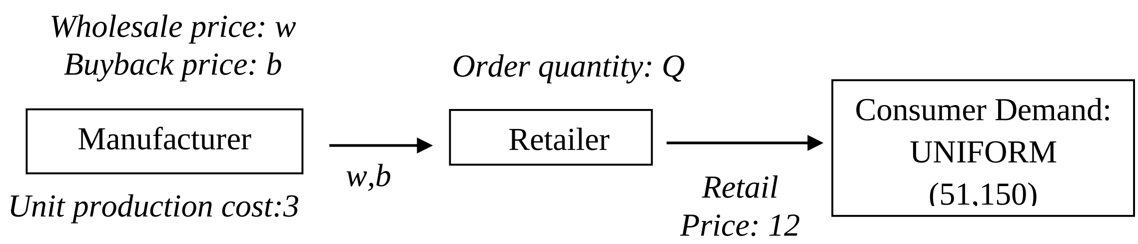


Figure 1: The supply chain under buyback contract

This is the same parameter setting used in Katok and Wu (2009) and Wu (2013). Under the buyback contract, the interaction between the manufacturer and the retailer proceeds in three stages:

Stage 1: The manufacturer offers the contract by determining two contract parameters:

- **Wholesale price, *w*:** This is the price that the retailer pays per unit she orders from the manufacturer. The wholesale price shall be an integer between the unit production cost *$3* and the retail price *$12*.
- **Buyback price, *b*:** This is the price that the manufacturer pays when buying the unsold products from the retailer at the end of the selling season. The buyback price shall be an integer not larger than the wholesale price.

Stage-2: In response to the manufacturer's contract offer, the retailer determines her order (stock) quantity. The retailer may "reject" the contract by setting zero order quantity, in which case both firms obtain zero profit. Otherwise, the manufacturer produces and delivers the order. The order quantity shall be an integer satisfying $51 \leq Q \leq 150$.

Stage-3: Random consumer demand is realized as "*D*". One of following happens:

- If demand is larger than or equal to the retailer's stock quantity, the retailer sells all *Q* units of her stock, and *D-Q* units of demand is lost. The retailer does not incur any loss of goodwill cost for unsatisfied demand.
- If demand is less than the stock quantity, the retailer sells as much as the demand. In this case *Q-D* products are left over. The manufacturer buys back these products by paying the buyback price of *b* per unit to the retailer.

3.3.2 The Theoretical Solution of the Model

We find the subgame-perfect equilibrium of the described three-stage game using backwards induction. The firms are assumed to be risk neutral, each aiming to maximize its own expected profit. The retailer faces the standard newsvendor problem in determining her order quantity at stage 2. Her expected profit function is *12*E[Sales] – w*Q + b*E[leftover]*, which is maximized at the order quantity $Q^*(w, \mathrm{b}) =$

$F^{-1}\left(\frac{c_u}{c_u+c_o}\right) = F^{-1}\left(\frac{p-w}{p-\mathrm{b}}\right) = 50 + \frac{12-\mathrm{w}}{12-\mathrm{b}}100$. The retailer is assumed to accept any contract that provides her with a nonnegative expected profit.

At stage 1, the manufacturer is assumed to foresee the retailer's order quantity choice *Q*(w,b)* in response to any contract *(w,b)* that he may offer. The manufacturer's expected profit function becomes *(w-3)*Q - b*E[leftovers]*. Because this function is not jointly concave in *w* and *b*, the manufacturer's optimal contract parameters are determined through a numerical procedure as *(w*=11, b*=10).*

In response to this contract, the retailer is predicted to order *Q=100* units. Expected profit values for the manufacturer, the retailer and the supply chain become $667.5, $75.5 and $753.0, yielding a contract efficiency value of 753/792=95.08%. Note that the realized profit values, which depend on the consumer demand realization, are likely to be different from these expected values. The manufacturer makes the first move, thus the subgame-perfect Nash equilibrium of the game corresponds to the manufacturer's optimal solution. Theory predicts this subgame-perfect equilibrium to be observed in every period when this game is repeated a finite number of periods.

So far, we explained the theoretical solution under the buyback (BB) contract *(w,b)*. The solution under the other two contract types are outlined below.

Wholesale price (WSP) contract *(w)*: This is a special case of the buyback contract with zero buyback price, i.e., *b=0*. Retailer's optimal order quantity is calculated as $Q^*(w) = F^{-1}\left(\frac{p-w}{p}\right) = 50 + \frac{12-\mathrm{w}}{12}100$.

Revenue sharing (RS) contract *(w, r)*: In addition to receiving the wholesale price of *w* per order, the manufacturer also receives a revenue share of *r* per unit sold to consumers in the market. Retailer's optimal order quantity is calculated as $Q^*(w,r) = F^{-1}\left(\frac{p-r-w}{p-r}\right) = 50 + \frac{12-r-w}{12-r}100$.

As the benchmark to compare contract performance, the order quantity that would maximize the total supply chain expected profit is calculated as $Q^* = F^{-1}\left((p-c)/p\right) = 125$.

The theoretical expected outcome under the three contracts are compared in Table 1. According to the theory, both buyback and revenue sharing contracts achieve an efficiency of 95%, whereas the wholesale price contract achieves only 74% efficiency. Contract efficiency comparison is directly reflected in the total profit comparison among the contracts as well. Thanks to the flexibility of using two contract parameters, the manufacturer's share of total profit is much higher under the BB and RS contracts than under the WSP contract. Having only one parameter, the WSP contract fails in this respect.

Table 1: Comparison of theoretical outcomes under different contracts

Contract Type	Total Profit	Contract Efficiency	Manufacturer's Profit	Retailer's Profit	w*	b*	r*	Q*
Wholesale Price	586.7	74.08%	469.0	117.7	10	-	-	67
Buyback	753	95.08%	677.5	75.5	11	10	-	100
Revenue Sharing	753	95.08%	677.5	75.5	1	-	10	100

3.3.3 Experimental Procedure

The experiment was announced to Sabanci University student body, and subjects were recruited using an online registration system. Prior to coming to the experimental session, the subjects were provided with experiment instructions. A sample of the experiment instructions can be found in the electronic companion. As detailed in Table 2, 12 experimental sessions were organized to which 132 subjects attended.

Table 2: Experimental design

Treatment	Acronym	# Sessions	# Subjects
Buyback	BB	3	44
Revenue sharing	RS	4	44*
Wholesale price	WSP	5	44

** One pair is found as an outlier and excluded from the analyses.*

The experiments were computer-based and were conducted at the CAFE (Center for Applied Finance Education) computer laboratory of Sabanci University. At the beginning of each session, the instructions were re-explained to ensure that they were clearly understood, and any remaining questions were answered. Before the actual experiment, the subjects were guided through three pilot periods. During the actual experiments the subjects were not allowed to communicate with each other. Each experimental session took around two hours. The subjects were motivated with real monetary payment proportional to subject's total profit in the experiment. Average payment was around $22. Each subject participated in only one treatment.

The experimental model was coded in HP MUMS software. The server computer in the lab randomly assigns the roles of manufacturer and retailer to the subjects, and determines the manufacturer-retailer pairs. Subject roles and pairings remain unchanged throughout the experiment, and subjects are aware of this.

Each experiment consists of 40 periods. In each period, the three-stage "game" between the manufacturer and retailer that was explained in Section 3.3.1 is played. The periods are independent of each other, and inventory is not carried from one period to the next. Subjects are provided with a decision support tool using which they can conduct what-if analysis to help their decisions. As shown in Figure 2, the tool illustrates the outcomes of the contract and quantity decisions under various demand realizations.

If my wholesale price is	9
and my buyback price is	3
and retailer's stock quantity is	120

If the total demand turns out to be	Retailer's sales quantity	Leftover products at the retailer	Units that I should buy back	My payoff	Retailer's payoff
51	51	69	69	513.0	-261.0
60	60	60	60	540.0	-180.0
70	70	50	50	570.0	-90.0
80	80	40	40	600.0	0.0
90	90	30	30	630.0	90.0
100	100	20	20	660.0	180.0
110	110	10	10	690.0	270.0
120	120	0	0	720.0	360.0
130	120	0	0	720.0	360.0
140	120	0	0	720.0	360.0
150	120	0	0	720.0	360.0

Figure 2: Sample manufacturer screen (Buyback contract)

3.4 Contract Comparisons with Theory and with Each Other

Here, we compare the results of the WSP, BB and RS treatments with the theoretical predictions, and among themselves. Recall that in our experimental setup a retailer can reject a contract offer, causing both firms to obtain zero profit for that period. Overall, 9.4% of the offered contracts were rejected. One particular retailer in the RS treatment rejected 21 out of 40 offered contracts. This retailer's pair is determined to be an outlier according to the Grubbs procedure (Grubbs 1969) at 1% significance level, and their data is not used in the analysis.

In this section we present the results over accepted contracts only. The unit of analysis in our hypothesis tests is a subject's average decision over 40 periods, ignoring the periods where the contract is rejected. Hence, each subject yields one independent data point. We have no prior assumptions on the distributions of the assessed variables; therefore we use non-parametric statistical tests (Siegel, 1956). These are the Wilcoxon Signed-Rank test for single-sample, and the Wilcoxon Rank-Sum test (the Mann-Whitney U test) for two-sample comparisons.

3.4.1 Comparisons with Theoretical Predictions

Hypothesis 1: *(For all three contracts) Subjects' decisions will be as predicted by theory.*

Table 3 compares averages of experimental data with theoretical predictions for all three contract treatments. We observe the manufacturer's contract decisions to be significantly different from the predicted values. Retailer's order quantity decision is also significantly different under the wholesale price and revenue sharing contracts. Under the WSP contract, the manufacturers set lower wholesale prices than theory, causing the retailers to order higher quantities. Thanks to these higher quantities, the WSP contract turns out to be more efficient than predicted, allowing a higher total supply chain profit. Under the buyback contract, the manufacturers offer much lower buyback prices than theory. However, the contract induces the theoretical predicted order quantity because the wholesale price is also lower than theory. Hence, the total chain profit and contract efficiency are close to the theoretical predictions. Note however that the retailer obtains a higher share of total profits than predicted. The observations for the revenue sharing contract are similar to the ones for the BB contract. Hence, overall, Hypothesis 1 is rejected.

Table 3: Comparison of the three contract type's experiment results with theory

	Wholesale Price (WSP)		**Buyback (BB)**		**Revenue Sharing (RS)**	
	Theory	**Experiment**	**Theory**	**Experiment**	**Theory**	**Experiment**
w	10	7.50 (0.59)***	11	8.69 (0.91)***	1	4.31 (1.65)***
b or r	-	-	10	4.95 (1.71)***	10	4.09 (1.76)***
Q	67	95.61 (13.56)***	100	99.77 (9.83)	100	95.74 (12.02)*
Retailer's Profit	117.68	304.82 (55.77)***	75.50	248.30 (81.86)***	75.50	267.5 (106.19)***
Mfg.'s Profit	469	415.31 (72.66)**	677.50	490.54 (96.28)***	677.50	452.85 (98.98)***
Total Chain Profit	586.68	720.13 (56.56)***	753	738.83 (42.22)	753	720.40 (55.20)**
Contract Efficiency	0.74	0.91 (0.07)***	0.95	0.93 (0.05)	0.95	0.91 (0.07)**
Mfg.'s Profit Share	0.80	0.63 (0.08)***	0.90	0.69 (0.12)***	0.90	0.65 (0.14)***

Standard deviations are reported in parenthesis. P values are from two-tailed Wilcoxon signed-rank test.
**** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1*

Under all three contracts, compared to theoretical predictions, the retailers obtain a higher, and the manufacturers obtain a lower share of the total chain profit. The retailer's ability to reject contracts is likely to be an important reason for this outcome. In some games the retailers even reject contracts that provide them with significant expected profit. This observation contradicts the theoretical assumption that the retailer accepts any contract that provides her a nonnegative profit.

3.4.2 Comparisons between the Contract Types

Here we compare the experimental performances of the contracts with each other. The theoretical solutions from Table 1 suggest the following hypotheses:

Hypothesis 2: *The efficiency of the buyback contract will be higher than that of the wholesale price contract.*

Hypothesis 3: *The efficiency of the revenue sharing contract will be higher than that of the wholesale price contract.*

Table 4 compares the WSP contract with the buyback and revenue sharing contracts separately. Reported p-values allow us to reject both Hypotheses 2 and 3. We observe the

WSP contract to be no less efficient than the other two contracts. The WSP contract leads to a lower critical fractile than the buyback and revenue sharing contracts, however, thanks to the higher Q/Q* ratio it achieves, the three contracts lead to similar order quantity choices. This causes the total supply chain profit and the efficiency of the WSP contract to be close to the other two contracts. In particular, the retailer obtains the highest profit and profit share under the WSP contract.

Table 4: Comparison of contracts with each other

Differences in Medians			
	(WSP – BB)	**(WSP – RS)**	**(BB – RS)**
Critical Fractile	-0.13***	-0.09*	0.05≠
Q	-2.62	0.53	3.15≠
Q/Q*	0.12*	0.13***	0.01≠
Retailer's Profit	76.29*	40.03	-36.26≠
Mfg.'s Profit	-86.3***	-26.81	59.49≠
Total Chain Profit	-8.82	-4.81	4.02≠
Contract Efficiency	-0.01	-0.01	0.01≠
Mfg.'s Profit Share	-0.09*	-0.08	0.01≠

P-values marked with ≠ are from a two tailed Mann Whitney U test, unmarked ones are from a one tailed test.
**** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1*

In its rightmost column, Table 4 also presents the comparison between the buyback and the revenue sharing contracts. We observe differences between these two theoretically-equivalent contracts, albeit not at a statistically significant level. The buyback contract causes higher order quantities, leading to higher total supply chain profit. This increase, however, benefits only the manufacturer, whereas the retailer's profit is lower than that under the revenue sharing contract.

3.5 Explaining Differences in Contract Performance

So far we compared contracts over their realized experimental outcome which depends on three factors: (1) Manufacturer's contract decision, (2) Retailer's quantity decision, (3) Realization of random demand, *D*. To understand the difference in contract performance, the effect of each factor needs to be isolated. To this end, we analyze experiment data at the following three levels we define:

- **Realized level (For a given contract, Q and D)**: The vales that are recorded as the experiment data.
- **Expected level (For a given contract and Q)**: The expected values with respect to probabilistic demand distribution.
- **Newsvendor-predicted level (for a given contract):** The expected values with respect to demand distribution, assuming that the retailer orders the newsvendor quantity Q*. These would be the expected values if a human manufacturer were playing against a computer that is programmed to order the newsvendor quantity. We often use the shorthand version "predicted" to denote the newsvendor-predicted values.

The plots in Figure 3 and Figure 4 illustrate the differences in manufacturer's cumulative profit on these three levels, under the three contracts. Each accepted contract decision constitutes one data point. Each plot in Figure 3 compares the three levels for a given contract type. The difference between the predicted and expected curves is due to retailer's deviation from the newsvendor-optimal quantity Q*, whereas the difference between the expected and realized curves is due to the random demand realization. As expected, the predicted profit distribution has the lowest variance and the realized distribution has the highest. For the WSP contract, the expected and realized curves overlap as the manufacturer's profit is not affected by demand realization under this contract.

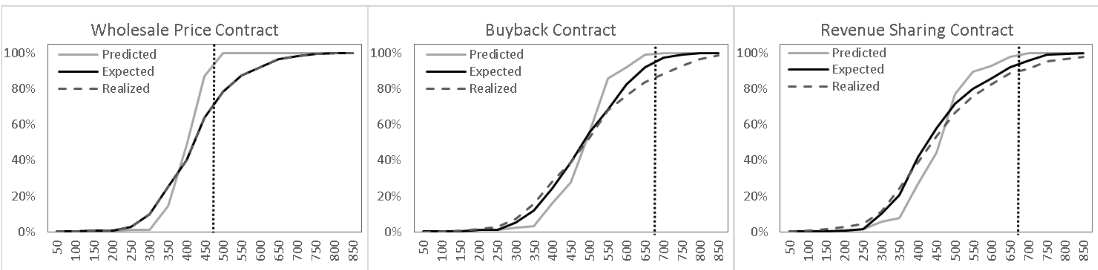


Figure 3: Inter-level comparison of the manufacturer's cumulative profit distribution

The dashed vertical line in each plot corresponds to the theoretical-optimal expected profit value. Note that around 70% of manufacturers' profit realizations are above the theoretical optimal under the WSP contract, whereas this figure is much lower for the BB and RS contracts.

Each plot in Figure 4 compares the three contracts for a given level. At each level, the buyback contract is observed to be the best, and the wholesale price contract is the weakest in terms of manufacturer profits. This is because contracts that allocate the manufacturer higher profit values are offered and accepted most frequently under the buyback contract, and least frequently under the wholesale price contract.

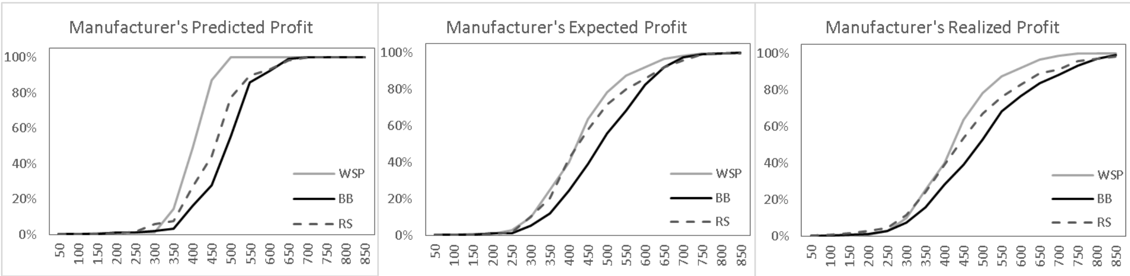


Figure 4: Inter-contract comparison of the manufacturer's cumulative profit distribution

These figures suggest that in order to understand the impact of the subjects' decisions, further analysis on newsvendor-predicted and expected levels is needed. In the following sections we separately focus on each level. Analysis is conducted on both subject averages and pooled data where each accepted contract generates a data point.

3.5.1 Effect of Manufacturer's Contract Parameter Decision

The impact of the manufacturer's decisions on contract performance can be isolated by focusing on the newsvendor-predicted level. At this level, differences are driven only by the differences in manufacturer's behavior, because the order quantity is assumed to be the newsvendor order quantity Q*.

The two-parameter design of the BB and RS contracts provide the manufacturers with a higher number of contract options and thus higher flexibility than the WSP contract. In fact, in our experiments, while only 10 contract options are available for WSP manufacturers, BB and RS manufacturers have 84 and 91 contract options (parameter combinations) respectively. As the flexibility of the contract increases, so does the manufacturer’s ability to extract more profit. Hence, we expect BB and RS manufacturers to have higher newsvendor-predicted profit than WSP manufactures.

Hypothesis 4: *BB and RS manufacturers will have higher predicted profit than WSP manufacturers.*

As shown in Table 5 through differences in medians, this hypothesis is supported under both subject averages and pooled data.

Table 5: Comparison of manufacturer's newsvendor-predicted profit

	Subject Averages			**Pooled Data**		
	Differences in Medians					
	(WSP – BB)	**(WSP – RS)**	**(BB – RS)**	**(WSP – BB)**	**(WSP – RS)**	**(BB – RS)**
Mfg Pred. Profit	-93.85***	-56.52***	57.18	-71.35***	-36***	35.35***

P values are from one-sided Mann Whitney U test. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1

Although they are mathematically equivalent, the framing of the BB and RS contracts are substantially different from each other. Under the BB contract, the manufacturer can obtain high newsvendor-predicted profit by offering a high wholesale price and by keeping a small difference between the wholesale price and the buyback price. However, to extract as much profit under the RS contract, the manufacturer needs to offer a low wholesale price, even lower than the unit production cost, and obtain most of his earnings through the revenue share from sold products. Even though the inventory risk to be undertaken by the manufacturer is mathematically equal, the RS contract setting is more prone to be affected by risk and loss aversion traits of manufacturer subjects. Thus we expect RS manufacturers to act more cautiously and have a lower newsvendor-predicted profit than BB manufacturers.

Hypothesis 5: *RS manufacturers will have lower predicted profit than BB manufacturers.*

The results of the comparison are presented again in Table 5. The sign of the difference between the BB and RS contracts conforms to our prediction. However, because the difference is significant only for the pooled data, Hypothesis 5 is only partially supported.

3.5.2 Effect of Retailer's Order Quantity Decision

Because the retailer is the second player in this game, her order quantity decision needs to be analyzed as a response to the manufacturer's contract parameter decision. The impact of the retailer's response can be studied through the relation between the newsvendor-predicted and the expected levels.

Table 6 compares the contracts in terms of the retailer's suboptimal underorder (defined as *max{Q*-Q,0}*) and overorder (defined as *max{Q-Q*,0}*) behavior. The table presents differences in medians between contract pairs.

Table 6: Comparison of retailer's deviation from Q*

Differences in Medians

	WSP-BB	**WSP-RS**	**BB-RS**
Underorder Quantity	-5.23*	-4.71*	$0.53^{\neq}$
Overorder Quantity	5.04*	8.17*	$3.13^{\neq}$
#Underorders	-1.5	-5*	$-3.5^{\neq}$
#Overorders	6.5*	8*	$1.5^{\neq}$

P-values marked with $^{\neq}$ are from a two tailed Mann Whitney U test, unmarked ones are from a one tailed test.
**** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1*

Under the WSP contract, retailers are observed to overorder significantly more frequently, and the average overorder quantity is significantly higher compared to the BB and RS contracts. While the underorder quantity of the WSP contract is significantly lower than both BB and RS contracts, underorders are significantly less frequent than only the RS contract. Between the BB and RS contracts, none of the comparisons is significant. However, on average, overorders are more frequent and larger, and underorders are less frequent under the BB contract.

Next, we study the manufacturers' ratio of expected profit to newsvendor-predicted profit, which is a measure of the manufacturer's benefit from the retailer's suboptimal ordering behavior. Under the WSP contract, an overorder always translates to the manufacturer earning higher expected profit than the newsvendor-predicted profit. Based on this fact, and the results of Table 6, we expect the WSP manufacturers to have a higher expected-to-predicted profit ratio than BB and RS manufacturers.

Hypothesis 6: *WSP manufacturers will have higher ratio of expected profit to predicted profit than BB and RS manufacturers.*

Next we compare the BB and RS contracts. From Table 6, the average overorder quantity and number of overorder occurrences are both higher under the BB contract than under the RS contract. Although the average underorder quantity under the BB contract is higher, the number of underorder occurrences is lower. For the majority of the contract parameter combinations, an overorder causes both BB and RS manufacturers to earn higher expected profit than the predicted value. Hence we conjecture that BB manufacturers will have higher expected-to-predicted profit ratio than RS manufacturers. That is, the retailers' suboptimal ordering behavior will work, on average, more in favor of BB than RS manufacturers.

Hypothesis 7: *BB manufacturers will have a higher ratio of expected profit to predicted profit than RS manufacturers.*

Table 7 presents the median differences in expected-to-predicted profit ratios between contract pairs for both subject averages and pooled data. Under the WSP contract, the ratios are significantly higher than that of BB and RS contracts, which supports Hypothesis 6. Between the BB and RS contracts, the sign of the difference conforms to our prediction, however it is only significant for the pooled data. Hence Hypothesis 7 is only partially supported.

Table 7: Comparison of the manufacturer's ratio of expected profit to predicted profit

	Subject Averages			**Pooled Data**		
	Differences in Medians					
	(WSP – BB)	**(WSP – RS)**	**(BB – RS)**	**(WSP – BB)**	**(WSP – RS)**	**(BB – RS)**
Manufacturer's Expected/Predicted Profit	12.28%*	14.21%**	1.93%	8.16%***	8.70%***	0.54%**

P values are from one-sided Mann Whitney U test. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1

Figure 5 compares the manufacturer's newsvendor-predicted (x-axis) and expected profit (y-axis) values for each contract type. The trend line, which regresses the expected profit on the newsvendor-predicted profit, and the regression equations are also shown. The comparison of regression equation coefficients support both Hypotheses 6 and 7.

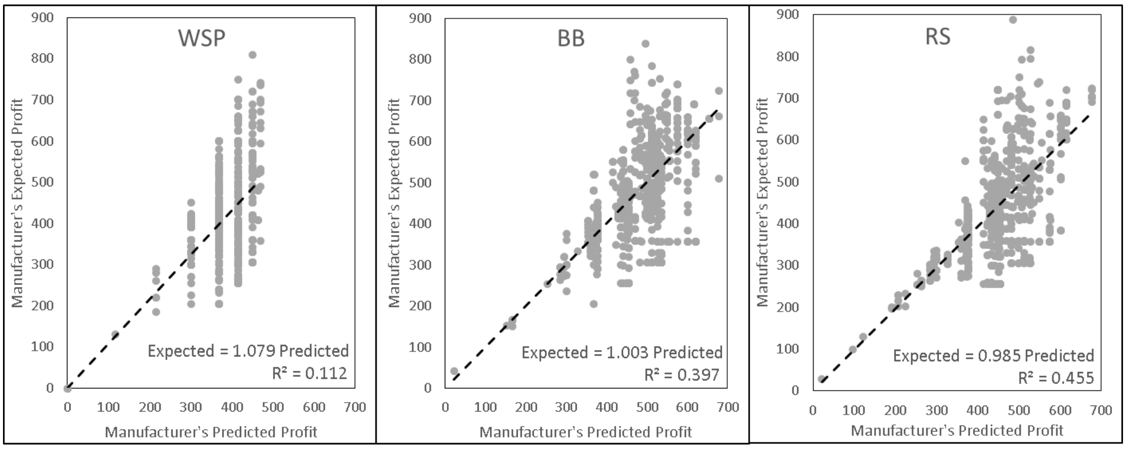


Figure 5: Comparison of manufacturer's expected and newsvendor-predicted profit values

3.5.3 Effect of Retailer's Under/Overorder Behavior on Manufacturer's Realized Profit

In the previous section, we discussed the impact of retailer's deviation from the newsvendor quantity on the manufacturer's expected profit. Here, we analyze the effect on manufacturer's realized profit.

Under the WSP contract, the manufacturer's profit is independent of the demand realization. Thus, the retailer's overorder is always beneficial, and underorder is always harmful for the manufacturer's profit (except when the wholesale price is equal to the unit production cost 3). Under the BB and RS contracts, however, the manufacturer's profit is also dependent on the demand realization. With these so-called risk-sharing contracts, the two firms share inventory risk in proportions determined by the contract parameters.

We first clarify what is meant by a gain/loss <u>of manufacturer</u> due to under/overordering of the retailer through the following definitions:

- **Counterfactual profit:** Denotes the profit that the manufacturer would make if the retailer ordered the newsvendor-predicted quantity Q* under the same demand realization.

- **Gain/Loss due to under/overorder**: Denotes the difference between the realized and counterfactual profit of the manufacturer (that is, realized minus counterfactual profit).

The following analysis is presented in two parts: First we discuss the number of under/overorder occurrences, then we discuss their magnitudes.

Table 8 presents the results by the number of under/overorder occurrences, when all subject pairs' data is pooled for each contract type. The first column present the percentage of underorder occurrences over all orders, whereas the second column, for example, shows the percentage of gain occurrences over all underorder occurrences. The number of relevant data points are provided in parenthesis.

Table 8: Underorder and overorder occurrences (Pooled data)

	# Underorders	**# Gain due to underorder**	**# Loss due to underorder**	**# Overorders**	**# Gain due to overorder**	**# Loss due to overorder**
WSP	36% (288/805)	0%	100% (288/288)	62% (500/805)	99% (497/500)	0%
BB	42% (324/780)	14% (46/324)	79% (256/324)	53% (411/780)	80% (327/411)	11% (47/411)
RS	47% (358/769)	6% (22/358)	90% (322/358)	49% (375/769)	79% (296/375)	12% (46/375)

Under the WSP contract, as expected, overorders were almost always beneficial (except when $w=3$), and underorders were harmful for the manufacturer. For the BB and RS contracts, the manufacturer gained from around 80% of the overorder occurrences. Only 11-12% of the overorders caused a loss compared to what the manufacturer would have made had the retailer ordered Q^*. The effect of underorders on the BB and RS contracts, however, are different. The number of loss occurrences due to underorder are higher under the RS contract. Also, overorders are more frequent under the BB contract, whereas underorders are more frequent under the RS contract (though the differences are not large). A possible explanation for this difference might be found in the structural framing differences of the two contracts. The security against inventory risk provided by the buyback clause might be encouraging the retailers to overorder. Meanwhile, the revenue sharing clause might be discouraging the retailers from placing higher order quantities with the reasoning that most of the high order's benefits will accrue to the manufacturer. Also note that the WSP contract leads to a much higher percentage of overorder occurrence compared to the other two contracts.

Next, we compare the magnitudes of underorder and overorders, and the gain/loss they cause to the manufacturer. The analysis is conducted for each manufacturer subject separately. The averages over all manufacturers are presented in Table 9. The gain-related definitions for the table are given below. Loss-related definitions are similar.

- **#Gain**: Number of gain occurrences for the manufacturer.
- **Average Gain**: Average gain of the manufacturer over all accepted contracts, calculated for under and overorder cases separately.
- **Overall Gain**: Total gain of the manufacturer due to retailer's under and overordering.

Table 9: Underorder and overorder magnitudes: Contract comparisons and p-values

	Experiment Averages						**P-values**					
	WSP		**BB**		**RS**		**WSP vs BB**		**WSP vs RS**		**BB vs RS**	
	Under order	**Over order**	**Under order**	**Over order**	**Under order**	**Over order**	**U**	**O**	**U**	**O**	**U**	**O**
#Gain	0.00	22.59	2.09	14.86	1.05	14.10		**.03**		**.01**	.36	.82
Average Gain	0.00	59.31	2.46	33.78	1.05	31.53		**.02**		**.01**	.40	.77
Total Gain / Total Profit	0.00%	13.26%	0.53%	6.14%	0.24%	5.97%		**.00**		**.00**	.30	.85
#Loss	13.09	0.00	11.64	2.14	15.33	2.19	.64		.20		**.07**	.46
Average Loss	-28.76	0.00	-34.55	-1.51	-39.05	-1.18	.42		**.09**		.50	.38
Total Loss / Total Profit	-7.97%	0.00%	-8.14%	-0.28%	-8.21%	-0.25%	.83		.34		.79	.69
Overall Gain / Total Profit	13.26%		6.67%		6.21%		**.00**		**.00**		.57	
Overall Loss / Total Profit	-7.97%		-8.42%		-8.46%		.72		.34		.96	

We observe that WSP manufacturers simultaneously enjoy the highest gain from overorders and the lowest loss from underorders. In fact, they make a net gain from the suboptimal ordering behavior of the retailer. The BB and RS manufacturers, on the other hand, lose more than what they gain due to under/overorders. These observations are supported with two-sided Mann Whitney U test p-values provided at the right half of the table. While there are some differences in the under/overordering behavior under the BB and RS contracts, the only significant difference is found in the average number of loss occurrences.

Our detailed analysis on the under/overordering behavior of the retailers sheds light on the difference between experiment data and theoretical predictions for the contracts. In the following section, in addition to an analysis over rejected contracts, we present a model to explain the factors affecting retailer's suboptimal ordering decisions.

3.6 Contract Rejections and Fairness Considerations

Up to this point, we only studied the data of accepted contracts. However, contract rejection is also an important component of the retailer's decision making process. In this section, we first present an overall analysis of rejected contracts. Then we present a fairness model to predict the retailer's contract rejection and ordering behavior.

3.6.1 Rejected Contracts Analysis

Being the party to offer the contract, the manufacturer has the first-mover advantage in our scenario. This advantage, however, is to some extent counterbalanced by the retailer's power to reject the contract. In this regard, the interaction within our subject pairs resembles the well-known Ultimatum Game (Kagel and Roth 1995, Wu 2013).

Overall, retailer subjects rejected 245 out of total 2600 offered contracts in the experiments. As summarized in Table 10, a large majority of retailer subjects have rejected at least one contract offer. The table also presents the average newsvendor-predicted retailer profit in rejected contracts. When compared with the average newsvendor-predicted retailer profit in accepted contracts (which is 304.8 for WSP, 248.3 for BB and 274.1 in RS contracts), we observe the retailers to forego a significant profit opportunity through contract rejections. This is especially the case for the buyback contract. In fact, the number of contract rejections per retailer under the buyback contract (median: 4) is found to be higher than both WSP (median: 2) and RS (median: 2) contracts (p-values 0.11 and 0.10).

Table 10: Comparison of rejected contracts

Contract	# Rejections	# Retailers with at least one rejection	Retailer's average predicted profit in rejected contracts	Manufacturer' average predicted profit in rejected contracts
WSP	75	20/22	200.76	438.08
BB	99	20/22	173.76	518.13
RS	71	16/21	161.24	504.86

Figure 6 presents the cumulative distribution of the newsvendor-predicted retailer profit share for the accepted and rejected contracts under each contract type. We observe the WSP manufacturers not to offer contracts that have very low retailer profit share. In fact, WSP contracts that do not offer at least around 30% of the total supply chain profit to the retailer are rejected. On the other hand, BB and RS manufacturers do offer contracts that suggest very low predicted profit share to the retailer, and some of these contracts were accepted by retailers. Manufacturers might be using the two parameters available in the BB and RS contracts to craft better-looking contracts that actually serve to their own interests.

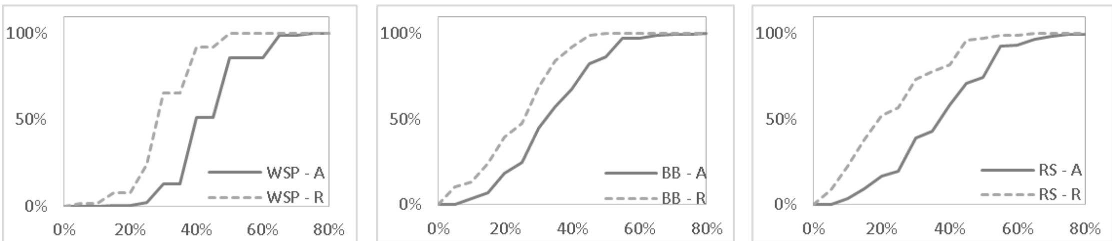


Figure 6: Retailer's cumulative newsvendor-predicted profit share (**A**: Accepted, **R**: Rejected)

Next, we present a more comprehensive analysis to understand the factors behind the retailers' contract rejection, as well as their under and overordering behavior.

3.6.2 Retailer's Response to Inequality

Here we present a regression analysis to understand the role of fairness concerns in retailer's ordering behavior. We measure the fairness of a contract through the share of the total newsvendor-predicted profit that the contract offers to the retailer. Because of the separation between the order quantity that represents a contract rejection *(Q=0)* and the possible quantities for accepted contracts *(Q between 51 and 150*), a linear regression model would not be suitable. Thus we use the following random effects ordered logistic regression model, which is explained over the BB contract as an example:

$$Y_{it} = \alpha + \beta_1 I_{it} + \beta_2 t + \beta_3 D_{t-1} + \beta_4 \Delta w_{it} + \beta_5 \Delta b_{it} + \eta_i + \varepsilon_{it}$$

The variables are described in Table 11. Note that the dependent variable takes on a value out of the response category {1,2,3,4} depending on the retailer's order quantity's relative position with respect to the newsvendor-optimal Q*{contract rejection (Y=1), underorder (Y=2), near optimal (Y=3), overorder (Y=4)}.

The first explanatory variable, inequality, is related to the fairness of the contract (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Wu, 2013). Fairness here is defined as the equal split of newsvendor-predicted profits between the retailer and the manufacturer. Inequality then is defined as the difference between the retailer's predicted profit share and a fair split. Demand realization in the previous period is another possible factor affecting retailer's quantity decisions. This effect might manifest itself as demand chasing behavior or a bias such as the gambler's fallacy. The repetitive and interactive component of the retailer-manufacturer relationship in the data is partially captured through the changes in the contract parameters from the previous period.

Table 11: Definition of variables for the regression equation

Variable	Definition
Indices	
i	Retailer ID
t	Period
Dependent Variable	
Y_{it}	1, if $Q=0$ (contract rejection) 2, if $Q<0.95Q^*$ (underorder) 3, if $0.95Q^* \leq Q \leq 1.05Q^*$ (near-optimal) 4, if $Q>1.05Q^*$ (overorder)
Independent Variables	
I_{it}	Inequality: (*Retailer's Newsvendor-Predicted Profit Share minus 50%)*
t	Period index
D_{t-1}	Demand realization in period $t-1$
Δw_t	Change in wholesale price, w_t-w_{t-1}
Δb_t	Change in buyback price, b_t-b_{t-1}
Error terms	

η_i	Individual-specific error for retailers
ε_{it}	Common error term with standard logistic distribution

Regression results are summarized in Table 12. The number of observations for each category is given at the top, and the coefficients of explanatory variables are given at the bottom of the table. A positive coefficient implies that an increase in the value of the variable will increase the probability of order behavior moving away from rejection (Y=1) towards overorder (Y=4).

For all contract types, the coefficient of inequality is found to be significant and positive. Thus, the retailer's chances of rejection and underorder increases as the retailer's predicted profit share decreases below 50%. Equivalently, as the retailer's predicted profit share increases over 50%, retailer's chances of placing a larger order quantity increases. These results are in line with our expectation that the retailer's order quantity responds to the perceived fairness of the contract offer.

Table 12: Random effects ordered logistic regression results for the retailer's decisions

Variables	**Dependent Category**	**WSP**	**BB**	**RS**
# Observations	Y=1	72	96	69
	Y=2	223	284	300
	Y=3	93	118	142
	Y=4	470	360	308
Log likelihood		-758.02	-934.77	-891.74
Wald's Chi^2		41.71	26.71	125.56
Intercept	Y=1	-4.57 (0.667)***	-4.01 (0.532)***	-3.809 (0.616)***
	Y=2	-1.881 (0.669)**	-1.59 (0.557)**	-0.965 (0.523)*
	Y=3	-1.143 (0.7)	-0.804 (0.59)	-0.018 (0.539)
Inequality		16.276 (2.914)***	6.182 (2.189)**	6.504 (2.59)*
Period		0.007 (0.01)	0 (0.011)	0.023 (0.013)*
Previous Demand		0.001 (0.003)	-0.001 (0.002)	-0.002 (0.002)
Δw		0.064 (0.101)	-0.146 (0.122)	-0.314 (0.152)*
Δb or Δr			0.003 (0.077)	-0.224 (0.105)*

Standard deviations are reported in parenthesis.

*** p-value < 0.001, ** p-value < 0.01, * p-value < 0.1

The coefficient of period number is significant only for the RS treatment. Period number has a small and positive coefficient for all three treatments, which implies the chances of rejection and underorders to slowly reduce over time. The impact of the most recent demand realization is not significant in any of the treatments and its sign is not consistent.

For the BB and RS contracts, the signs of coefficients for the changes in the contract parameters are in line with our expectations. As the wholesale price increases, retailer's cost of underage decreases, thus we expect the retailer to be more inclined to reject the contract or to place a small order. As the buyback price increases, retailer's cost of overage decreases, thus we expect the more secure-feeling retailer to place a higher order. The effect of an increase in the revenue share is similar to that of an increase in the wholesale price w. Changes in the wholesale price are significant for the BB and RS contracts. For the RS contract, changes in r is also significant. Because b is relevant to cost of overage and r is relevant to cost of underage, retailers in general may be more sensitive to the changes in the cost of underage than the cost of overage.

In this ordered logistic regression model, the impact of explanatory variables is assumed to be same over all categories. The intercept values given in the table indicate the cutoff levels between response categories of the dependent variable. These intercept values help determine the probability of each category.

Table 13 presents the results of a sample computation for the regression results. For this example, assume that in period 20 the manufacturer has increased the wholesale price by 1 relative to period 19, the other contract parameter (if any) is unchanged, and the demand realization in period 19 was 100. Given the different values of the inequality measure, the table lists the probabilities of each response category under the three contracts. Unless the contract is highly unequal, the model predicts WSP retailers to be more likely to place an overorder than BB and RS retailers are. Also, BB retailers are predicted to be more likely to place overorders than RS retailers. These predictions are in parallel with Hypotheses 6 and 7.

Table 13: Sample computation for regression results

	Inequality = 0.2			Inequality = 0.1			Inequality = -0.1			Inequality = -0.3		
	WSP	**BB**	**RS**	**WSP**	**BB**	**RS**	**WSP**	**BB**	**RS**	**WSP**	**BB**	**RS**
Rejection	0.0%	0.7%	0.6%	0.1%	1.2%	1.2%	3.7%	4.1%	4.3%	50.2%	12.9%	14.1%
Underorder	0.4%	6.4%	9.2%	2.0%	11.1%	16.1%	32.7%	28.5%	39.2%	43.5%	49.6%	59.7%
Near-optimal	0.5%	7.2%	12.1%	2.3%	11.3%	17.8%	18.1%	18.9%	23.0%	3.2%	16.0%	14.1%
Overorder	99.1%	85.7%	78.0%	95.6%	76.4%	64.9%	45.5%	48.5%	33.5%	3.1%	21.5%	12.1%

3.7 Conclusions

In this paper, we study the experimental performances of the wholesale price, buyback and revenue sharing contracts. Our primary findings are in line with the findings of the existing literature. Specifically, the simple wholesale price contract can be as efficient as the other two theoretically more efficient contracts. The buyback contract performs better than the theoretically-equivalent revenue sharing contract. Under all three contracts, total supply chain profit is more equitably shared between the firms than predicted by theory. In addition to these findings, our setup allows decoupling the effects of the manufacturer's and retailer's decisions on contract performance. This is because both firms in our experiments are represented by human subjects (unlike Katok and Wu 2009), and the retailer subjects' order quantity decision is not restricted to the optimal quantity or the minimum demand (unlike Wu 2013).

To isolate the effect of manufacturer's contract decisions, we conduct the "newsvendor-predicted" level analysis. At this level, for a given contact, the order quantity is assumed to be the newsvendor optimal quantity. We find the buyback and revenue sharing contract manufacturers to offer less advantageous contracts to retailers than the wholesale price contract manufacturers. The comparison between the buyback and revenue sharing manufacturers find the former to offer less favorable contracts, and undertake higher inventory risk than the latter. We conjecture that the framing of the prices in the two contracts, and risk/loss aversion biases of the manufacturers can be the drivers behind the observed performance difference between these two contracts. In fact, an investigation on the effects of these drivers offers a promising research direction.

The effect of retailer's decisions are investigated through retailers' underorder and overorder behavior with respect to the newsvendor order quantity. We characterize how the retailer's ordering behavior affect the manufacturer's "expected" (with respect to demand distribution) profit level, and compare this with the manufacturer's newsvendor-predicted profit level. Wholesale price contract manufacturers obtained higher expected profit than their newsvendor-predicted profit due to retailers' tendency to overorder. Similarly, we find the buyback manufacturers to enjoy higher expected profits than revenue sharing manufacturers because of the retailers' inclination to place more frequent and larger overorders under the buyback contract. The difference, again, may be due to the framing of the contracts, this time from the retailers' point of view. Next, we characterize the impact of retailer's suboptimal behavior on manufacturer's "realized" profit. In particular, we find the retailers' suboptimal decisions to benefit the wholesale price contract manufacturers.

Finally, we develop a random effects ordered logit model to predict the retailer's contract rejection, underorder and overorder tendencies. We demonstrate the fairness of the contract, which is measured with the retailer's predicted profit share, to have a major effect on the retailer's decisions. The tendency to overorder increases if the contract is fair, whereas the tendency to underorder or contract rejection increases otherwise. Furthermore, in agreement with our experimental findings, the model predicts that the retailers' tendency to overorder will be the highest under the wholesale price contract, and lowest under the revenue sharing contract.

Acknowledgements

This research was funded by TÜBİTAK Grant 111K454. The authors would like to thank Elena Katok and Diana Wu for sharing their demand stream.

3.8 References

Bendoly, E., K. Donohue, K. L. Schultz. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of Operations Management, 24, 737-752.

Bolton, G. E., Ockenfels. A. (2000). ERC: a theory of equity, reciprocity and competition. The American Economic Review, 90(1), 166-193.

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

Corbett, C. J., Fransoo, J. C. (2007). Entrepreneurs and newsvendors: do small businesses follow the newsvendor logic when making inventory decisions? Working Paper.

Croson, D. C., Croson, R., Ren, Y. (2013). How to Manage an Overconfident Newsvendor. Management Science, 59(11), 2502-2517.

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

Fehr, E., Schmitdt, K. M. (1999). A Theory of Fairness, Competition and Cooperation. Quarterly Journal of Economics, 114(3).

Fisher, M., Raman, A. (1996). Reducing the cost of demand uncertainty through accurate response to early sales. Special Issue on New Directions in Operations Management, 44(1), 87-99.

Gavirneni, S., Isen A. M. (2010). Anatomy of a Newsvendor Decision: Observations from a Verbal Protocol Analysis, Production and Operations Management, 19, 453-462.

Gigerenzer, G., Selten, R. (2002). Bounded Rationality. Cambridge: MIT Press

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. Technometrics 11(1), 1-21.

Ho, T., Zhang, J. (2008). Designing Pricing Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter? Management Science, 54(4), 686-700.

Kagel, J.H., Roth, A. E. (1995). Introduction to experimental economics. ed: Kagel. J. H. and Roth, A. E., The Handbook of Experimental Economics, Princeton University Press, Princeton, 3-109.

Kalkanci, B., Chen, K., Erhun, F., (2014). Complexity as a Contract Design Factor: A Human-to-Human Experimental Study. Production and Operations Management, 23(2), 269–284.

Kalkanci, B., Chen, K., Erhun, F. (2011). Contract Complexity and Performance under Asymmetric Demand Information: An Experimental Evaluation. Management Science, 57(4), 689-704.

Katok, E., Wu, D. Y. (2009). Contracting in Supply Chains: A Laboratory Investigation. Management Science, 55(12), 1953-1968.

Katok, E., Pavlov, V. (2013). Fairness in Supply Chain Contracts: A Laboratory Study. Journal of Operations Management, 31, 129–137.

Keser, C., Paleologo, G. A. 2004. Experimental Investigation of Supplier-Retailer Contracts: The Wholesale Price Contract. Working paper, CIRANO Scientific Series

Lau, N., Bearden, J. N., Hasija, S. (2014). Newsvendor Pull-to-Center Reconsidered. Decision Support Systems, 58, 68-73.

Loch, C. H., Wu, Y. (2008). Social Preferences and Supply Chain Performance: An Experimental Study, Management Science, 54(11), 1835-1849.

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

Schultz, K. L., McClain, J. O.,Robinson, L., Thomas, L. J. (2007). The Use of Framing in Inventory Decisions. Working paper, Johnson School Research Papers #02, Cornell University.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Siegel, S. (1956). Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

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

Su, X. (2008). Bounded Rationality in Newsvendor Models. Manufacturing and Service Operations Management, 10(4), 566-589.

Tversky, A., Kahneman. D. (1981). The Framing of Decisions and the Psychology of Choice. Science, 211, 453-458.

Wu, D. Y. (2013). The Impact of Repeated Interactions on Supply Chain Contracts: A Laboratory Study. International Journal of Production Economics, 142, 3-15.

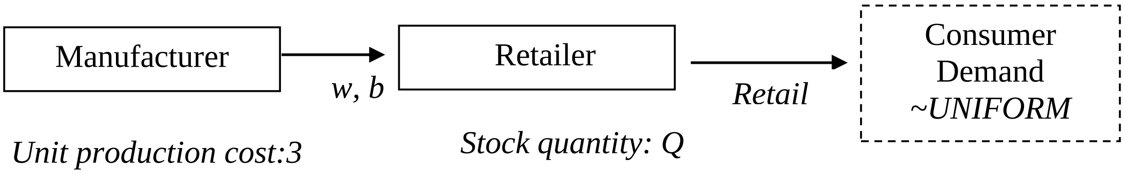
Zhang, Y., Donohue, K., Cui, H. T. (2016). Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing. Management Science, 62(6), 1734–1754.

3.9 Appendix

3.9.1 Sample Experiment Instructions (Buyback Contract)

Scenario

We consider a manufacturer who produces a certain product, and a retailer who buys the product from the manufacturer and sells it to consumers. Consumer demand is uncertain. It is a random number distributed uniformly *between 51 and 150*. That is, there is a 1/100 chance that demand will be equal to any of the integers between 51 and 150.



We consider a three-stage game between the manufacturer and the retailer:

Stage-1: The manufacturer determines the two contract parameters:

- ***Wholesale price, w***. This is the price at which the manufacturer sells his product to the retailer. The wholesale price has to be an integer between the ***unit production cost of the manufacturer 3,*** and the ***retail price 12***. Retail price is the price at which the retailer sells the product to consumers.
- ***Buyback price, b***: This is the price at which the manufacturer will buy back unsold products from the retailer. Buyback price should be lower than or equal to the wholesale price.

Stage-2: The retailer observes the wholesale price and buyback price offers of the manufacturer, and determines his ***stock quantity, Q*** for the product. The retailer may ***reject*** the manufacturer's offer by setting ***Q=0***. In this case, both firms earn zero profit. Otherwise, the retailer orders Q products from the manufacturer. The manufacturer produces this order by incurring the unit production cost 3 per product, and delivers them to the retailer. The retailer stocks these products prior to the selling season. Because consumer demand can be between 51 and 150, the retailer's stock quantity Q decision also has to be between these values (if it is not equal to zero).

Stage-3: Random consumer demand is determined as "*d*". Using his stock of product, the retailer satisfies this demand as much as possible. The ***sales quantity*** of the retailer is the minimum of his stock quantity and the realized demand. Two cases are possible:

- If demand is higher than retailer's stock quantity (i.e., $d>Q$), then retailer will sell all Q units, and $(d-Q)$ units of demand will be unsatisfied. (***unsatisfied demand***).
- If demand is less than the retailer's stock quantity, (i.e., $d<Q$), then the retailer will sell d units, and $(Q-d)$ products will be unsold (***leftover products***). These products have zero value. The manufacturer buys back these products by paying the buyback price b per product to the retailer.

Each firm aims to maximize its payoff (or, profit) in the game.

The retailer's payoff is calculated as the retail price times the sales quantity, minus the payment to the manufacturer, plus buyback payment from the manufacturer. That is, $\boldsymbol{12 * sales - w * Q + b * leftovers}$**.**

The manufacturer's payoff is calculated as the payment received from the retailer, minus the production cost, minus the buyback payment to retailer. That is, $\boldsymbol{w * Q - 3 * Q - b * leftovers}$.

Note that there are three decisions in the game: The manufacturer determines w and b; afterwards, the retailer determines his stock quantity Q. Both firms' decisions affect the payoff of both firms.

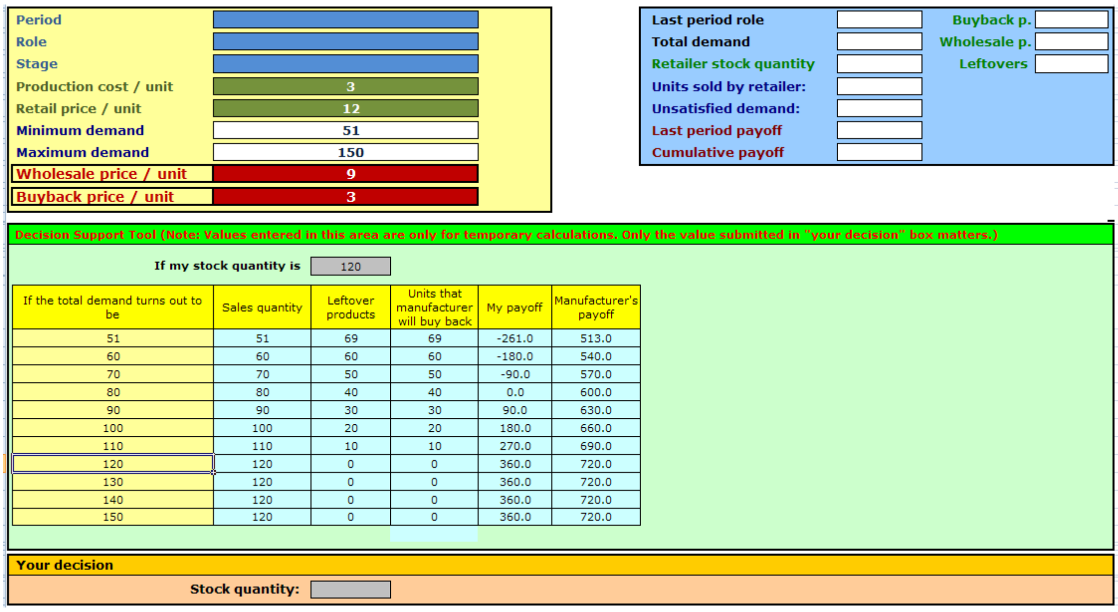
Preparation for Our Experiments

- The experiments will take place at the CAFÉ 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.

Our Experiment

- You will be randomly assigned to the role of either a manufacturer or a retailer. Manufacturer and retailers will be randomly matched with each other. You will play with the same partner throughout the experiment.
- The experiment repeats for a number of ***periods***, which are independent of each other. That is, a large or small demand realization in a period does not affect the demand in the later periods. Leftover products from a period are returned to the manufacturer, and thrown away. They cannot be used to satisfy demand in following periods. Only your payoff will accumulate over periods.

A Sample Screenshot: The following figure illustrates how the retailer's screen will look like at stage 2:



If the total demand turns out to be	Sales quantity	Leftover products	Units that manufacturer will buy back	My payoff	Manufacturer's payoff
51	51	69	69	-261.0	513.0
60	60	60	60	-180.0	540.0
70	70	50	50	-90.0	570.0
80	80	40	40	0.0	600.0
90	90	30	30	90.0	630.0
100	100	20	20	180.0	660.0
110	110	10	10	270.0	690.0
120	120	0	0	360.0	720.0
130	120	0	0	360.0	720.0
140	120	0	0	360.0	720.0
150	120	0	0	360.0	720.0

Figure 1: Retailer's screen 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 and the wholesale price and buyback price that the manufacturer set at stage 1.

- 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). Note that the submit button will be activated only after you enter a valid decision and hit enter (or, click somewhere in the screen). Invalid entries will cause warnings.
- The cells in which you can enter values are labeled with "gray" background.
- You can check the results of previous periods by clicking the ***Historical Results*** tab in the bottom of the screen. This will open a second worksheet with the titles seen below (for manufacturer):

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

Figure 2: Historical results table (manufacturer)

The Decision Support Tool

Before you submit a decision, you can use the decision support tool that is in the middle of the screen. This tool allows you to calculate the outcome for certain values of your decision, the other firm'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 we record is the one you submit in the "stock quantity" box at the bottom of the screen.

Retailer's decision support tool at stage-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 summarizes the outcome for different consumer demand realizations (*d=51, 60, …, 150*) each in a row.

In the example in Figure 1, the retailer's stock quantity is entered as 120. We observe from the table that if consumer demand turns out to be, for example, 80, you (retailer) will sell 80 units because the demand is smaller than the stock quantity. Your leftover inventory will be 120-80=40 units. These units will be bought back by the manufacturer. 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 140. In this case, you (the retailer) will sell all of your 120 units, and there will be zero leftover inventory. Unsatisfied demand will be 140-120=20. The last two columns provide your payoff and the manufacturer's payoff.

Manufacturer's decision support tool at stage-1

At stage-1, 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. Figure 3 below illustrates what the outcome will be if you set 6 as your wholesale price, 3 as your buyback price, and if the retailer sets 120 as his stock quantity (i.e., if he orders 120 products from you).

Decision Support Tool (Note: values entered in this area are only for temporary calculations

If my wholesale price is 9
and my buyback price is 3

and retailer's stock quantity is 120

If the total demand turns out to be	Retailer's sales quantity	Leftover products at the retailer	Units that I should buy back	My payoff	Retailer's payoff
51	51	69	69	513.0	-261.0
60	60	60	60	540.0	-180.0
70	70	50	50	570.0	-90.0
80	80	40	40	600.0	0.0
90	90	30	30	630.0	90.0
100	100	20	20	660.0	180.0
110	110	10	10	690.0	270.0
120	120	0	0	720.0	360.0
130	120	0	0	720.0	360.0
140	120	0	0	720.0	360.0
150	120	0	0	720.0	360.0

Figure 3: Manufacturer's decision support tool at Stage 1.

3.9.2 Comparison Results when Rejected Contracts Included

Comparison of experiment results with theory.

	Wholesale Price		Buyback		Revenue Sharing	
	Theory	**Experiment**	**Theory**	**Experiment**	**Theory**	**Experiment**
w	10	7.62 (0.55)***	11	8.83 (0.81)***	1	4.38 (1.63)***
b/r	-	-	10	5.08 (1.68)***	10	4.13 (1.72)***
Q	67	87.87 (15.77)***	100	88.6 (12.14)**	100	87.79 (14.92)**
Retailer's Profit	117.68	278.7 (58.05)***	75.5	217.54 (65.18)***	75.5	245.87 (103.78)***
Mfg.'s Profit	469	382.52 (85.19)***	677.5	438.86 (108.13)***	677.5	414.98 (104.19)***
Total Chain Profit	586.68	661.22 (92.83)**	753	656.4 (76.01)***	753	660.85 (92.26)***
Contract Efficiency	0.74	0.83 (0.12)**	0.95	0.83 (0.1)***	0.95	0.83 (0.12)***

Standard deviations are reported in parenthesis. P values are from a two tailed Mann Whitney U test.
**** p-value<0.001, ** p-value<0.01, * p-value<0.1*

Comparison of WSP contract with BB and RS contracts.

	Experiment – Mean (Std. Dev.)			P - Values	
	WSP	**BB**	**RS**	**WSP vs BB**	**WSP vs RS**
Critical fractile	0.36 (0.05)	0.49 (0.13)	0.46 (0.14)	**0**	**0.01**
Q	87.87 (15.77)	88.6 (12.14)	87.79 (14.92)	0.39	0.4
Q/Q*	1 (0.18)	0.91 (0.17)	0.91 (0.17)	**0.01**	**0.01**

Retailer's Profit	278.7 (58.05)	217.54 (65.18)	245.87 (103.78)	**0**	0.19
Manufacturer's Profit	382.52 (85.19)	438.86 (108.13)	414.98 (104.19)	**0.05**	0.25
Total Chain Profit	661.22 (92.83)	656.4 (76.01)	660.85 (92.26)	0.18	0.44
Contract Efficiency	0.83 (0.12)	0.83 (0.1)	0.83 (0.12)	0.18	0.44
Underorder Quantity	11.65 (9.18)	18.87 (11.33)	15.77 (10.49)	**0.01**	**0.05**
Overorder Quantity	13.06 (7.86)	8.65 (6.44)	7.54 (6.27)	**0.03**	**0.01**
#Underorders	16.55 (9.72)	19.18 (10.3)	20.43 (8.25)	0.18	**0.03**
#Overorders	22.73 (9.98)	18.59 (10.34)	17.86 (9.05)	**0.08**	**0.02**

P values are from one tailed Mann Whitney U test.

Comparison of BB and RS contract scenarios

	BB	RS	P-Value
Critical fractile	0.49 (0.13)	0.46 (0.14)	0.52
Q	88.6 (12.14)	87.79 (14.92)	0.90
Q/Q*	0.91 (0.17)	0.91 (0.17)	0.95
Retailer's Profit	217.54 (65.18)	245.87 (103.78)	0.34
Manufacturer's Profit	438.86 (108.13)	414.98 (104.19)	0.38
Total Chain Profit	656.4 (76.01)	660.85 (92.26)	0.69
Contract Efficiency	0.83 (0.1)	0.83 (0.12)	0.69
Underorder Quantity	18.87 (11.33)	15.77 (10.49)	0.35
Overorder Quantity	8.65 (6.44)	7.54 (6.27)	0.41
#Underorders	19.18 (10.3)	20.43 (8.25)	0.52
#Overorders	18.59 (10.34)	17.86 (9.05)	0.69

Standard deviations are reported in parenthesis. P values are from a two tailed Mann Whitney U test.

Chapter 4

4. DIFFERENT UTILITY FUNCTIONS

Different Utility Functions

4.1 Introduction

In this chapter we develop behavioral models involving risk-aversion, loss-aversion, inventory error aversion and social preferences. Our aim is to come up with better explanations of the retailer and manufacturer behavior observed in the experimental study of Chapter 3, than standard theory.

Due to the high heterogeneity of the subjects, each model is estimated for each retailer and manufacturer separately. To illustrate this heterogeneity, Figure 7 displays the order quantity decisions of all buyback contract retailers. Each plot corresponds to a retailer and each x represents an ordering decision. The x-axis corresponds to the standard newsvendor optimal order and the y-axis is the actual experiment orders of the retailer.

In the remaining of the chapter we first present the solution of the standard theory through a "general contract" notion. Then we develop and analyze retailer's risk-aversion (Section 3), loss-aversion (Section 4), inventory error aversion (Section 5) and social preference models (Section 6) for all three contract types. We find that for most retailers, these models yield a better prediction than the standard newsvendor model. In Section 7, we develop social preference models for the manufacturers, and show these models to explain pricing decisions of the WSP manufacturers with relatively high accuracy.

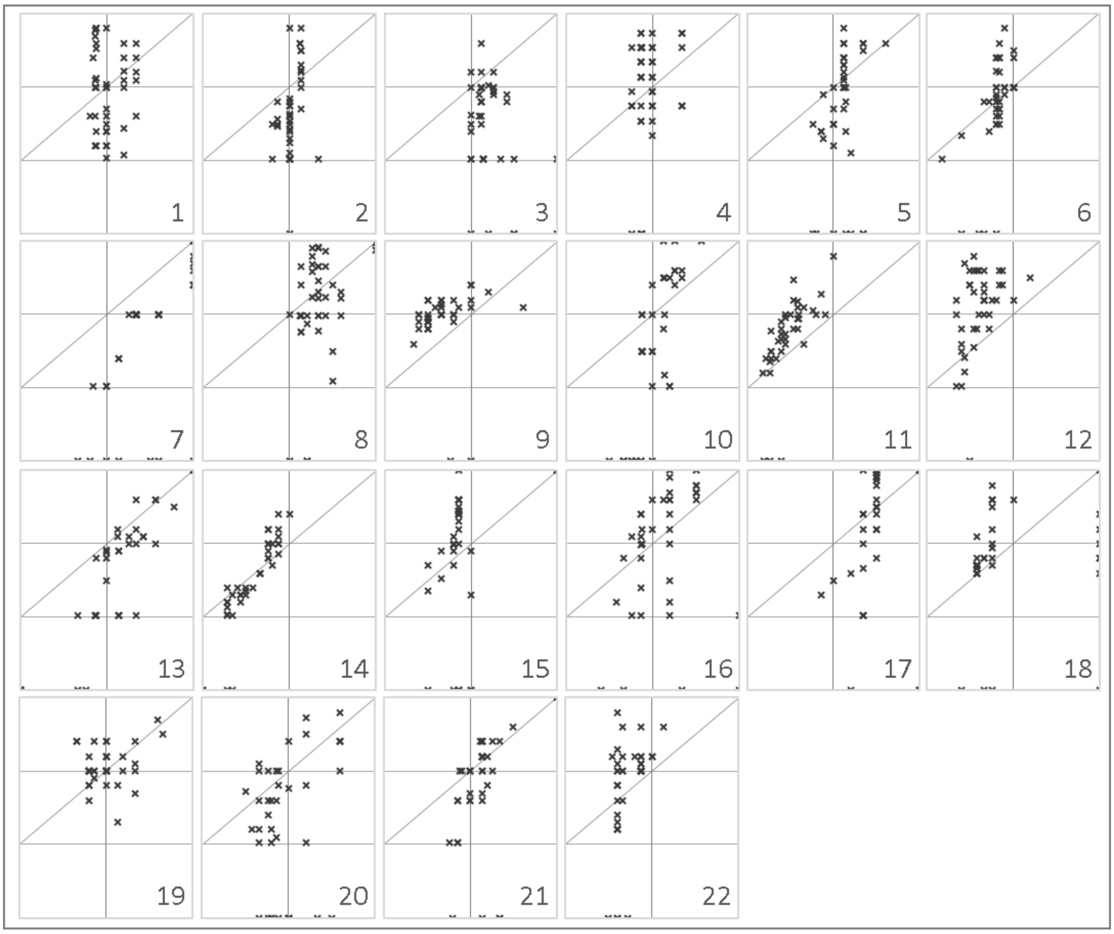


Figure 7: Order decision vs *Q*(w,b)* under the buyback contract

4.2 Theoretical Background

Throughout the theoretical model development process we adopt a "general contract" notation, which is a generalized contract format that covers all three contract types we consider. The results for this general contract can be easily adapted to wholesale price, buyback or revenue sharing contracts by eliminating the redundant parameter(s). Here we provide the solution of the standard theory with this general contract form.

<u>Notation:</u>

- *Q:* Retailer's order quantity decision.
- *Q*:* The newsvendor-optimal order quantity.
- *w*: Wholesale price of the general contract.
- *b*: Buyback price of the general contract.
- *r*: Revenue share of the general contract.

- p: Unit selling price to the consumer. *(p=$12)*
- c: Unit production cost. *(c=$3)*
- L: Lower bound of the demand distribution. *(L=51)*
- U: Upper bound of the demand distribution. *(U=150)*
- *F(.)*: The distribution function of consumer demand.

For a given wholesale price w, buyback price b and revenue share r, and for a given order quantity decision Q, the retailer's expected profit can be found as:

$$\begin{aligned} E(\pi_R \mid Q) &= (p-r)E(\text{Sales} \mid Q) + bE(\text{Leftovers} \mid Q) - wQ \\ &= (p-r)\big(Q - E(\text{Leftovers} \mid Q)\big) + bE(\text{Leftovers} \mid Q) - wQ \\ &= (p-r-w)Q + (b-p+r)\int_L^Q (Q-x)F(x)dx \end{aligned}$$

The optimal order quantity to maximize the above expected profit function is derived as

$$Q^* = F^{-1}\left(\frac{p-r-w}{p-r-b}\right)$$

For any order quantity Q, the manufacturer's expected profit can be found as:

$$\begin{aligned} E(\pi_M \mid Q) &= wQ + rE(\text{sales} \mid Q) - bE(\text{leftover inv.} \mid Q) - cQ \\ &= (w-c)Q + r(Q - E(\text{leftover inv.} \mid Q)) - bE(\text{leftover inv.} \mid Q) \\ &= (w-c+r)Q - (b+r)\int_L^Q (Q-x)F(x)dx \end{aligned}$$

Given that the retailer will order *Q*(w,b,r)*, the manufacturer's optimal contract parameters can be found through backwards induction by maximizing the below expected profit function:

$$E(\pi_M) = (w-c+r)Q^* - (b+r)\int_L^Q (Q^*-x)F(x)dx$$

Under the general contract, the above function is <u>not</u> jointly concave in all contract parameters. Figure 8 and Figure 9 visualize manufacturer's expected profit function for buyback and revenue sharing contracts under our experimental settings.

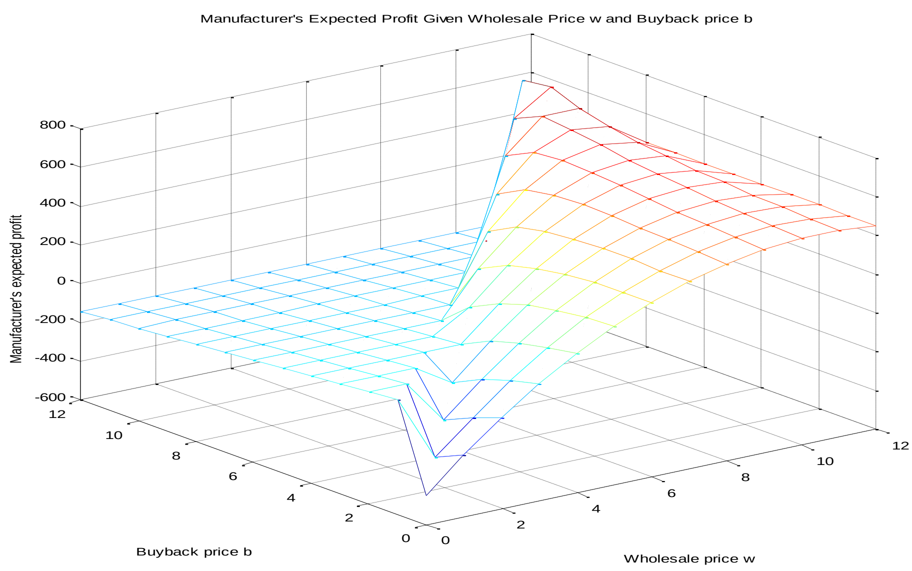


Figure 8: Manufacturer's expected profit under the buyback contract

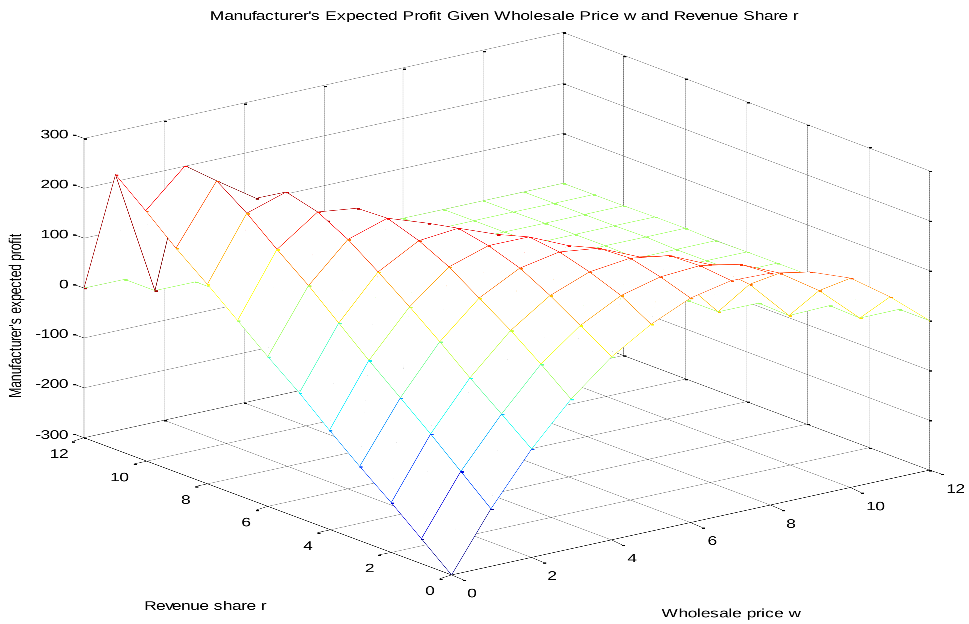


Figure 9: Manufacturer's expected profit under the revenue sharing contract

Under buyback and revenue sharing contracts, we can’t derive closed form expressions for the manufacturer’s best contract. Instead, we determine these through a numerical grid search. Under the wholesale price contract, manufacturer’s expected utility function is concave in *w* and hence the optimal wholesale price for the manufacturer can be derived as:

$$w^* = \frac{U}{(U - L + 1)} \frac{p}{2} + \frac{c}{2}$$

Figure 10 displays the expected profit of the supply chain under the wholesale price contract, given that the retailer orders *Q*(w)*. In all derived results, the retailer's utility function is assumed to be known by the manufacturer.

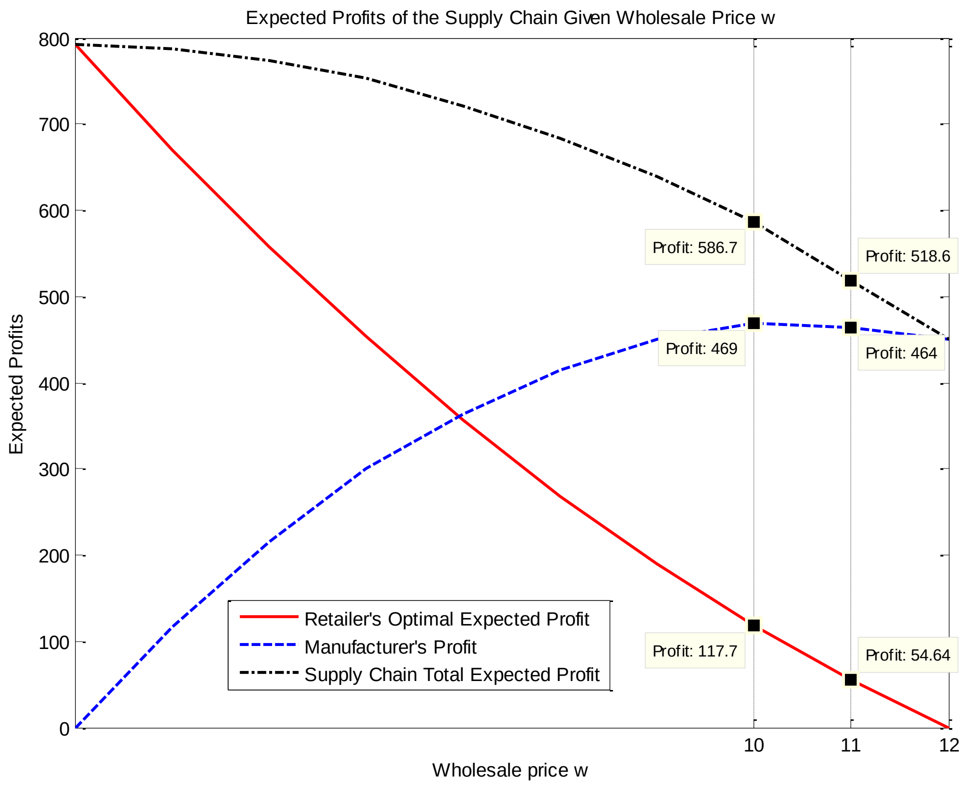


Figure 10: Expected profits of the supply chain under the wholesale price contract

4.3 Retailer's Risk Aversion Models

Standard newsvendor model assumes the retailer is risk-neutral. However, the fact that decision makers have various levels of sensitivity towards risk is widely acknowledged in today's economic literature. Schweitzer and Cachon (2000) explain that a risk-averse retailer facing the newsvendor problem will place an order below, and a risk-seeking retailer will place an order above the optimal order quantity. The authors base their argument on the second derivative of the utility function being positive or negative without using a specific utility function. However the too-low-too-high pattern or the pull-to-center effect observed in their study, which is also supported by our findings, suggests that retailers are not completely risk-averse or completely risk-seeking. In this regard, Schweitzer and Cachon propose that prospect theory, which asserts that risk attitude of individuals depend on being in loss or gain domains, may explain the pull-to-center effect.

According to prospect theory (Kahneman and Tversky 1979), in gains domain where gains are highly probable and risk of losses is small, individuals exhibit risk-aversion, while in loss domain they exhibit risk-seeking behavior. Thus, according to prospect theory, in our context, a high profit margin will induce risk-aversion and a low profit margin will induce risk-seeking behavior in retailers.

Within the existing behavioral operations literature, studies on retailer's risk attitude do not use a specific utility function. Rather, they base their risk attitude argument on the position of the order quantity decision relative to the optimal order quantity.

In this study we develop a risk attitude utility function where the definition of risk is based on the variability in the profit to be earned. Even though the definition of risk as variability in the profit is not new, the utility function used in this section is our contribution to the literature. The risk of an order decision is measured as the difference between the highest and lowest possible profit values to be earned given that order decision. Accordingly, the utility function of the retailer is defined as

$$U(\pi) = \pi - \delta([\pi_{\max} \mid Q] - [\pi_{\min} \mid \mathrm{Q}]).$$

For a risk-averse retailer, as the utility will decrease with risk, the δ coefficient will take a positive value. Similarly for a risk-seeking retailer the δ coefficient will be negative. The order quantity to maximize retailer's expected utility function is derived as:

$$Q^r = F^{-1}\left(\frac{(p-r-w)-\delta(p-r-b)}{p-r-b}\right) = F^{-1}\left(\frac{p-r-w}{p-r-b} - \delta\right).$$

Proof:

The minimum profit will occur when the demand realization is at the lower bound of the demand distribution, and the maximum profit will occur when the demand realization is at the upper bound. Here we don't assume any distributional restriction on the consumer demand. We only assume that the demand is restricted to the interval [L, U].

$$[\pi_{\max} \mid Q] = (p-r)\min(U,Q) + b[Q - \min(U,Q)] - wQ = (p-r-b)Q + (b-w)Q$$

$$[\pi_{\min} \mid \mathrm{Q}] = (p-r)\min(L,Q) + b[Q - \min(L,Q)] - wQ = (p-r-b)L + (b-w)Q$$

$$[\pi_{\max} \mid Q] - [\pi_{\min} \mid \mathrm{Q}] = (p-r-b)(Q-L)\ .$$

Combining the above results with earlier newsvendor results, we obtain:

$$E[U(\pi \mid Q)] = (p-r-w)Q - (p-r-b)\int_0^Q (Q-x)dF(x) - \delta(p-r-b)(Q-L)\,.$$

The proof follows as in the standard newsvendor solution. If δ is positive, we have $Q^r = F^{-1}\left(\frac{(p-r-w)-\delta(p-r-b)}{p-r-b}\right) \le F^{-1}\left(\frac{p-r-w}{p-r-b}\right) = Q*$ which is the risk neutral optimal order quantity. If δ is negative then we have $Q^r \ge Q*$. ■

4.3.1 Model Estimation

Combining the above findings with the assumption that consumer demand is uniform on the interval [51,150], the expected utility maximizing order quantity is found as

$$Q^r = 50 + \left(\frac{p-r-w}{p-r-b} - \delta\right) * 100\,.$$

To estimate the δ coefficient, this equation is transformed into the linear equation below:

$$\frac{p-r-w}{p-r-b} - \frac{(Q-50)}{100} = \delta + \varepsilon\ .$$

In this equation we assume the error term ε is normally distributed with 0 mean and independent of each other and other variables.

Table 14 to Table 16 display the results of the estimation that is performed for each retailer separately in MS Excel. Each row corresponds to a single retailer. The results of three different analyses are presented; only the high profit margin (HPM) contracts, only the low profit margin (LPM) contracts, and all accepted contracts together (All). Contracts with profit margin exactly equal to 50% are included in the all accepted contracts analysis.

Estimation results which conform to our sign expectations on δ regarding risk behavior are highlighted with purple background. The last four columns compare these model estimations with the standard theory using average absolute percentage deviation. To this end, the predicted order decision is computed using the estimated model coefficients for each retailer and for each period. Then the difference between the actual order and this prediction is calculated.

For the WSP contract, most of the retailers are found to satisfy the risk-seeking behavior expectation under low profit margin, and most of the retailer's models are statistically significant. However, there are also retailer that have significant models which don't satisfy the risk-seeking behavior expectation. Overall, for 12 of the retailers, the model estimation provides a better prediction than the standard theory.

Due to their two parameter design, high profit margin periods are observed more under the BB and RS contracts than under the WSP contract. Under the BB contract, the risk-attitude models are observed to be better and more significant at explaining the changes in the order decisions than under WSP or RS contracts. Additionally, the models fit data with higher R^2 values under BB and RS contracts than under WSP contract.

To the right of the estimation results, the tables present the average absolute percent deviation of the model-predicted order decisions from the experiment data. From these, we observe the risk-aversion model to have better prediction accuracy than the standard newsvendor model for the majority of the retailers under all three contracts. The models with the highest prediction accuracy are printed in boldface. Based on all of these findings, we can conclude that the risk attitude model is partially successful at explaining retailer subjects' ordering decisions.

Table 14: Estimation results of the risk attitude model for WSP retailers

	HPM			LPM			All Contracts			Avg. Absolute % Deviation			
	#	δ	R^2	#	δ	R^2	#	δ	R^2	HPM	LPM	All	Q*(w)
R1				31	0.21**	**93%**	35	0.16**	**62%**		69%	**13%**	33%
R2				38	-0.13**	25%	38	-0.13**	25%		27%	23%	**22%**
R3				33	0.00*	8%	34	-0.08	10%		26%	26%	25%
R4				36	-0.18**	44%	39	-0.19**	48%		32%	**17%**	21%

R5	2	-0.25**	**100%**	19	0.05*	3%	38	-0.15**	24%	33%	32%	**26%**	28%
R6				35	-0.15**	40%	37	-0.16**	43%		28%	**16%**	18%
R7	1			19	0.16**	**67%**	40	0.02	1%		48%	**18%**	**18%**
R8				33	0.00*	0%	39	-0.04	3%		**21%**	22%	**21%**
R9				38	-0.26**	**86%**	39	-0.25**	**85%**		46%	**9%**	24%
R10				39	-0.18**	42%	39	-0.18**	42%		34%	**20%**	22%
R11				37	-0.21**	34%	37	-0.21**	34%		37%	31%	**30%**
R12				34	-0.07*	12%	39	-0.10**	20%		21%	**19%**	20%
R13				33	0.04**	26%	40	0.01	2%		12%	**8%**	9%
R14	3	-0.08	44%	23	0.14**	**50%**	29	0.08*	18%	**14%**	46%	21%	24%
R15	3	-0.08	17%	35	0.00*	0%	38	-0.01	0%	**15%**	24%	23%	23%
R16				37	-0.09**	17%	38	-0.09**	17%		22%	19%	**18%**
R17				37	-0.08*	15%	37	-0.08*	15%		19%	18%	**17%**
R18	1			4	-0.06*	5%	33	-0.32**	**85%**		29%	**8%**	27%
R19				34	-0.04*	7%	34	-0.04	7%		15%	**14%**	15%
R20				33	-0.08**	27%	38	-0.08**	27%		15%	**12%**	**12%**
R21				36	-0.17**	**54%**	37	-0.17**	**54%**		31%	**10%**	17%
R22				25	0.21**	**66%**	27	0.21**	**68%**		71%	**21%**	38%

**: Significant at 5% level. **: Significant at 1% level.*

Table 15: Estimation results of the risk attitude model for BB retailers

	HPM			LPM			All Contracts			Avg Absolute % Deviation			
	#	δ	R^2	#	δ	R^2	#	δ	R^2	HPM	LPM	All	Q*(w.b)
R1	12	0.10	16%	16	-0.12	17%	40	0.03	1%	**20%**	26%	27%	28%
R2	12	-0.03	1%	6	0.18*	**74%**	39	0.17**	31%	**23%**	54%	26%	39%
R3	28	0.26**	48%				31	0.26**	47%	74%		**29%**	44%
R4	5	0.02	1%	19	-0.18**	49%	34	-0.13**	31%	21%	31%	**18%**	19%
R5	23	-0.01	1%	5	0.17*	**73%**	26	0.07	11%	**16%**	49%	19%	20%
R6				30	-0.06	10%	32	-0.05	10%		14%	14%	**12%**
R7	30	0.20**	**83%**	1			33	0.23**	**79%**	36%		**11%**	27%
R8	36	0.03	2%				37	0.03	2%	19%		**17%**	18%
R9	2	0.10	31%	32	-0.20**	**89%**	38	-0.18**	**75%**	19%	40%	**7%**	19%
R10	16	-0.02		8	0.13*	**60%**	29	0.05	4%	31%	35%	**28%**	29%
R11				36	-0.16**	**69%**	37	-0.16**	**68%**		33%	**10%**	17%
R12	1			37	-0.26**	**61%**	39	-0.26**	**61%**		45%	**18%**	25%
R13	22	0.18**	40%	5	0.33*	**83%**	31	0.20**	47%	44%	122%	**23%**	33%
R14				35	-0.01	1%	36	-0.02	2%			12%	**11%**
R15	2	0.00**	**100%**	28	-0.22**	**68%**	32	-0.18**	49%		38%	**14%**	19%
R16	20	-0.01		14	0.00	0%	36	-0.02	0%	32%	**24%**	28%	28%
R17	34	0.01		1			36	0.03	1%	**23%**		**23%**	**23%**
R18	7	0.44**	**80%**	28	-0.14**	44%	36	-0.03	1%	47%	27%	**18%**	20%
R19	10	0.13*	45%	17	-0.12**	48%	40	-0.03	4%	12%	**9%**	11%	11%
R20	9	0.10	14%	18	0.09	19%	29	0.08	14%	30%	33%	**21%**	25%
R21	23	0.00		7	0.19*	**51%**	36	0.05	12%	**8%**	66%	15%	15%
R22	1			33	-0.12**	29%	36	-0.12	32%	35%	24%	**14%**	17%

**: Significant at 5% level. **: Significant at 1% level.*

Table 16:Estimation results of the risk attitude model for RS retailers

	HPM			LPM			All Contracts			Average Absolute % Deviation			
	#	**δ**	**R^2**	**#**	**δ**	**R^2**	**#**	**δ**	**R^2**	**HPM**	**LPM**	**All**	**Q*(w.r)**
R1	33	-0.10*	18%	2	-0.37*	**100%**	37	-0.46**	19%	29%		**21%**	26%
R2	30	0.00	0%	6	-0.15**	**96%**	38	-0.56**	29%	**11%**	40%	13%	14%
R3	3	-0.49	**88%**	33	-0.09**	23%	40	-0.21**	27%	136%	32%	**22%**	29%
R4	8	0.05	11%	13	-0.04	25%	32	-0.26*	14%	**11%**	12%	12%	12%
R5				38	0.25**	**71%**	38	0.43**	40%		46%	**15%**	23%
R6	22	-0.05	8%	13	-0.18**	**74%**	38	-0.25	7%	**17%**	57%	**17%**	22%
R7	28	-0.01	0%				29	-0.10	1%	**27%**		29%	29%
R8	25	-0.09	7%	6	0.14	32%	37	0.36	4%	30%	**25%**	27%	27%
R9				39	0.09**	24%	39	0.15*	13%		20%	**15%**	17%
R10				38	0.25**	**80%**	40	0.35**	**67%**		43%	**10%**	26%
R11	1			31	-0.09*	14%	32	-0.11	5%		40%	**24%**	30%
R12				36	-0.04	9%	39	-0.07	8%		19%	**15%**	16%
R13	2	-0.03	2%	28	-0.01	0%	35	0.21	2%	**21%**	27%	24%	25%
R14				17	0.05	8%	37	-0.11	1%		**13%**	15%	15%
R15	31	0.02	1%	3	-0.12	**61%**	39	-0.16	5%	**15%**	32%	**15%**	16%
R16	8	0.05	3%	5	-0.28*	**69%**	19	-0.24	8%	**25%**	107%	29%	35%
R17	20	0.06**	33%	14	-0.11	21%	40	0.06	2%	**11%**	41%	19%	19%
R18	5	0.05	42%	23	-0.14**	35%	34	-0.07	1%	**8%**	46%	22%	24%
R19				14	-0.14	14%	27	1072.00	2%		70%	**40%**	50%
R20	32	-0.04*	13%	2	-0.05	23%	38	-0.10	2%	12%		**11%**	13%
R21				26	-0.01	1%	40	0.13	7%		**15%**	**15%**	**15%**
R22	21	-0.14*	27%	6	-0.02	4%	40	-0.17	6%	36%	**14%**	17%	20%

**: Significant at 5% level. **: Significant at 1% level.*

4.4 Retailer's Loss Aversion Models

Loss-aversion is another fundamental decision making concept which was introduced by Kahneman and Tversky (1984). Out of a loss and gain of equal magnitude, the loss has a stronger effect and importance on a loss-averse decision maker. Specifically, if W_o denotes the initial and W denotes the current wealth, the utility function of a loss-averse individual has the following form:

$$U(\mathrm{W}) = \begin{cases} W - W_o & if\ \ \mathrm{W} \geq W_o \\ \lambda(\mathrm{W} - W_o) & if\ \ \mathrm{W} \leq W_o \end{cases}$$

where λ is greater than 1. Using this utility function, Wang and Webster (2009), and Ho and Zhang (2008) study loss-averse retailer, and Zhang et al. (2016) study loss averse supplier behavior in supply chain contracting experiments.

In this section, we model retailer's loss-aversion with three different models. In the first one, we adopt the above utility function. In the second and third, we redefine what is being considered as a loss by the retailer.

4.4.1 Loss-Aversion Model -1

In this model the retailer considers the cost of unsold products and the revenue share to be paid to the manufacturer as losses. We assume that the retailer has a utility function of the form

$U(Q) = (\mathrm{p} - \mathrm{w})\,\mathrm{Q} - \lambda_1(\mathrm{p} - \mathrm{b})(Q - D)^+ - \lambda_2 r[Q - (Q - D)^+]$.

The order quantity to maximize the retailer's utility function is derived as:

$$Q^l = F^{-1}\left(\frac{p - \lambda_2 r - w}{\lambda_1(p - b) - \lambda_2 r}\right).$$

Proof:

We rearrange the terms in the expected utility function as follows:

$$E[U(Q)] = (\mathrm{p} - \mathrm{w})\,\mathrm{Q} - \lambda_1(\mathrm{p} - \mathrm{b})\int_0^Q (Q - x)dF(x) - \lambda_2 r\left(Q - \int_0^Q (Q - x)dF(x)\right)$$

$$E[U(Q)] = (\mathrm{p} - \mathrm{w} - \lambda_2 r)\,\mathrm{Q} - [\lambda_1(\mathrm{p} - \mathrm{b}) - \lambda_2 r]\int_0^Q (Q - x)dF(x)$$

The proof follows as the standard newsvendor solution. For a loss-averse retailer, as losses will have greater impact than their monetary value, λ_1 and λ_2 coefficients will both be greater than 1, which will cause Q^l to be less than Q^*.■

4.4.2 Loss-Aversion Model -2

In this model, retailer's utility function increases with the lowest possible profit that can be earned given the order quantity decision. As the lowest possible profit value increases, the probability of loss decreases, thus the utility increases. We define the utility function as $U(\pi \mid Q) = [\pi \mid Q] + \lambda[\pi_{\min} \mid \mathrm{Q}]$.

The order quantity to maximize the expected utility is derived as

$$Q^{l} = F^{-1}\left(\frac{p-r-w}{p-r-b} - \lambda\frac{w-b}{p-r-b}\right).$$

Proof:

Rearranging the terms in the expected utility function we obtain

$$E[U(\pi \mid Q)] = E[\pi \mid Q] + \lambda\left((p-r-b)L + (b-w)Q\right)$$

$$E[U(\pi \mid Q)] = \left[p-r-w+\lambda(b-w)\right]Q - (p-r-b)\int_0^Q (Q-x)dF(x) + \lambda(p-r-b)L$$

The proof follows as the standard newsvendor solution. Again for a loss-averse retailer, λ will be positive, which will cause Q^l to be less than Q^*.■

4.4.3 Loss-Aversion Model -3

In this model, in addition to the factors in the second loss-aversion model, retailer's utility function is affected positively by the highest possible profit value. We define the utility function as $U(\pi \mid Q) = [\pi \mid Q] + \lambda_1[\pi_{\min} \mid \mathrm{Q}] + \lambda_2[\pi_{\max} \mid \mathrm{Q}]$.

The order quantity to maximize the expected utility is derived as

$$Q^{l} = F^{-1}\left(\frac{p-r-w}{p-r-b} - \lambda_1\frac{w-b}{p-r-b} + \lambda_2\frac{p-r-w}{p-r-b}\right).$$

Proof:

Rearranging the terms in the expected utility function we get:

$$\begin{aligned} E[U(\pi \mid Q)] &= E[\pi \mid Q] + \lambda_1[\pi_{\min} \mid \mathrm{Q}] + \lambda_2[\pi_{\max} \mid Q] \\ &= E[\pi \mid Q] + \lambda_1\left((p-r-b)L + (b-w)Q\right) + \lambda_2\left((p-r-b)Q + (b-w)Q\right) \\ &= \left[p-r-w+(\lambda_1+\lambda_2)(b-w)+\lambda_2(p-r-b)\right]Q - (p-r-b)\int_0^Q (Q-x)dF(x) \\ &\quad + \lambda_1(p-r-b)L \end{aligned}$$

The proof follows as the standard newsvendor solution. For a loss-averse retailer both λ_1 and λ_2 will take positive values. However, depending on the values of the λ coefficients and the contract parameters, Q^l may be less than or greater than Q^*.■

4.4.4 Model Estimation

We rearrange and linearize the optimal order quantity equations and run linear regression analyses. The linear regression equations we use are given below:

Model 1: $p - w = \lambda_1 \frac{Q-50}{100}(p-b) + \lambda_2 \frac{150-Q}{100} r + \varepsilon$

Model 2: $\frac{p-r-w}{p-r-b} - \frac{Q-50}{100} = \lambda_1 \frac{w-b}{p-r-b} + \varepsilon$

Model 3: $\frac{p-r-w}{p-r-b} - \frac{Q-50}{100} = \lambda_1 \frac{w-b}{p-r-b} - \lambda_2 \frac{p-r-w}{p-r-b} + \varepsilon$

Estimation results for these three loss aversion models are presented in Table 17 to Table 19. Loss-aversion model-1 seems to be better fitting to the data for all contract types as the R^2 values obtained for all retailers are higher than 50%. R^2 values greater than 50% are obtained for very few of the retailers for the second model and for about half of the retailers for the third model. In terms of significance, again, the first model is the strongest among the three.

Loss-aversion behavior is expected to yield $\lambda \geq 1$ for the first model and $\lambda \geq 0$ for the second model. However, we observe that these expectations are satisfied for only few of the retailers for WSP and BB contracts, signaling that loss-aversion is not strong under these contracts. Under the RS contract for the first model, λ_1 is estimated to be greater than 1 for all retailers while λ_2 is estimated to be less than 1. This can be interpreted as aversion from losses due to leftovers being stronger than the aversion due to lost revenue share that is paid to the manufacturer.

The average absolute percent deviations of the model-predicted order quantities from the experiment data are provided on the rightmost columns of the tables. The model with the highest prediction accuracy is printed in boldface. For more than half of the retailers, one of the three loss aversion models is seen to provide an order quantity prediction better than *Q**. Hence we can claim the loss-aversion models to be partially successful in explaining retailer behavior.

Table 17: Estimation results of loss aversion models for WSP retailers

	Model1		Model2		Model3			Average Absolute % Deviation			
	λ	**R^2**	**λ**	**R^2**	**λ_1**	**λ_2**	**R^2**	**Model1**	**Model2**	**Model3**	**Q*(w)**
R1	1.06**	**67%**	0.27**	**72%**	0.69**	0.62*	**90%**	31%	10%	**4%**	33%
R2	0.6**	**85%**	-0.19**	20%	1.91**	0.88**	**54%**	25%	25%	45%	**22%**
R3	0.65**	**82%**	-0.11	7%	3.12**	1.9**	**80%**	28%	26%	42%	**25%**
R4	0.58**	**91%**	-0.29**	43%	1.38**	0.5**	**64%**	**18%**	22%	51%	21%
R5	0.64**	**87%**	-0.23**	18%	1**	0.6**	44%	**25%**	30%	37%	28%
R6	0.6**	**90%**	-0.22**	37%	0.66**	0.10	**51%**	**16%**	**16%**	43%	18%
R7	0.87**	**87%**	0.06	4%	1.48**	1.29**	**67%**	21%	23%	24%	**18%**
R8	0.73**	**83%**	-0.04	1%	0.35	0.16	11%	24%	26%	25%	**21%**
R9	0.52**	**93%**	-0.35**	**85%**	0.1	-0.31**	**85%**	**12%**	13%	42%	24%
R10	0.53**	**83%**	-0.26**	43%	-0.27	-0.38	43%	23%	23%	**22%**	**22%**
R11	0.48**	**82%**	-0.3**	29%	2.48**	0.97**	**60%**	33%	31%	45%	**30%**
R12	0.69**	**88%**	-0.15*	15%	1.75**	1.05**	**55%**	21%	24%	47%	**20%**
R13	0.95**	**95%**	0.03	5%	0.56**	0.38**	48%	10%	14%	25%	**9%**
R14	0.9**	**78%**	0.16**	24%	0.59**	0.52**	44%	27%	19%	**15%**	24%
R15	0.73**	**79%**	0.01	0%	0.54**	0.3*	18%	29%	25%	27%	**23%**
R16	0.67**	**86%**	-0.12*	15%	0.88**	0.33	30%	20%	21%	42%	**18%**
R17	0.66**	**85%**	-0.1*	12%	1.38**	0.56**	41%	21%	18%	40%	**17%**
R18	0.59**	**98%**	-0.59**	**77%**	1.44**	0.73**	**97%**	**5%**	11%	50%	27%
R19	0.78**	**87%**	-0.06	6%	0.33	0.13	8%	16%	**14%**	22%	15%
R20	0.78**	**94%**	-0.12**	25%	0.26	0.04	28%	**12%**	14%	23%	**12%**
R21	0.63**	**92%**	-0.26**	**53%**	0.38	-0.04	**54%**	13%	**10%**	40%	17%
R22	1.33**	**58%**	0.34**	**66%**	-0.6	-0.02	**70%**	28%	**22%**	96%	38%
**: Significant at 5% level. **: Significant at 1% level.*											

Table 18: Estimation results of loss aversion models for BB retailers

	Model1		Model2		Model3			Average Absolute % Deviation			
	λ	**R^2**	**λ**	**R^2**	**λ_1**	**λ_2**	**R^2**	**Model1**	**Model2**	**Model3**	**Q*(w.b)**
R1	0.74**	**74%**	0.04	0%	-0.70	-0.43	8%	40%	31%	33%	**28%**
R2	0.9**	**74%**	0.35**	33%	0.35	0.71	34%	44%	32%	**2%**	39%
R3	1.18**	**79%**	0.51**	36%	-0.86**	-0.61*	**64%**	35%	**21%**	42%	44%
R4	0.7**	**88%**	-0.22**	23%	-0.31	-0.49*	27%	21%	26%	20%	**19%**
R5	0.91**	**89%**	0.13	8%	0.32	0.49	15%	22%	**17%**	25%	20%
R6	0.78**	**90%**	-0.09	10%	0.33	0.13	17%	16%	14%	18%	**12%**
R7	1.33**	**87%**	0.7**	**72%**	-0.15**	0.43**	**83%**	**12%**	46%	53%	27%
R8	0.92**	**92%**	0.04	0%	-0.19	-0.3	8%	20%	16%	17%	18%
R9	0.66**	**91%**	-0.28**	**85%**	-0.22**	-0.38**	**89%**	**10%**	43%	17%	19%

R10	0.84**	**83%**	0.15	6%	0.43	0.63	15%	33%	**23%**	34%	29%
R11	0.59**	**93%**	-0.2**	**63%**	0.46**	-0.06	**72%**	**9%**	17%	40%	17%
R12	0.5**	**88%**	-0.36**	**59%**	0.32	-0.23	**61%**	**20%**	28%	47%	25%
R13	1.11**	**86%**	0.42**	36%	-0.18*	0.23	47%	28%	**21%**	54%	33%
R14	0.82**	**94%**	-0.01	0%	0.49**	0.19**	**51%**	13%	12%	25%	**11%**
R15	0.6**	**96%**	-0.32**	50%	0.06	-0.27*	**50%**	**14%**	31%	30%	19%
R16	0.78**	**85%**	-0.06	1%	-0.12	-0.19	3%	29%	**27%**	29%	28%
R17	0.88**	**88%**	0.21	5%	0.17	0.57*	11%	**26%**	40%	35%	23%
R18	0.63**	**86%**	-0.2**	18%	-0.4**	-0.39**	**57%**	20%	27%	**16%**	20%
R19	0.89**	**90%**	-0.1*	10%	-0.34**	-0.4**	34%	12%	25%	16%	**11%**
R20	0.96**	**77%**	0.14	12%	-0.11	0.05	14%	27%	**22%**	31%	25%
R21	0.99**	**92%**	0.15**	18%	0.19	0.36**	26%	**15%**	17%	21%	**15%**
R22	0.67**	**88%**	-0.19**	31%	0.14	-0.1	32%	18%	**16%**	24%	17%

**: Significant at 5% level.   **: Significant at 1% level.*

Table 19: Estimation results of loss aversion models for RS retailers

	Model1			Model2		Model3			Average Absolute % Deviation			
	λ1	λ2	R^2	λ	R^2	λ_1	λ_2	R^2	Model1	Model2	Model3	Q*(w,r)
R1	1.71**	0.86**	**95%**	0.36**	27%	0.03	0.41	27%	28%	**19%**	45%	26%
R2	1.72**	0.92**	**94%**	0.18*	12%	0.03	0.22*	13%	37%	**14%**	20%	**14%**
R3	1.59**	0.77**	**95%**	0.17**	22%	-0.82**	-0.24**	**59%**	28%	29%	95%	29%
R4	1.12**	0.92**	**95%**	0.01	0%	0.2*	0.15	15%	61%	**11%**	12%	12%
R5	1.97**	0.47**	**95%**	-0.33**	**69%**	1.07**	0.06	74%	**17%**	18%	49%	23%
R6	1.45**	0.9**	**96%**	0.23**	30%	0.06	0.29**	31%	32%	**17%**	33%	22%
R7	1.74**	0.79**	**98%**	0.12	3%	1.41**	2.26**	44%	33%	**29%**	56%	**29%**
R8	1.34**	0.85**	**96%**	-0.17	5%	-0.36**	-0.57**	35%	36%	**27%**	**27%**	**27%**
R9	2**	0.55**	**93%**	-0.13**	25%	-2.24*	-1.29*	33%	20%	20%	70%	**17%**
R10	1.27**	0.57**	**98%**	-0.31**	**79%**	0.24	-0.25**	**81%**	38%	**17%**	47%	26%
R11	1.33**	0.69**	**93%**	0.14*	14%	-0.19	0.04	15%	56%	**24%**	45%	30%
R12	1.59**	0.7**	**94%**	0.07	10%	-0.13	0	11%	26%	19%	27%	**16%**
R13	1.73**	0.68**	**93%**	0.03	1%	0.13	0.11	2%	27%	23%	**21%**	25%
R14	1.21**	0.82**	**97%**	-0.11	9%	0.27	0.07	13%	48%	**15%**	22%	**15%**
R15	1.48**	0.9**	**95%**	0.11	5%	0.16*	0.36*	15%	29%	19%	26%	**16%**
R16	1.85**	0.71**	**91%**	0.22	14%	0.49	0.69	24%	39%	**27%**	38%	35%
R17	1.92**	0.77**	**96%**	0.08	5%	0.67**	0.74**	24%	22%	25%	25%	**19%**
R18	1.87**	0.78**	**96%**	0.18**	23%	0.51**	0.58**	40%	26%	21%	**19%**	24%
R19	1.54**	0.71**	**96%**	0.2	9%	-0.34	-0.07	10%	57%	**40%**	63%	50%
R20	2.14**	0.84**	**97%**	0.15**	19%	0.2	0.43	23%	17%	14%	27%	**13%**
R21	2.21**	0.64**	**97%**	-0.02	0%	0.71*	0.57	10%	16%	21%	25%	**15%**
R22	2.11**	0.77**	**96%**	0.16*	13%	-0.69**	-0.61**	39%	**18%**	24%	24%	20%

**: Significant at 5% level.   **: Significant at 1% level.*

4.5 Retailer's Inventory Error Models

For a newsvendor retailer, having unsold products and having stock-outs are both undesirable outcomes. In addition to the standard underage and overage costs, these outcomes may have additional psychological costs to the retailer. To capture such psychological costs, Schweitzer and Cachon (2000) mention minimizing ex-post inventory error model as a candidate to explain newsvendor behavior. While Schweitzer and Cachon don't differentiate leftovers and lost sales in their utility function, Ho et al. (2010) develop a model with different psychological costs associated with leftovers and lost sales.

In this section we study three utility functions related to the inventory errors. 1) Waste (unsold product) aversion model, 2) Stock-out (lost sales) aversion model and 3) Minimizing ex-post inventory error model which includes both waste and stock-out aversion.

4.5.1 Waste Aversion Model

In this model, the retailer has a psychological cost associated with leftover products (i.e., a psychological overage cost) in addition to the standard costs of underage and overage. We define the utility function as $U(\pi \mid Q) = \pi - \delta(\text{Leftovers}) = \pi - \delta(Q-D)^+$. For a waste-averse retailer, the psychological cost of overage coefficient δ will be positive. Thus, a waste-averse retailer will place a lower order quantity than the newsvendor quantity. The order quantity to maximize the retailer's utility function is derived as $Q^w = F^{-1}\left(\dfrac{p-r-w}{p-r-b+\delta}\right)$.

Proof:

Rearranging the expected utility function, we get:

$$
\begin{aligned}
E[U(\pi \mid Q)] &= E[\pi \mid Q] - \delta E(Q-D)^+ = (p-r-w)Q - (p-r-b)E(Q-D)^+ - \delta E(Q-D)^+ \\
&= (p-r-w)Q - (p-r-b-\delta)E(Q-D)^+ \\
&= (p-r-w)Q - (p-r-b-\delta)\int_0^Q (Q-x)\,dF(x)
\end{aligned}
$$

The proof follows as the standard newsvendor solution. ■

4.5.2 Stock-out Aversion Model

In this model, the retailer has a psychological cost associated with lost sales (i.e., a psychological underage cost) in addition to the standard costs of underage and overage. We define the utility function as $U(\pi \mid Q) = \pi - \delta(\text{Lost Sales}) = \pi - \delta(D-Q)^{+}$. For a stock-out averse retailer, the psychological cost of underage δ will be positive. Thus, a stock-out averse retailer will place a higher order quantity than the newsvendor quantity. The order quantity to maximize the retailer's utility function is derived as $Q^{so} = F^{-1}\left(\dfrac{p-r-w+\delta}{p-r-b+\delta}\right)$.

Proof:

Using the relation *Lost Sales = Demand – Sales= Demand – (Order Quantity-Leftovers)* we can write the expected utility of the retailer as:

$$
\begin{aligned}
E[U(\pi \mid Q)] &= E[\pi \mid Q] - \delta E(D-Q)^{+} \\
&= (p-r-w)Q - (p-r-b)E(Q-D)^{+} - \delta\left[E(D) - Q + E(Q-D)^{+}\right] \\
&= (p-r-w+\delta)Q - (p-r-b+\delta)E(Q-D)^{+} - \delta E(D)
\end{aligned}
$$

The proof follows as the standard newsvendor solution.■

4.5.3 Minimizing Ex-post Inventory Error Model

This model is a combination of the aforementioned waste-aversion and stock-out-aversion models. Here we assume that the retailer has psychological costs of both overage and underage in addition to the standard underage and overage costs. We define the utility function as

$$U(\pi \mid Q) = \pi - \delta_1(\text{Leftovers}) - \delta_2(\text{Lost Sales}) = \pi - \delta_1(Q-D)^{+} - \delta_2(D-Q)^{+}.$$

The order quantity to maximize the expected utility function of the retailer is derived as $Q^{inv} = F^{-1}\left(\dfrac{p-r-w+\delta_2}{p-r-b+\delta_1+\delta_2}\right)$.

Proof:

Rearranging the terms in the expected utility function, we get:

$$E[U(\pi \mid Q)] = E[\pi \mid Q] - \delta_1 E(Q-D)^+ - \delta_2 E(D-Q)^+$$
$$= (p-r-w+\delta_2)Q - (p-r-b+\delta_1+\delta_2)E(Q-D)^+ - \delta_2 E(D)$$

The proof follows as the standard newsvendor solution. ■

4.5.4 Model Estimation

Similar to earlier models, we linearize the optimal order quantity equations and run a regression analysis over the accepted contracts. Below are the resulting linear regression equations:

Waste aversion model: $\frac{Q-50}{100}(p-r-b)-(p-r-w) = \delta \frac{50-Q}{100} + \varepsilon$

Stock-out aversion model: $\frac{Q-50}{100}(p-r-b)-(p-r-w) = \delta \frac{150-Q}{100} + \varepsilon$

Minimizing ex-post inventory error model:

$$\frac{Q-50}{100}(p-r-b)-(p-r-w) = \delta_1 \frac{50-Q}{100} + \delta_2 \frac{150-Q}{100} + \varepsilon$$

Table 20 to Table 22 present the estimated δ coefficients, R^2 values and the average absolute percent deviation of the resulting model predictions from the actual order decisions. For most of the retailers, the models turn out to be significant, and provide a better prediction than Q*.

First, we observe that the minimizing ex-post inventory error model is the best of the three models. This is expected as the other two models are special cases of this model. In addition, the model provides a significant fit for almost all retailers.

Under the WSP contract, the waste aversion model turns out to be stronger than the stock-out aversion model. This may be due to the whole inventory risk being undertaken by the retailers, and the retailers being offered mostly low profit margin contracts. Under the BB contract, where some portion of the left-over inventory risk is being undertaken by the manufacturer, the stock-out aversion model is observed to be stronger than the waste-aversion model, as expected. Yet under the RS contract, the waste-aversion model seems to be stronger than the stock-out aversion model. This might be explained by the fact that manufacturers offer higher than theoretical-optimal wholesale prices, and therefore the cost of overage of the retailers is higher than their cost of underage.

Overall, for most of the retailers under all three contracts, at least one of these three inventory-error models is seen to provide a better prediction than the standard newsvendor model. Even though the minimizing ex-post inventory error model is the strongest of the three, the best prediction accuracy is obtained by the stock-out aversion model. Recall that the regression model minimizes the sum of squared deviations while we measure accuracy with absolute deviation.

Table 20: Estimation results of inventory error models for WSP retailers

	1. Waste Aversion		**2. Stock-out Aversion**		**3. Minimizing Inv. Error**			**Avg. Absolute % Deviation**			
					Waste	**Stock -out**					
	δ1	**R^2**	**δ2**	**R^2**	**δ1**	**δ2**	**R^2**	**Waste**	**Stock -out**	**Inv. Error**	**Q*(w)**
R 1	0.74	1%	-2.69**	**83%**	-3.32**	-4.05**	**98%**	31%	**8%**	10%	33%
R.2	-4.82**	**71%**	0.18	0%	-3.52**	-6.7**	**98%**	25%	**22%**	44%	**22%**
R.3	-4.14**	**56%**	-0.83	3%	-3.95**	-6.39**	**100%**	29%	26%	42%	**25%**
R.4	-5**	**83%**	1.81	7%	-3.25**	-6.64**	**98%**	18%	**1%**	49%	21%
R.5	-4.26**	**67%**	-1.02	2%	-4.24**	-5.54**	**96%**	**25%**	29%	39%	28%
R.6	-4.77**	**80%**	1.56	9%	-2.43**	-6.34**	**93%**	16%	**15%**	46%	18%
R.7	-1.52*	12%	-1.67**	22%	-4.4**	-5.09**	**98%**	21%	**17%**	25%	18%
R.8	-3.2**	39%	-0.82	4%	-3.11**	-5.45**	**83%**	24%	**21%**	37%	**21%**
R.9	-5.82**	**93%**	5.53**	**66%**	-0.95	-6.56**	**93%**	12%	**8%**	50%	24%
R.10	-5.62**	**80%**	1.91*	10%	-3.6**	-8.26**	**98%**	23%	**19%**	46%	22%
R.11	-6.21**	**84%**	-0.17	0%	-3.34**	-7.27**	**99%**	33%	**30%**	45%	**30%**
R.12	-3.69**	**60%**	0.09	0%	-3.96**	-6.15**	**98%**	21%	**20%**	46%	**20%**
R.13	-0.56	4%	-0.64*	13%	-2.86**	-4.27**	**89%**	10%	**7%**	28%	9%
R.14	-1.14	4%	-2.18**	46%	-3.61**	-4.82**	**92%**	27%	**18%**	22%	24%
R.15	-3.26**	35%	-1.28*	12%	-3.13**	-5.55**	**88%**	29%	**21%**	31%	23%
R.16	-3.95**	**59%**	0.3	1%	-3.24**	-6.7**	**95%**	20%	**18%**	42%	**18%**
R.17	-4.05**	**59%**	0.19	0%	-3.1**	-6.87**	**97%**	21%	**17%**	40%	**17%**
R.18	-4.96**	**97%**	6.05*	16%	-3.1**	-5.52**	**100%**	**5%**	17%	50%	27%
R.19	-2.61**	35%	-0.03	0%	-4**	-6.94**	**94%**	16%	**14%**	42%	15%
R.20	-2.6**	**53%**	0.93	9%	-3.57**	-5.99**	**89%**	**12%**	**12%**	46%	**12%**
R.21	-4.39**	**80%**	2.58**	23%	-3.74**	-7.1**	**97%**	13%	**10%**	49%	17%
R.22	3.99	8%	-3.29**	**82%**	-4.42**	-6.97**	**96%**	28%	20%	**19%**	38%

**: Significant at 5% level.   **: Significant at 1% level.*

Table 21: Estimation results of inventory error models for BB retailers

	1. Waste Aversion		2. Stock-out Aversion		3. Minimizing Inv. Error			Avg. Absolute % Deviation			
	δ1	R^2	δ2	R^2	Waste δ1	Stock-out δ2	R^2	Waste	Stock-out	Inv. Error	Q*(w)
R.1	-1.59**	19%	-1.55**	20%	-3.4**	-3.58**	**87%**	40%	**26%**	34%	28%
R.2	-0.56	2	-1.76**	**56%**	-2.54**	-2.59**	**91%**	45%	**18%**	41%	39%
R.3	1.67	9%	-2.95**	**82%**	-3.82**	-2.46**	**94%**	32%	**24%**	64%	44%
R.4	-1.76**	49%	0.31	1%	-2.64**	-3.14**	**80%**	22%	**19%**	41%	**19%**
R.5	-0.66	6%	-1.47**	30%	-3.15**	-2.65**	**88%**	22%	**17%**	32%	20%
R.6	-2.45**	42%	0.13	0%	-3.35**	-5.22**	**87%**	16%	**12%**	42%	**12%**
R.7	1.53**	25%	-3.07**	**90%**	-3.39**	-0.47*	**92%**	**14%**	15%	54%	27%
R.8	-0.44	5%	-1.56**	26%	-3.34**	-1.75**	**77%**	20%	16%	35%	**18%**
R.9	-1.6**	**70%**	1.66**	**67%**	0.59	-1.07	**71%**	13%	**7%**	39%	19%
R.10	-1.35*	15%	-2.08**	38%	-3.36**	-2.89**	**95%**	35%	**22%**	27%	29%
R.11	-3.4**	**80%**	1.75**	41%	-0.39	-3.82**	**81%**	12%	**10%**	40%	17%
R.12	-4.82**	**85%**	2.79**	19%	-1.73**	-5.73**	**90%**	25%	**20%**	53%	25%
R.13	0.88	6%	-2.63**	**76%**	-2.83**	-0.52	**78%**	27%	**16%**	57%	33%
R.14	-1.55**	41%	-0.1	1%	-1.14**	-2.93**	**84%**	13%	**11%**	29%	11%
R.15	-2.24**	**74%**	1.64*	17%	-1.05*	-2.69**	**78%**	20%	**17%**	50%	19%
R.16	-1.17**	26%	-1.1*	15%	-2.16**	-1.94**	**74%**	29%	30%	39%	**28%**
R.17	-0.39	10%	-1.67**	**57%**	-2.07**	-0.73**	**89%**	26%	**22%**	44%	23%
R.18	-1.73**	37%	0.59	5%	-1.31*	-2.73**	**48%**	19%	**17%**	37%	20%
R.19	-0.57*	11%	-0.01	0%	-2.4**	-2.44**	**48%**	13%	**11%**	31%	**11%**
R.20	-0.23	0%	-1.6**	42%	-2.45**	-2.08**	**68%**	26%	**21%**	31%	25%
R.21	-0.05	0%	-1.33**	39%	-2.61**	-1.85**	**80%**	15%	**14%**	32%	15%
R.22	-2.94**	**60%**	0.91	6%	-2.78**	-5.04**	**85%**	18%	**14%**	45%	17%

**: Significant at 5% level.   **: Significant at 1% level.*

Table 22: Estimation results of inventory error models for RS retailers

	1. Waste Aversion		2. Stock-out Aversion		3. Minimizing Inv. Error			Avg, Absolute % Deviation			
	δ1	R^2	δ2	R^2	Waste δ1	Stock-out δ2	R^2	Waste	Stock-out	Inv. Error	Q*(w)
R.1	0.12	0%	-2.71**	**70%**	-3.7**	-1.74**	**93%**	26%	**19%**	53%	26%
R.2	-0.21	1%	-2.17**	**52%**	-2.74**	-0.85**	**71%**	14%	**11%**	23%	14%
R.3	0.36	0%	-1.61**	47%	-2.89**	-5.52**	**81%**	28%	**22%**	60%	29%
R.4	-0.35	4%	-0.65*	16%	-1.43**	-1.38**	**51%**	**10%**	21%	20%	12%
R.5	-6.1**	**93%**	4.02**	36%	-2.41**	-7.86**	**98%**	17%	**15%**	51%	23%
R.6	0.00	0%	-1.79**	**51%**	-2.54**	-1.6**	**72%**	22%	**15%**	35%	22%
R.7	-1.49**	22%	-2.49**	38%	-3.84**	-2.68**	**98%**	35%	**28%**	50%	29%
R.8	-0.88*	17%	-1.16*	14%	-2.63**	-1.88**	**68%**	30%	31%	30%	**27%**
R.9	-2.67**	52%	0.62	5%	-3.49**	-6.46**	**99%**	16%	**15%**	52%	17%
R.10	-3.76**	**75%**	2.67**	**55%**	0.37	-3.41**	**76%**	23%	**11%**	42%	26%
R.11	-2.35**	20%	-1.12**	23%	-2.22**	-4.8**	**86%**	58%	**21%**	85%	30%
R.12	-0.46	1%	-0.82**	21%	-2.42**	-4.62**	**75%**	17%	**14%**	40%	16%

R.13	-2.04**	22%	-1.29**	19%	-2.73**	-4.1**	**83%**	32%	**20%**	35%	25%
R.14	-1.08**	30%	-0.23	2%	-1.46**	-2.04**	**67%**	19%	17%	43%	**15%**
R.15	-0.43	6%	-1.69**	37%	-2.71**	-1.31**	**77%**	**16%**	**16%**	28%	**16%**
R.16	-2.04	17%	-2.61**	44%	-3.69**	-3.74**	**94%**	49%	**23%**	28%	35%
R.17	-0.9	8%	-1.47**	25%	-3.69**	-3.62**	**93%**	20%	**19%**	21%	**19%**
R.18	-0.6	2%	-1.81**	46%	-3.12**	-3.34**	**92%**	25%	20%	**16%**	24%
R.19	-2.4*	21%	-2.37**	**50%**	-3.02**	-3.68**	**95%**	73%	**26%**	49%	50%
R.20	0.22	1%	-1.66**	39%	-4.27**	-2.99**	**90%**	13%	**11%**	45%	13%
R.21	-1.38**	20%	-0.59	5%	-3.78**	-4.59**	**96%**	17%	**15%**	45%	**15%**
R.22	0.34	1%	-1.94**	42%	-4.76**	-4.27**	**91%**	19%	**17%**	40%	20%

: Significant at 5% level.* *: Significant at 1% level.*

4.6 Retailer's Social Preference Models

Standard theory assumes decision makers are rational profit maximizers who are only interested in their own profit. However, where there is human interaction, decisions of the individuals may be affected by their relationship with other individuals or by their attitude towards other individuals. As supply chain contracting relations involve manufacturer-retailer interaction, it is natural to expect social preferences of individuals to affect their ordering or contracting behavior. In the behavioral operations literature, Pavlov and Katok (2011), Cui et al. (2007), and Loch and Wu (2008) study social preferences in supply chain relationships. However these studies base their analysis on a linear deterministic demand setting where, unlike our study, the effect bounded rationality is restricted because of the lack of uncertainty in demand realization. Hence as bounded rationality is somewhat suppressed, social preferences have stronger effect on pricing and ordering decisions.

In this section, we study fairness, status seeking, group identity and a general model of social preference. Similar models have been used in Wu (2006), Loch and Wu (2008), Pavlov and Katok (2011), and Cui and Malluci (2016).

4.6.1 Fairness Model

In this model, we assume that the retailer has a concern for equitable allocation of supply chain profit. Hence, the retailer's utility function is increasing in both her own profit and also in the manufacturer's profit. We assume that the retailer has a utility function of the form:

$$U(\pi_R \mid Q) = (\pi_R \mid \mathrm{Q}) + \delta(\pi_M \mid \mathrm{Q})$$

A positive δ signals a fairness-concerned retailer whose utility increases with manufacturer's profit. The parameter δ taking a value close to 1 indicates that retailer's fairness ideal is close to equality. A negative δ may be associated with rivalry behavior. The order quantity to maximize the expected utility function of a retailer is derived as:

$$Q^f = F^{-1}\left(\frac{p-r-w+\delta(\mathrm{w}-\mathrm{c}+\mathrm{r})}{p-r-b+\delta(b+r)}\right).$$

<u>**Proof:**</u>

Rearranging the terms in the expected utility function we get:

$$\begin{aligned}
E[U(\pi_R \mid Q)] &= E[\pi_R \mid \mathrm{Q}] + \delta E(\pi_M \mid \mathrm{Q}) \\
&= (p-r)E(\mathrm{Sales} \mid Q) + bE(\mathrm{Leftovers} \mid Q) - wQ \\
&\quad + \delta\left(wQ + rE(\mathrm{sales} \mid Q) - bE(\mathrm{leftover\ inv.} \mid Q) - cQ\right) \\
&= \left[p - r(1-\delta)\right]E(\mathrm{Sales} \mid Q) + b(1-\delta)E(\mathrm{Leftovers} \mid Q) - \left[w(1-\delta) + c\delta\right]Q
\end{aligned}$$

The proof follows as the standard newsvendor solution. ■

4.6.2 Status Seeking Model

In this model we assume that the retailer is a competitive individual whose utility increases if she earns more than the manufacturer. In individuals with strong competitiveness, utility from gains increase if they gain more than other individuals (Loch and Wu 2008). To capture this, the retailer's utility function is modeled as:

$$U(\pi_R \mid Q) = (\pi_R \mid \mathrm{Q}) + \delta(\pi_R - \pi_M \mid \mathrm{Q}).$$

The order quantity to maximize the expected utility function becomes:

$$Q^{ss} = F^{-1}\left(\frac{p-r-w+\delta(\mathrm{p}-2r-2w+c)}{p-r-b+\delta(p-2b-2r)}\right).$$

<u>**Proof:**</u>

Rearranging the terms of the expected utility function we obtain:

$$
\begin{aligned}
E[U(\pi_R \mid Q)] &= E[\pi_R \mid \mathrm{Q}] + \delta E[\pi_R - \pi_M \mid \mathrm{Q}] \\
&= (p-r)E(\mathrm{Sales} \mid Q) + bE(\mathrm{Leftovers} \mid Q) - wQ \\
&\quad + \delta\left[(p-2r)E(\mathrm{Sales} \mid Q) + 2bE(\mathrm{Leftovers} \mid Q) - (2w-c)Q\right] \\
&= \left[p - r + \delta(p-2r)\right]E(\mathrm{Sales} \mid Q) + b(1+2\delta)E(\mathrm{Leftovers} \mid Q) - \left[w + \delta(2w-c)\right]Q
\end{aligned}
$$

The proof follows as the standard newsvendor solution.■

4.6.3 Group Identity Model

Here we assume that the retailer is concerned with the total supply chain profit, in addition to her own profit (Loch and Wu 2008). The retailer's utility function become:

$U(\pi_R \mid Q) = (\pi_R \mid \mathrm{Q}) + \delta(\pi_R + \pi_M \mid \mathrm{Q})$.

The order quantity to maximize the expected utility function is

$$
Q^{gi} = F^{-1}\left(\frac{p - r - w + \delta(\mathrm{p} - c)}{p - r - b + \delta(p)}\right).
$$

<u>Proof:</u>

Rearranging the terms in the expected utility function we obtain:

$$
\begin{aligned}
E[U(\pi_R \mid Q)] &= E[\pi_R \mid \mathrm{Q}] + \delta E[\pi_R + \pi_M \mid \mathrm{Q}] \\
&= (p-r)E(\mathrm{Sales} \mid Q) + bE(\mathrm{Leftovers} \mid Q) - wQ + \delta\left(pE(\mathrm{sales} \mid Q) - cQ\right) \\
&= \left[p - r + \delta p\right]E(\mathrm{Sales} \mid Q) + bE(\mathrm{Leftovers} \mid Q) - (w-c)Q
\end{aligned}
$$

The proof follows as the standard newsvendor solution.■

4.6.4 Model Estimation

We re-arrange the optimal order quantity equations and obtain the linear regression models as

- **Fairness model:**

  $$(Q-50)(\mathrm{p}-\mathrm{r}-\mathrm{b}) - 100(\mathrm{p}-\mathrm{r}-w) = \delta\left[100(w-c+r) - (Q-50)(b+r)\right]$$

- **Status seeking model:**

  $$(Q-50)(\mathrm{p}-\mathrm{r}-\mathrm{b}) - 100(\mathrm{p}-\mathrm{r}-w) = +\delta\left[100(p-2b-2r) - (Q-50)(p-2r-2w+c)\right]$$

- **Group identity model:**

$$(Q-50)(\mathrm{p}-\mathrm{r}-\mathrm{b})-100(\mathrm{p}-\mathrm{r}-w)=+\delta\left[100(\mathrm{p}-\mathrm{c})-(Q-50)(p)\right]+\varepsilon$$

Estimation results are shown in Table 23 to Table 25. Under the WSP contract, the fairness model is found to be significant for about half of the retailers. The estimated δ coefficient is mostly positive, indicating that most of these retailers are concerned with fairness. The estimated values of δ coefficients for these retailers are around 0.4, which can be interpreted as the retailers valuing their own profit more than twice as much as the profit of the manufacturer. Under the BB and RS contracts, the fairness model is not as strong and the estimated coefficients are mostly negative, indicating the retailer is in a competitive behavior. Under these two contracts, due to the more flexible two-parameter design, the manufacturer receives a higher profit share than the WSP contract, and this may be the cause of the retailer's negative δ coefficients.

The status-seeking model yields mostly positive δ coefficients for WSP and BB retailers, and mostly negative δ coefficients for RS retailers. This can be interpreted as WPS and BB retailers being mostly status-seeking decision makers while RS retailers are mostly equity concerned decision makers. Additionally, the R^2 values obtained from this model are below 50% for most retailers indicating the model does not fit to the data very well.

For the group identity model, majority of the δ coefficients are negative for all three contract types indicating status-seeking behavior. The model is significant for majority of the retailers, especially under RS contract, and the R^2 values are higher compared to fairness and status-seeking models.

The rightmost four columns of Table 23-Table 25 provide the average absolute percent deviation of the model-predicted order decisions from the actual experiment decisions. Compared to the standard newsvendor model's prediction accuracy, presented on the rightmost column, these three models obtain at least a slightly better prediction for the majority of the retailers. Hence, we conclude that the fairness models explain the data better than the standard newsvendor model for most of the retailers.

Note that these three models are very similar to each other in the sense that in each one, the utility functions can be written as weighted sums of the retailer's and the manufacturer's expected profits. However, due to the model structures, the ratio of these weights are subject to different restrictions. Specifically, the ratio of the weight of the retailer's profit to the weight of the manufacturer's profit in the fairness model is $1/\delta$, in the status seeking model $(1+\delta)/(-\delta)$ and in the group identity model $(1+\delta)/\delta$. These differences result in the three models to have different interpretation and different performances.

Table 23: Estimation results of social preference models for WSP retailers

	Fairness Concerns		**Status Seeking**		**Group Identity**		**Avg. Absolute % Deviation**			
	δ	**R^2**	**δ**	**R^2**	**δ**	**R^2**	**Fairness**	**Status**	**Group**	**Q*(w)**
R1	-0.44**	**77%**	-0.17**	**67%**	-0.33**	**89%**	8%	46%	**6%**	33%
R2	0.29**	18%	0.13**	23%	-0.21	8%	23%	23%	25%	**22%**
R3	0.15	5%	0.07	8%	-0.29**	20%	26%	25%	25%	25%
R4	0.45**	39%	0.19**	44%	-0.12	2%	**18%**	24%	24%	21%
R5	0.35*	12%	0.14*	16%	-0.51**	33%	**28%**	**28%**	37%	**28%**
R6	0.33**	33%	0.15**	41%	0	0%	43%	21%	**18%**	**18%**
R7	-0.15	7%	-0.04	4%	-0.31**	44%	17%	19%	**16%**	18%
R8	0.04	1%	0.03	2%	-0.2**	17%	**21%**	**21%**	**21%**	**21%**
R9	0.53**	**84%**	0.23**	**87%**	0.71**	43%	**8%**	33%	9%	24%
R10	0.41**	43%	0.17**	45%	0.04	0%	**19%**	26%	21%	22%
R11	0.46**	26%	0.21**	33%	-0.39*	16%	31%	31%	39%	**30%**
R12	0.23*	13%	0.1**	16%	-0.20	9%	**20%**	**20%**	22%	**20%**
R13	-0.07	7%	-0.02	4%	-0.11**	23%	**7%**	9%	8%	9%
R14	-0.28**	28%	-0.1**	22%	-0.31**	**61%**	19%	28%	**17%**	24%
R15	-0.04	1%	0.00	0%	-0.24**	27%	**22%**	23%	**22%**	23%
R16	0.18*	13%	0.08**	17%	-0.08	2%	19%	19%	19%	**18%**
R17	0.15*	10%	0.08*	15%	-0.08	2%	18%	**17%**	**17%**	**17%**
R18	1.06**	**68%**	0.38**	**79%**	-1.11**	39%	**14%**	30%	20%	27%
R19	0.10	6%	0.04	7%	-0.08	3%	**14%**	15%	15%	15%
R20	0.2**	24%	0.08**	26%	0.05	1%	**12%**	13%	**12%**	**12%**
R21	0.42**	**52%**	0.16**	**53%**	0.16	4%	**10%**	21%	14%	17%
R22	-0.55**	**64%**	-0.21**	**68%**	-0.4**	**86%**	22%	55%	**20%**	38%

**: Significant at 5% level.   **: Significant at 1% level.*

Table 24: Estimation results of social preference models for BB retailers

	Fairness Concerns		**Status Seeking**		**Group Identity**		**Avg. Absolute % Deviation**			
	δ	**R^2**	**δ**	**R^2**	**δ**	**R^2**	**Fairness**	**Status**	**Group**	**Q*(w)**

R1	-0.17	9%	0.14*	10%	-0.3**	46%	**26%**	**26%**	29%	28%
R2	-0.27**	**53%**	0.14	7%	-0.24**	**72%**	**17%**	30%	**17%**	39%
R3	-0.56**	**76%**	-0.19**	19%	-0.38**	**91%**	**25%**	49%	**25%**	44%
R4	0.11	4%	0.29**	45%	**-0.15**	9%	**18%**	25%	23%	19%
R5	-0.24**	27%	0.01	0%	-0.27**	57%	**16%**	19%	17%	20%
R6	0.12	7%	0.07*	12%	-0.09	4%	13%	**12%**	13%	**12%**
R7	-0.58**	**67%**	-0.28**	24%	-0.41**	**88%**	30%	30%	34%	**27%**
R8	-0.34**	29%	0.15	10%	-0.35**	**66%**	21%	**18%**	28%	**18%**
R9	0.23**	**76%**	0.24**	23%	0.26**	**60%**	**6%**	40%	**6%**	19%
R10	-0.35**	35%	-0.02	0%	-0.34**	**68%**	22%	30%	20%	29%
R11	0.21**	**52%**	0.18**	**67%**	0.19**	27%	**11%**	30%	**11%**	17%
R12	0.44**	48%	0.29**	**71%**	0.07	1%	**20%**	35%	23%	25%
R13	-0.38**	**52%**	-0.17	12%	-0.32**	**76%**	24%	39%	**23%**	33%
R14	-0.01	0%	0.05	10%	-0.03	5%	11%	**10%**	**10%**	11%
R15	0.24**	23%	0.4**	**76%**	-0.02	0%	**17%**	46%	19%	19%
R16	-0.17*	16%	0.15*	14%	-0.23**	45%	28%	28%	**26%**	28%
R17	-0.29**	**81%**	0.14*	14%	-0.23**	**89%**	22%	27%	**21%**	23%
R18	0.13*	14%	0.25**	**62%**	0.01	0%	**16%**	35%	19%	20%
R19	0.04	1%	0.19*	14%	-0.06	3%	**11%**	20%	14%	**11%**
R20	-0.21**	30%	-0.11	10%	-0.23**	**56%**	**22%**	28%	24%	25%
R21	-0.23**	40%	-0.06	4%	-0.23**	**62%**	**13%**	16%	14%	15%
R22	0.22**	21%	0.15**	34%	-0.01	0%	**14%**	20%	17%	17%
*: Significant at 5% level. **: Significant at 1% level.										

Table 25: Estimation results of social preference models for RS retailers

	Fairness Concerns		Status Seeking		Group Identity		Avg. Absolute % Deviation			
	δ	R^2	δ	R^2	δ	R^2	Fairness	Status	Group	Q*(w)
R1	-0.58**	**75%**	-0.27**	22%	-0.4**	**90%**	**23%**	32%	25%	26%
R2	-0.33**	45%	-0.16**	21%	-0.3**	**71%**	**10%**	17%	**10%**	14%
R3	-0.19**	31%	-0.08*	10%	-0.21**	**54%**	**23%**	35%	**23%**	29%
R4	-0.09*	14%	-0.09*	16%	-0.12**	30%	21%	15%	30%	**12%**
R5	0.49**	**66%**	0.24**	**75%**	0.33*	11%	16%	33%	**15%**	23%
R6	-0.27**	41%	-0.1	7%	-0.26**	**66%**	**15%**	26%	16%	22%
R7	-0.52**	42%	0.12	2%	-0.43**	**78%**	**25%**	29%	28%	29%
R8	-0.18*	12%	0.19**	23%	-0.28**	46%	31%	31%	32%	**27%**
R9	0.14*	16%	0.08**	17%	0.02	0%	**15%**	18%	17%	17%
R10	0.3**	**68%**	0.28**	**74%**	0.32**	40%	11%	41%	**10%**	26%
R11	-0.11*	12%	0.08	8%	-0.16**	34%	23%	25%	**20%**	30%
R12	-0.09*	12%	-0.03	4%	-0.12**	28%	**14%**	17%	15%	16%
R13	-0.12	7%	-0.02	0%	-0.22**	35%	**21%**	25%	**21%**	25%
R14	-0.03	1%	0.06	4%	-0.09*	13%	17%	19%	23%	**15%**
R15	-0.34**	46%	-0.09	5%	-0.3**	**72%**	**14%**	17%	16%	16%
R16	-0.38*	29%	-0.08	3%	-0.38**	**68%**	23%	38%	**22%**	35%
R17	-0.2**	16%	-0.04	2%	-0.28**	**50%**	**18%**	19%	19%	19%
R18	-0.27**	38%	-0.1*	12%	-0.27**	**64%**	**19%**	27%	**19%**	24%
R19	-0.34**	39%	-0.01	0%	-0.32**	**69%**	29%	51%	**23%**	50%
R20	-0.31**	32%	-0.07	9%	-0.32**	**65%**	11%	14%	**10%**	13%

R21	-0.02	0%	0.03	2%	-0.17**	19%	**15%**	**15%**	17%	**15%**
R22	-0.31**	31%	-0.1*	13%	-0.32**	**64%**	**18%**	21%	19%	20%
**: Significant at 5% level. **: Significant at 1% level.*										

4.7 Manufacturer's Social Preference Models

Here we study social preference models of the manufacturer while assuming that the retailer conforms to the standard newsvendor solution. We only consider the wholesale price contract in this study. This is because one cannot obtain closed-form solutions for the manufacturer's optimal buyback and revenue sharing contract parameters, as discussed earlier.

4.7.1 Fairness Model

The manufacturer's utility function is $U(\pi_M \mid w) = (\pi_M \mid w) + \delta(\pi_R \mid w)$.

The wholesale price to maximize this function is calculated as

$$w^f = \frac{\left(U - \delta U - \frac{\delta}{2}\right)p}{(U-L+1)(2-\delta)} + \frac{c}{(2-\delta)}.$$

4.7.2 Status Seeking Model

The manufacturer's utility function is $U(\pi_M \mid w) = (\pi_M \mid w) + \delta(\pi_M - \pi_R \mid w)$.

The wholesale price to maximize this function is found as:

$$w^{ss} = \frac{\left((1-2\delta)U + \frac{\delta}{2}\right)p}{(U-L+1)(2+3\delta)} + \frac{c(1+\delta)}{(2+3\delta)}.$$

4.7.3 Group Identity Model

The manufacturer's utility function is $U(\pi_M \mid w) = (\pi_M \mid w) + \delta(\pi_M + \pi_R \mid w)$.

The wholesale price to maximize this function is found as

$$w^{gi} = \frac{\left(U - \frac{\delta}{2}\right)p}{(U - L + 1)(2 + \delta)} + \frac{c(1 + \delta)}{(2 + \delta)}.$$

4.7.4 Model Estimation

Rearranging the optimal wholesale price equations, we obtain the linear regression models as

- **Fairness model:**

$$Up + c(U - L + 1) - 2w(U - L + 1) = \delta\left((U + \frac{1}{2})p - w(U - L + 1) \right) + \varepsilon$$

- **Status seeking model:**

$$Up + c(U - L + 1) - 2w(U - L + 1) = \delta\left((3w - c)(U - L + 1) - 2Up - \frac{p}{2} \right) + \varepsilon$$

- **Group identity model:**

$$Up + c(U - L + 1) - 2w(U - L + 1) = \delta\left((w - c)(U - L + 1) + \frac{p}{2} \right) + \varepsilon$$

Estimation results are presented in Table 26. All three models are found to be significant for all retailers. In addition, the models fit the data quite well. The estimated wholesale prices $\hat{w}$ are shown in the table. On the right hand side of the table, the average absolute percentage deviations of these estimated wholesale prices from the actual experiment decisions are shown. On the right-most column, the average absolute percentage deviation of the theoretical optimal wholesale price from the experiment decisions is given. With these three social preference models, the manufacturer's pricing decisions are observed to be predicted with higher accuracy than the standard theory. The percent deviation of the model providing the highest prediction accuracy is printed in boldface.

As it is the case with retailer's social preference models, these three models are very similar to each other in the sense that in each of them, the utility functions can be written as weighted sums of the retailer's and the manufacturer's expected profits. This time, however, the performances difference between the three models are found to be relatively small. This may be because unlike the retailers, the manufacturers' analyses reveal strong fairness concerns and negative status-seeking.

Table 26: Estimation results of social preference models for WSP manufacturers

	Fairness Concerns			Status Seeking			Group Identity			Exp. Avg. w	Avg. Absolute % Deviation			
	δ	R^2	$\hat{w}$	δ	R^2	$\hat{w}$	δ	R^2	$\hat{w}$		Fairness	Status	Group	w*=10
M1	0.50	87%	7.96	-0.35	94%	7.79	0.75	54%	8.45	8.13	**12%**	13%	14%	27%
M2	0.55	97%	7.63	-0.36	99%	7.60	1.13	88%	7.76	7.68	8%	8%	**7%**	31%
M3	0.57	95%	7.48	-0.37	98%	7.40	1.12	75%	7.78	7.58	9%	**9%**	11%	34%
M4	0.56	97%	7.55	-0.36	99%	7.50	1.17	85%	7.72	7.60	9%	9%	**8%**	33%
M5	0.70	97%	6.40	-0.42	99%	6.32	1.89	72%	6.82	6.50	14%	14%	**15%**	60%
M6	0.52	91%	7.88	-0.35	96%	7.75	0.85	65%	8.25	8.00	**12%**	12%	13%	28%
M7	0.70	99%	6.47	-0.41	100%	6.44	2.12	89%	6.61	6.50	**9%**	**9%**	**9%**	55%
M8	0.59	90%	7.35	-0.38	96%	7.16	0.98	52%	8.02	7.55	12%	**11%**	17%	39%
M9	0.41	85%	8.54	-0.30	92%	8.40	0.56	59%	8.84	8.65	**11%**	**11%**	**11%**	17%
M10	0.45	97%	8.33	-0.31	99%	8.30	0.77	91%	8.39	8.35	**5%**	**5%**	6%	20%
M11	0.51	95%	7.95	-0.34	98%	7.89	0.93	82%	8.11	8.00	**7%**	**7%**	**7%**	26%
M12	0.62	99%	7.07	-0.39	99%	7.04	1.56	91%	7.18	7.10	**5%**	**5%**	6%	42%
M13	0.60	97%	7.24	-0.38	99%	7.18	1.35	82%	7.45	7.30	10%	**9%**	10%	39%
M14	0.57	92%	7.50	-0.37	97%	7.37	1.00	62%	7.97	7.65	14%	14%	**13%**	35%
M15	0.56	89%	7.57	-0.37	96%	7.39	0.91	54%	8.14	7.75	16%	16%	**16%**	36%
M16	0.51	95%	7.89	-0.34	98%	7.82	0.94	80%	8.08	7.95	8%	8%	**8%**	27%
M17	0.48	95%	8.10	-0.33	98%	8.05	0.85	84%	8.24	8.15	6%	**6%**	7%	24%
M18	0.69	95%	6.55	-0.41	98%	6.43	1.59	62%	7.15	6.70	14%	**12%**	19%	54%
M19	0.55	99%	7.60	-0.36	99%	7.58	1.19	93%	7.68	7.63	6%	6%	**6%**	32%
M20	0.59	97%	7.32	-0.37	99%	7.27	1.31	84%	7.51	7.38	9%	**9%**	10%	37%
M21	0.54	97%	7.71	-0.35	99%	7.68	1.09	88%	7.83	7.75	7%	7%	**6%**	30%
M22	0.54	94%	7.74	-0.35	97%	7.65	0.98	74%	8.01	7.83	9%	10%	**9%**	30%

All estimations are significant at 1% level.

4.8 Conclusion

In this chapter we develop and estimate various behavioral models based on different utility functions for the retailer and the manufacturer subjects of our supply chain experiments. For each model, we rearrange related nonlinear equations to make them linear, and estimate the models through linear regression analysis. Although these linear regression models are easier to interpret and yield relatively high R^2 values, their prediction accuracy for the retailer's behavioral models is not at desired levels. Thus, there is potential in estimating advanced nonlinear regression models such as maximum likelihood estimation. This offers an interesting extension to our study. Another future study might be developing constrained regression models given that both the order quantity and the pricing decisions are restricted in our experiment scenario.

4.9 References

Akbay, Ü., Kaya, M. (2016). Experiments on Supply Chain Contracting: Effects of contract type. Working paper.

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

Cui, T. H., Malluci P. (2016). Bounded Rationality and Fairness Ideals in Distribution Channels. Working Paper, to appear in Journal of Marketing Research.

Ho, T., Lim N., Cui H. T. (2010). Reference Dependence in Multilocation Newsvendor Models: A Structural Analysis. Management Science, 56, 1891-1910.

Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-292.

Kahneman, D., Tversky, A. (1984). Choices, Values, and Frames. American Psychologist 39 (4): 341–350

Loch, C. H., Wu Y. (2008). Social Preferences and Supply Chain Performance: An Experimental Study, Management Science, 54(11), 1835-1849.

Pavlov, V., Katok E., Fairness and Coordination Failures in Supply Chain Contracts, Working Paper, (2011).

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

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

Wu, Y. (2006). Social Preferences in Coordination and Delegation Decisions- Models and Empirical Tests, (Doctoral Dissertation), INSEAD.

Zhang, Y., K. Donohue, H. T. Cui. 2016. Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing. Management Science, 62(6), 1734–1754.

Chapter 5

5. POWER OF COMMITMENT AND IMPACT OF PRIMINT ON SUPPLY CHAIN RELATIONSHIPS

Experiments on the Power of Commitment and Fairness Priming in Supply Chain Decisions

Abstract

We conduct decision experiments to study the effects of commitment and fairness priming in a manufacturer-retailer supply chain. In each period, the manufacturer offers a wholesale price contract, and the retailer, who faces the newsvendor problem, reacts by choosing her order quantity to satisfy probabilistic demand. The retailer has the chance to reject the contract, in which case both firms receive zero profits. The firms are represented by human subjects who interact repeatedly throughout the experiment, capturing a long-run relationship. Our first study focuses on the power of commitment to a decision (wholesale price decision and quantity decision) for five periods. We find the committing firm to obtain strategic advantage over the other, and to increase its profit share.

Our second study explores what happens if the subjects are primed with an exogenously-given fair contract at the beginning of the interaction. We conjecture the relatively fair profit shares during these initial periods to act as an anchor that will affect the subjects' subsequent decisions. Contrary to our expectations, following fairness priming, the manufactures acted more aggressively and captured an even higher share of total profits while the retailers failed to counteract. To shed light on the findings of both studies, we present regression models that explain the retailer's contract rejection, underorder and overorder tendencies.

Keywords: Commitment, Priming, Fairness, Newsvendor model; Supply chain contracts; Behavioral operations; Experiments

5.1 Introduction

In this paper, we investigate two phenomena related to strategic decision making in a supply chain context: The power of commitment and the effect of priming. To this end, we consider a canonical supply chain model consisting of a manufacturer that produces-to-order, and a retailer that faces the standard newsvendor problem. The manufacturer offers a contract to the retailer, which the retailer may or may not accept. If the contract is accepted, the retailer determines her order quantity. Lastly, demand is realized and the payoffs of the firms are calculated. The two firms interact repeatedly over a number of periods, capturing a long-run relationship.

Based on this scenario, we present two experimental studies in which we conduct a number of computerized decision experiments with human subjects playing the roles of manufacturer and retailer. Similar scenarios have already been used in a handful of behavioral supply chain studies such as Keser and Paleologo (2003), Wu (2013), Cui et al. (2007).

Our first experimental study is concerned with the power of commitment to a decision. Commitment is a central theme of strategic decision theory. Contrary as it sounds, credibly limiting one's options in a multi-person game can bring advantages, as illustrated with the proverbial "burning the ships" strategy. In our experiments, commitment translates into not being able to change decision for five periods. We first study the effect of manufacturer's commitment to a contract (i.e., a wholesale price). Then we investigate what happens when both the manufacturer and the retailer simultaneously commit to their respective decisions. We find commitment to a wholesale price to provide the manufacturer with an advantage over the retailer. The retailer can mitigate this advantage when she too commits to her order quantity decision. These changes in the power balances are reflected in profit allocations.

Our second experimental study is concerned with the effects of priming the subjects with a fair contract anchor. In particular, in the first five periods of the interaction, the wholesale price is exogenously set to a value that provides almost equal expected profits to the two firms. This fair anchor may correspond to an "already existing" contract in practice. According to standard economic theory, the existence of such a contract should not affect the way subsequent contracts are offered and perceived. However, psychological and behavioral studies have repeatedly shown decision makers to be affected by priming through such information anchors. Anchoring and adjustments theory (Tversky and Kahneman, 1974) suggests that individuals evaluate options based on the readily available information that may be related to their previous experiences or expectations. Our study on priming yielded a rather unexpected result: Fairness priming did not lead to a more equitable sharing of profits. On the contrary, manufacturers were strongly primed against the fair anchor and retailers failed to counteract.

The rest of the paper is organized as follows: Section 5.2 summarizes the relevant literature. Section 5.3 presents the analytical background of the model and briefly summarizes our experimental design. Sections 5.4 and 5.5 focus on the power of commitment and fairness priming studies respectively. We summarize our findings and conclude in Section 5.6.

5.2 Related Literature

We base our experiments on a single-manufacturer single-retailer supply chain scenario where the retailer is faced with a newsvendor problem. The cost parameters of the retailer's problem are determined by the manufacturer's contract offer. The newsvendor problem has been the subject of numerous experimental studies since the pioneering study by Schweitzer and Cachon (2000). The authors observe the subjects' order quantity decisions to deviate systematically from the theoretical optimal, and propose a number of heuristics such as demand chasing to explain the observed subject behavior. Bendoly et al. (2015) provides a recent review of this literature.

In our first experimental study, we investigate the effect of power of commitment on manufacturer and retailer decisions. To this end, we restrict the manufacturer's and the retailer's decisions to stand effective for a five periods rather than the standard single period. The effect of standing decisions on newsvendor performance has been investigated by several studies. Standing order decisions is expected to eliminate the lawof-small-numbers bias by reducing the observed variance of the random demand. In their experiments, Bolton and Katok (2008) restrict the orders to be effective for ten periods. After each decision, a combined feedback about demand realization of the ten periods is provided to the subjects. The authors find standing orders to improve order decisions significantly towards the optimal. Lurie and Swaminathan (2009) compare the newsvendor performance under different feedback and order frequencies (once in every one, three and six periods). Similar to Katok and Wu (2008), they find reduced feedback frequency to improve decision quality. Ockenfels et al. (2015) show that the improvement in the newsvendor decisions through standing orders can be predicted by an impulse-balance-equilibrium model. Contrary to these findings, Bostian et al. (2008) report no significant improvement in the newsvendor performance due to five-period standing orders. In their experiments, the subjects receive feedback about demand realization either every period or once every five periods. Neither treatment results in improved decisions; in fact, the average order decisions are found to be closer to the demand mean than the theoretical optimal.

Further studies on standing decisions can be found in behavioral economics literature. For instance Beartzi and Thaler (1995), and Fellner and Sutter (2009) study the effect of longer-standing decisions in an investment context. They report improvement over the returns when decision frequency is reduced.

Placing a standing decision is an act of commitment in the strategic context, often providing an advantage to the decision maker. The power of commitment concept has been abundantly studied in game theory literature. Some researchers consider a single-shot interaction where one of the players commit to a decision before the others, obtaining the first mover advantage (See for instance Van Huyck et al. 1995, Vardy 2004, Morgan and Vardy 2007). Other researchers consider repetitive games where the commitment is to a strategy type throughout the whole game. Schmidt (1993) studies a repeated bargaining game in a seller-buyer setting where the seller determines and offers the selling price, and the buyer decides how much to buy if she accepts the offer. Schmidt proves that when the buyer commits to the strategy of rejecting prices higher than her valuation, the seller will reduce the price in the long-run, causing the buyer to earn the reward of commitment.

To the best of our knowledge, there exists no study on the effect of standing decisions or power of commitment in a supply chain contracting setting. Our study substantially differs from the aforementioned works due to the presence of factors such as strategic interaction and social preferences among multiple human decision makers.

Our second study is about the effect of priming the subjects with a fair contract. Priming, in psychology, is the process of invoking a temporary sensitivity on a subconscious level by exposing the subject to certain stimuli (Bargh and Chartrand 2000). Experimental economics and psychology studies illustrate how expectations determine the way outcomes are perceived, and that humans depend heavily on the first piece of information they receive while making decisions, making it an anchor point (Tversky and Kahneman 1974). Sanfey (2009) studies the effect of expectation on ultimatum game responses. When the subjects are primed with an expectation of higher/fair earnings, their rate of rejection for unfavorable offers is shown to increase. Maxwell et al. (1999) study the effect of priming in a seller-buyer context and demonstrate fairness-primed buyers to be more cooperative and more satisfied. The authors assume the sellers to reciprocate buyers' cooperative behavior. In their subsequent study (Maxwell et al. 2003), the authors assume a similar setting except that the sellers do not reciprocate, and observe the buyers to become significantly competitive. They conclude that priming the subjects for fairness invokes a sensitivity of fairness. To the best of our knowledge, the effect of priming has so far been investigated by only one study in the behavioral operations literature: The study by Villa et al. (2014) finds power-primed subjects to make better newsvendor decisions on average, though the difference is not statistically significant.

The scenario on which we base our experiments is basically a Stackelberg setting that strategically resembles an ultimatum game (Kagel and Roth 1995). In an ultimatum game, the first-mover proposes a way to split a certain amount of money between himself and the second-mover. The second-mover then either rejects or accepts the offer. If the offer is accepted the money is split accordingly, otherwise neither player earns a payoff. Standard economic theory suggests that the second-mover, being a profit-maximizing rational decision-maker, should accept any offer where she receives a positive amount of money. However, ultimatum game experimental studies has repeatedly shown that offers that allocate small amounts to the second-mover are usually rejected. Due to the fear of rejection, the average offer by the first-mover usually turns out to be comparably equitable (See Güth et al. 1982 and Roth 1995). Beyond the similarity in strategic interaction, our experimental setting is more complicated than the ultimatum game. Rather than making a simple accept-or-reject decision, the retailer decides on the order quantity. In addition, the firms' profits are affected by the realization of probabilistic demand.

Inequity aversion and fairness concerns are shown to significantly affect subject decisions in supply chain experiments. Keser and Paleologo (2003) are the first to conduct such an experimental study. They use a wholesale price contract setting and observe the manufacturers to offer prices that lead to almost equitable profit allocation. To explain retailer's behavior, Wu (2006), Cui et al. (2007), Ho and Zhang (2008) and Katok and Pavlov (2013) study fairness concerns and inequity aversion through linear utility functions that decline with the inequity between the firms' profits. Their findings indicate inequity aversion to have a strong effect on retailer's contract rejection and order quantity decisions. Wu (2013), and Akbay and Kaya (2016)'s experiments also indicate a more equitable split of supply chain profit between the firms than predicted by theory. Using regression models, the authors of both papers demonstrate inequity aversion to be a significant factor for the retailer's quantity decisions.

In our experiments, we use the parameter setting and demand stream of Katok and Wu (2009) and Akbay and Kaya (2016). In fact, the wholesale price contract treatment of Akbay and Kaya (2016) corresponds to our base treatment. In our two experimental studies, we compare the results of the base treatment with the "power of commitment" and "fairness priming" treatments.

5.3 Analytical Background and Experimental Procedure

In this section, we first present the analytical model on which we base our experiments, and its solution. Then we briefly summarize our experimental procedure.

5.3.1 Analytical Background

We adopt the same manufacturer-retailer supply chain and parameter setting used in Katok and Wu (2009), Wu (2013) and Akbay and Kaya (2016). The retailer faces a standard newsvendor problem where the purchase cost is determined by the manufacturer's wholesale price offer. Consumer demand is discrete uniformly distributed between 51 and 150. At the end of the selling season, unsold products lose their value and unsatisfied demand is lost.

The sequence of events in a selling season (a period) is as follows:

- ***Stage 1:*** The manufacturer determines the wholesale price w and offers it as the contract to the retailer. The wholesale price is bounded by the unit production cost (c=\$3) and the unit selling price (p=\$12) in the consumer market.
- ***Stage 2:*** Upon receiving the wholesale price offer, the retailer determines her order quantity Q. If she finds the contract unfavorable, she may reject it by placing an order of zero, otherwise the order quantity needs to be between 51 and 150. The manufacturer instantaneously produces and delivers the units to the retailer prior to the selling season.
- ***Stage 3:*** Random consumer demand d is realized and payoffs of the firms are computed as follows:

$$\pi_{\mathrm{R}} = p \min(d, Q) - wQ = 12 \min(d, Q) - wQ \,, \tag{1}$$

$$\pi_{\mathrm{M}} = (w - c)Q = (w - 3)Q \,. \tag{2}$$

Both firms are assumed to be risk-neutral and all decisions are restricted to integer values. The theoretical optimal decisions for the manufacturer and the retailer is determined by solving for the subgame-perfect equilibrium of the game through backwards induction. Given the wholesale price w and the distribution function *F(.)* of the consumer demand, the retailer's optimal order quantity $Q*$ is provided by the classical newsvendor solution:

$$Q^*(\mathrm{w}) = F^{-1}\left(\frac{p - w}{p}\right) = F^{-1}\left(\frac{12 - w}{12}\right) = 50 + \frac{12 - w}{12} 100 \,. \tag{3}$$

Assuming the retailer will act according to the theoretical prediction, the manufacturer's problem is to maximize the following profit function:

$$\pi_M = (w-3)Q^*(w) = (w-3)(50 + \frac{12-w}{12}100)\,. \qquad (4)$$

The optimal wholesale price is calculated as $w^* = 10$, leading to a retailer order quantity of $Q^*(\mathrm{w}) = 67$. The expected (with respect to the demand realization) outcomes of these two decisions are given in Table 27. Contract efficiency is calculated as the ratio of total expected profit under these decisions to the optimal total expected profit of the centralized system. For the centralized system, the optimal order quantity is 125, resulting in an expected total supply chain profit of $781.

Table 27: Theoretical solution of wholesale price contract scenario

w*	Q*	Manufacturer's Expected Profit	Retailer's Expected Profit	Total Expected Profit	Contract Efficiency
10	67	469.0	117.7	586.7	74.08%

5.3.2 Experimental Procedure

The experiments were conducted with Sabanci University students that were recruited through an online application system. In each treatment 22 pairs (44 students) participated, making a total of 176. The experiment was implemented in HP MUMS software and conducted in CAFÉ (Center for Applied Finance Education) computer laboratory at Sabanci University. Experiment instructions were sent to the subjects one day before each session. Before the actual experiment, a short tutorial was given and a pilot run of three periods was performed. Subjects were randomly and anonymously assigned the role of manufacturer or the retailer, and randomly matched with each other. They were informed that their roles and the manufacturer-retailer pairs will remain unchanged throughout the main experiment. Each session consisted of 40 periods. The actual length of the experiment was not disclosed to the subjects to prevent end-of-game effects. The duration of each session, including the tutorial and the pilot run, was about 2 hours. Subject were motivated by monetary payment ($15 on average) based on their total experimental profit. Sample experiment instructions can be found in the electronic companion.

5.4 Experimental Study 1: Power of Commitment

Here we study the effect of the power of commitment on decisions. We conjecture that commitment to a decision, whether it is the contract price or the order quantity, provides the owner of the decision with an advantage over the other firm. To test this, we compare results from the following three treatments.

1. Base Treatment (Base)

In this treatment, the scenario that was explained in Section 5.3.1 is played for 40 periods. First the manufacturer determines the wholesale price, then the retailer either determines the order quantity or rejects the contract. Finally, consumer demand is realized and both firm's payoffs are calculated. This treatment corresponds to the wholesale price contract treatment of Akbay and Kaya (2016).

2. Manufacturer's Commitment: Standing Price (SP) Treatment

In this treatment, the manufacturer is allowed to determine the wholesale price only once every five periods. In other words, the manufacturer offers a new price only in periods 1, 6, 11, …, 36, and each price stands effective for five periods. In each of these five periods, the retailer could order any quantity she prefers (or, she can reject the contract), demand is realized and the results are shared with both firms. This setup is designed to give the manufacturer the power of commitment, and to motivate the manufacturer towards developing a better understanding of each wholesale price he offers.

3. Both Firms' Commitment: Standing Price and Quantity (SPQ) Treatment

In this treatment, both the manufacturer's wholesale price and the retailers order quantity are determined once every five periods. Hence, both subjects make decisions only in periods 1, 6, 11, …, 36. In the other periods, they simply observe the outcome based on that period's particular demand realization. This setup is designed to give both firms the power of commitment to a decision, and opportunity to have a better understanding of their choices.

Table 28 compares the treatments with regard to the number of periods for which each decision is valid.

Table 28: Comparison of treatments

Treatment	Length of Effectiveness (number of periods)	
	Manufacturer's Contract Price	**Retailer's Order Quantity**
Base Treatment (Base)	1	1
Standing Price Treatment (SP)	5	1
Standing Price & Quantity Treatment (SPQ)	5	5

Note that a scenario where the retailer commits to her order quantity but the manufacturer is free to change the wholesale price at every period is not considered because it is not likely to be accepted by the retailer.

We distinguish between the accepted and rejected contracts in our analysis. In total, 8.1% of all contract offers were rejected by the retailers. For statistical comparisons, the unit of analysis is average of each subject's decisions over the accepted contracts, which yields 22 data points for each treatment. Comparison results including the rejecting contracts are presented in the electronic companion.

In what follows, we first compare the accepted contracts' data with theoretical predictions in Section 5.4.1. Then in Sections 5.4.2 and 5.4.3, we study the effect of commitment over the accepted contracts. Next we analyze the rejected contracts in Section 5.4.4. Finally in Section 5.4.5, combining all data, we present a random-effects logistic regression model to explain the retailer's decision.

5.4.1 Comparison with Theoretical Benchmarks (Accepted Contracts)

We first compare each of the three treatments with theoretical predictions.

Hypothesis 1: *For all treatments, subject decisions will be as predicted by the theory.*

In theory, our standing price or standing order manipulations should not affect the outcome of the game. Therefore the theoretical prediction of all treatments is as explained in Section 5.3.1. From Table 29, however, we observe all treatments' results to be significantly different from theoretical predictions. In particular, the wholesale price is lower and the order quantity is higher in experiments, leading to higher total supply chain profits than predicted. The total profit is shared more equitably, with retailers getting a higher and the manufacturers getting a lower profit share than predicted. The only measure which is not significantly different from theoretical predictions turn out to be the manufacturer's profit.

Table 29: Comparison of experiment results with theory (Base, SP and SPQ treatments)

	Theory	Base Treatment	Standing Price	Standing Price & Quantity
w	10	7.50 (0.59)***	8.07 (0.78)***	8.00 (1.06)***
Q	67	95.61 (13.56)***	91.97 (11.71)***	89.56 (9.28)***
Retailer's Profit	117.68	304.82 (55.77)***	256.27 (75.99)***	272.70 (102.35)***
Manufacturer's Profit	469	415.31 (72.66)**	456.00 (83.36)	440.09 (106.05)
Total Chain Profit	586.68	720.13 (56.56)***	712.27 (62.40)***	712.79 (46.53)***
Contract Efficiency	0.74	0.91 (0.07)***	0.90 (0.08)***	0.90 (0.06)***
Mfg.'s Profit Share	0.80	0.63 (0.08)***	0.69 (0.11)***	0.66 (0.15)***

Standard deviations are reported in parenthesis. P values are from a two tailed Wilcoxon signed-rank test.
*** p-value<0.001, ** p-value<0.01, * p-value<0.1

5.4.2 Manufacturer's Commitment to the Wholesale Price

Here, we compare the results from the Base treatment and Standing Price (SP) treatment. In each period of the base treatment, the manufacturer makes a take-it-or-leave-it wholesale price offer to the retailer with no room for negotiation. However, because the two firms interact repeatedly over periods, the relationship has a "bargaining" aspect as well. Each decision sends a signal to the other firm that might affect its subsequent periods' decisions. In particular, to earn more profits in the future periods, a firm may opt to reward or penalize its partner at the expense of current period profit. For instance, the retailer may prefer to reject a wholesale price offer (or place a small order) to warn the manufacturer, hoping that the manufacturer will offer lower prices and thus better profits in the future. Likewise, the manufacturer can reciprocate a contract rejection by the retailer with a high price offer in the subsequent period.

When the manufacturer commits to the contract price for a certain number of periods, five in our setting, the retailer's capacity to impact the manufacturer's decisions is reduced. This is because even if the retailer doesn't like the current price and protests it through a contract rejection (or through a small order quantity), the manufacturer won't be able to change the contract until the next contract change period. Retailer's bargaining power is diminished as her potential gains arising from a current period sacrifice are now further in the future. Hence, we expect the retailer to act in a more complaisant fashion and exhibit weaker protests to contracts. Meanwhile, we expect the manufacturer to gain power from commitment, act more aggressively, and earn higher profits. Thus, we conjecture that:

In the SP treatment, compared to the Base treatment,
Hypothesis 2A: *Retailers' underorder quantity will be lower and overorder quantity will be higher.*
Hypothesis 2B: *Manufacturers' profit and profit share will be higher.*

Here, an underorder is defined as the difference between the newsvendor optimal quantity Q^* and retailer's order quantity when the order quantity is below Q^*, computed as $max(Q^*-Q,0)$. Similarly, an overorder is the difference between the retailer's order quantity and Q^* when the order quantity is above Q^*, computed as $max(Q-Q^*,0)$. Table 30 summarizes the comparison based on our experiment data. We observe the manufacturer to offer higher wholesale prices in the SP treatment. However the retailer fails to react to this increase by significantly reducing her order quantity. Underorder quantity is significantly lower whereas we observe no increase in overorder quantity; in fact it is slightly lower. Thus, Hypothesis 2A is only partially supported. The manufacturer's profit is higher, retailer's profit is lower and hence, the manufacturer's profit share is higher when the manufacturer commits to price, supporting Hypothesis 2B. The total supply chain profit, and hence contract efficiency is a little bit lower in the SP treatment, but not significantly so.

Table 30: Comparison of experiment results between the Base and SP treatments

	Base Treatment	**Standing Price**	**P-Value**
w	7.50 (0.59)	8.07 (0.78)	**0.01**$^{\neq}$
Q	95.61 (13.56)	91.97 (11.71)	0.39$^{\neq}$
Q/Q*	1.09 (0.15)	1.11 (0.15)	0.72$^{\neq}$
Retailer's Profit	304.82 (55.77)	256.27 (75.99)	**0.02**$^{\neq}$
Manufacturer's Profit	415.31 (72.66)	456.00 (83.36)	**0.08**$^{\neq}$
Total Chain Profit	720.13 (56.56)	712.27 (62.40)	0.63$^{\neq}$
Contract Efficiency	0.91 (0.07)	0.90 (0.08)	0.63$^{\neq}$
Mfg.'s Profit Share	0.63 (0.08)	0.69 (0.11)	**0.01**
Underorder Quantity	5.97 (5.33)	3.89 (3.59)	**0.07**
Overorder Quantity	14.06 (8.6)	13.15 (8.55)	0.38

#Underorders	13.14 (8.24)	12 (6.93)	0.39
#Overorders	22.73 (9.98)	22.64 (10.47)	0.50

Standard deviations are reported in parenthesis. P-values marked with ≠ are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

5.4.3 Retailer's Commitment to Quantity

Here, we compare the results from the Standing Price (SP) and Standing-Price-and-Quantity (SPQ) treatments. When the retailer's order decision, in addition to the manufacturer's price offer, stand for a certain number of periods, the retailer's capacity to impact the manufacturer's decisions is somewhat restored. This is because if the retailer rejects a contract or places a small order, the manufacturer will earn zero or a small profit for five periods. Hence the retailer too is expected to earn some commitment power, and the manufacturer is expected to act less aggressive compared to the SP treatment. Hence we hypothesize that:

In the SPQ treatment, compared to the SP treatment
Hypothesis 3A: *Retailer's underorder quantity will be higher and overorder quantity will be lower.*
Hypothesis 3B: *Manufacturer's profit and profit share will be lower.*

The comparisons in Table 31 support both Hypothesis, albeit not at a statistically significant level. In particular, the retailer's profit is higher and the manufacturer's profit and profit share are lower when the retailer commits to an order quantity.

Table 31: Comparison of experiment results for SP and SPQ treatments

	Standing Price	**Standing Price & Quantity**	**P-Value**
w	8.07 (0.78)	8.00 (1.06)	0.66≠
Q	91.97 (11.71)	89.56 (9.28)	0.31≠
Q/Q*	1.11 (0.15)	1.09 (0.17)	0.42≠
Retailer's Profit	256.27 (75.99)	272.70 (102.35)	0.37≠
Manufacturer's Profit	456.00 (83.36)	440.09 (106.05)	0.45≠
Total Chain Profit	712.27 (62.40)	712.79 (46.53)	0.45≠
Contract Efficiency	0.90 (0.08)	0.90 (0.06)	0.81≠
Mfg.'s Profit Share	0.69 (0.11)	0.66 (0.15)	0.13
Underorder Quantity	3.89 (3.59)	4.37 (4.61)	0.38

Overorder Quantity	13.15 (8.55)	10.69 (8.83)	0.13
#Underorders	12 (6.93)	11.59 (10.51)	0.26
#Overorders	22.64 (10.47)	25 (12.15)	0.18

Standard deviations are reported in parenthesis. P-values marked with ≠ are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

Another interesting comparison is on the variances of profits. From the table, we observe both firms' profit variances to be higher under SPQ than under SP. More extreme profit values are observed in the SPQ treatment, due possibly to the firms' strategic commitments to their decisions.

5.4.4 Comparison of Rejected Contracts

Table 32 provides the details of the rejected contracts analysis. We observe SPQ retailers to be the least likely to reject a contract. The table also compares the "newsvendor-predicted" profit values of the rejected contracts, which are computed assuming that the retailer places the newsvendor-optimal order *Q*(w)* for a given wholesale price offer *w*. We observe the Base treatment retailers to reject more favorable contracts, i.e., contracts offering higher newsvendor-predicted profit, than SP and SPQ retailers.

Table 32: Comparison of rejected contracts (Base, SP, and SPQ treatments)

	# Rejections	**# Retailers with at least one rejection**	**Retailer's average newsvendor-predicted profit for the rejected contracts**	**Manufacturer's average newsvendor-predicted profit for rejected contracts**
Base Treatment	75	20	200.76	438.08
Standing Price	87	16	188.52	440.03
Standing Price & Quantity	55	11	183.60	437.73

Figure 11 presents the cumulative distribution of retailer's newsvendor-predicted profit share for all accepted and rejected contracts. In particular, we observe that SP manufacturers can make less favorable contracts accepted by the retailer more frequently than Base treatment manufacturers can make.

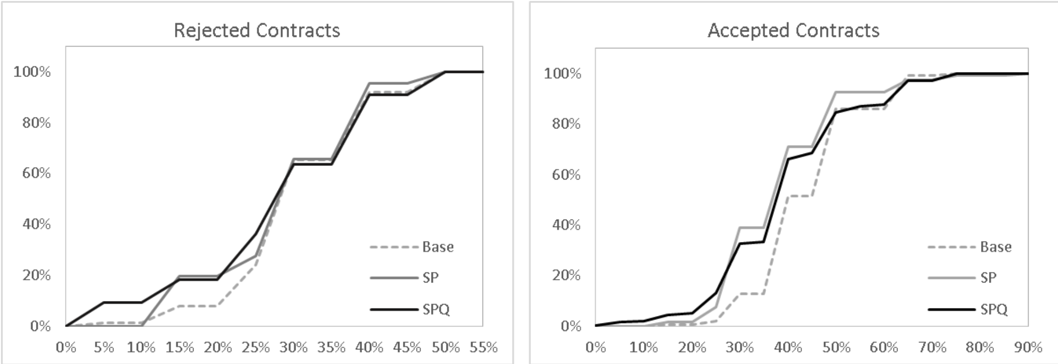


Figure 11: Cumulative distribution of newsvendor-predicted profit share of the retailer (Base, SP, SPQ)

Table 33 compares the proportion of rejected contracts by the offered wholesale price. Compared to the Base treatment, the proportion of rejected contracts for high wholesale price values is lower in the SP and SPQ treatments. This observation illustrates another benefit of the commitment to the manufacturer.

Table 33: Rejection rates (Base, SP, SPQ treatments)

Wholesale Price	**7**	**8**	**9**	**10**	**11**	**12**
Base Treatment	0.02	0.06	0.26	0.50	0.50	1.00
Standing Price	0.02	0.09	0.12	0.13	0.57	-
Standing Price & Quantity	0.03	0.05	0.08	0.13	0.20	0.33

Next we consider the manufacturer's reaction to a contract rejection. To this end, we analyze how manufacturers changed the wholesale price offer in the period following a rejection. Figure 12 presents the distribution of this wholesale price change for low (7, 8) and high (9, 10, 11, 12) rejected wholesale prices separately. Here, "-1" in the horizontal axis corresponds to a wholesale price reduction of 1. We observe the manufacturers to usually reduce their wholesale price offer following a rejection if the rejected wholesale price was relatively high. On the other hand, when a relatively low (7, 8) wholesale price offer is rejected, manufacturers are less likely to reduce it. In the latter case, some manufacturers even offered a higher price in the subsequent period than the rejected price. Consistent with our power of commitment argument, SP manufacturers are the least likely, whereas the SPQ manufacturers are the most likely to reduce the wholesale price following a contract rejection.

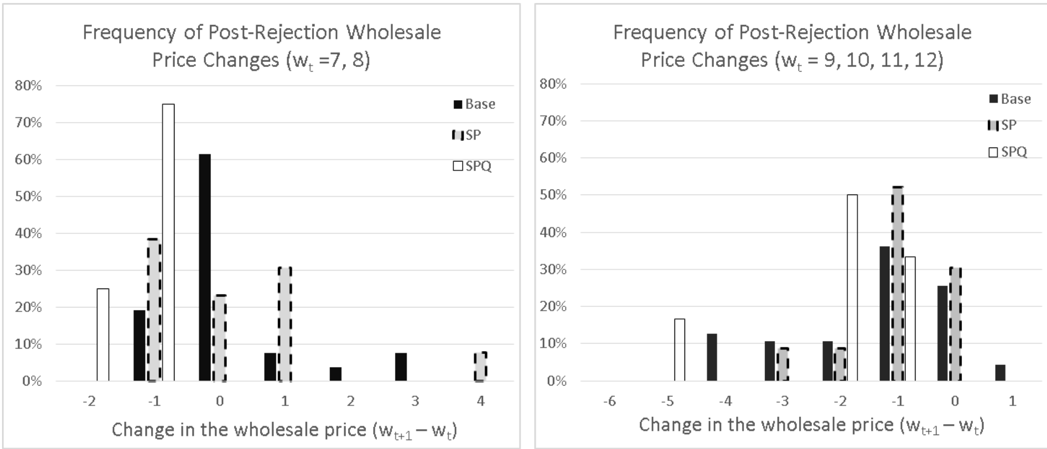


Figure 12: Post-rejection changes in wholesale price (Base, SP and SPQ treatments)

5.4.5 Modelling Retailer's Behavior

To further inspect the impact the power of commitment on retailer's order quantity decisions, we apply the following random-effects ordered-logit model to the data

$$Y_{it} = \alpha + \beta_1 w_{it} + \beta_2 \Delta w_{it} + \beta_3 D_{t-1} + \beta_4 t + \eta_i + \varepsilon_{it} \qquad (5)$$

The analysis is not applied to the SPQ treatment because the number of data points is not sufficient (only eight distinct order decisions per subject). Definition of the variables and indices used in the analysis are presented in Table 34. Rather than quantity itself, the dependent variable Y_{it} is defined as a categorical variable on the order behavior of the retailer as: Contract rejection (Y=1), underorder (Y=2), near-optimal (Y=3) and overorder (Y=4). This approach allows us to derive broad conclusions on the retailers' decision-making behavior.

Table 34: Variable definitions

Variable	Definition
Indices	
i	Retailer ID
t	Period
Dependent Variable	
Y_{it}	1, if $Q=0$ (contract rejection) 2, if $Q<0.95Q^*$ (underorder) 3, if $0.95Q^* \leq Q \leq 1.05Q^*$ (near-optimal) 4, if $Q>1.05Q^*$ (overorder)
Independent Variables	
w_t	Wholesale price in period t
Δw_t	Change in wholesale price, w_t-w_{t-1}
D_{t-1}	Demand realization in period t-1

t	Period index
Error terms	
η_i	Individual-specific error for retailers
ε_{it}	Common error term with standard logistic distribution

The results of this analysis are presented in Table 35.

Table 35: Regression results for the retailer's behavior (power of commitment)

Variables	**Dependent Category**	**Base Treatment**	**Standing Price (SP)**
# Observations	Y=1	72	86
	Y=2	223	195
	Y=3	93	128
	Y=4	470	449
Log likelihood		39.85	-885.731
Wald's Chi^2		-758.238	11.22
Intercept	Y=1	-15.757 (2.487)***	-7.18 (1.381)***
	Y=2	-13.057 (2.417)***	-5.29 (1.470)***
	Y=3	-12.32 (2.412)***	-4.429 (1.498)**
w		-1.633 (0.300)***	-0.541 (0.167)***
Δw		-1.571 (0.267)***	-0.471 (0.193)*
D_{t-1}		0.002 (0.003)	0.00 (0.003)
t		0.008 (0.010)	0.004 (0.012)

Standard deviations are reported in parenthesis.
*** p-value<0.001, ** p-value<0.01, * p-value<0.1

A positive coefficient for an independent variable implies that as that variable increases, the chances of the retailer placing an order with a higher *Q/Q** ratio also increase. For instance, as the wholesale price increases, the retailer is found to be more likely to reject a contract *(Y=0)* or to underorder *(Y=1)* than to, say, overorder *(Y=2)*. For the SP treatment the coefficients of *w* and Δ*w* are found to be much smaller in magnitude compared to those for the Base treatment. That is, for the same wholesale price or wholesale price increase, SP retailers are less likely to reject contract or to underorder than Base treatment retailers. This is again consistent with our power of commitment hypotheses.

A concrete example can help clarify the comparison between the treatments. Imagine in the 20th period the most recent demand realization to be $D_{19}=100$, and the increase in the wholesale price from period 19 to be $\Delta w = 1$. Consider two cases, with the current wholesale price being 9 and 10. The estimated regression model suggests the probabilities of each possible order behavior category as shown in Table 36.

Table 36: Sample estimation of probabilities for results in Table 8

	W_{19}=8, W_{20}=9		W_{19}=9, W_{20}=10	
Order Behavior Category	**Base**	**SP**	**Base**	**SP**
Rejection (Y=1)	0.538	0.128	0.856	0.201
Underorder (Y=2)	0.408	0.365	0.133	0.424
Near-Optimal (Y=3)	0.028	0.204	0.006	0.173
Overorder (Y=4)	0.027	0.303	0.005	0.202

The model predicts that, for instance, when the wholesale price increases from 8 to 9 (or from 9 to 10), SP retailers are less likely to react adversely than Base treatment retailers. This prediction of the estimated model is consistent with our previous experimental findings.

5.5 Experimental Study 2: Fairness Priming

Here we study the impact of priming the subjects with a fair contract anchor. To this end, we compare the results of the Base treatment with the results of the "fairness priming" treatment (FP). In the FP treatment, the subject pairs use an exogenously given wholesale price contract with *w=7* in the first five periods. During these periods the manufacturer does not make any decision. Under the assumption that the retailer will choose the newsvendor quantity *Q*(w=7),* the wholesale price *w=7* allocates the supply chain profit almost equally (subject to integer-constraints) between the manufacturer and the retailer, emulating a fair contract. Starting with period six, the firms engage in the regular game where the manufacturer determines the wholesale price as in the Base treatment. This decision structure is known to both subjects.

In what follows, we first compare the experiment results with theoretical benchmarks in Section 5.5.1. Then we compare the Base treatment and FP treatment over the accepted contracts in Section 5.5.2. The rejected contracts analysis is presented in Section 5.5.3. A regression model of the retailer's behavior is given in Section 5.5.4.

5.5.1 Comparison with Theoretical Benchmarks (Accepted Contracts)

Here, we compare the experiment results of the Base and FP treatments with the theoretical predictions. We use the data of periods 6-40 for both treatments. Recall that according to the standard theory, priming should have no impact on the subjects' decisions. Hence, the theoretical prediction for the FP treatment is the same as the prediction for the Base treatment.

Hypothesis 4: *For the FP treatment, subject decisions will be as predicted by the theory.*

Table 37 exhibits the comparison results. Parallel with our previous results, the FP treatment results are found to be significantly different from theoretical predictions.

Table 37: Comparison of FP treatment results with theory

	Theory	**Fairness Priming**
w	10	8.26 (1.08)***
Q	67	92.26 (9.02)***
Retailer's Profit	117.68	244.45 (104.44)***
Manufacturer's Profit	469	468.72 (93.41)
Total Chain Profit	586.68	713.16 (423.75)***
Contract Efficiency	0.74	0.90 (0.06)***
Manufacturer's Profit Share	0.8	0.70 (0.15)***

5.5.2 Priming Effect

Here we compare the Base treatment results with the FP treatment. The exogenously-given fair wholesale price 7 in the first five periods leads to a higher profit margin for the retailer than what she would typically get in the Base experiment scenario. In the Base treatment, the average wholesale price in all offered contracts is 7.62 and the average for the accepted contracts is 7.50 (See Table 30). A closer look at the distribution of these decisions reveals 55% of all contracts to have a wholesale price 8 or higher, and 13% to have wholesale price 6 or lower.

In the FP treatment, we conjecture the relatively fair profit shares during the first five periods to act as an anchor that will affect the subjects' subsequent decisions, causing the retailers to be primed on receiving a good profit share. Hence starting from the 6th period, we expect the retailer to claim a better profit share than what is typically achieved in the base treatment.

In the FP treatment, compared to the Base treatment
Hypothesis 5: *Retailer's profit and profit share will be higher*

Table 38 presents the results. Contrary to our expectation, this hypothesis is strongly rejected: It is the manufacturers' profit and profit share that increased following fairness priming. Manufacturers offered more aggressive prices, yet retailers did not react by reducing their order quantities sufficiently. In fact, we observe a significant reduction in retailer's underorder quantities and an increase (albeit not significant) in Q/Q*.

Table 38: Comparison of experiment results for Base treatment and Fairness Priming treatment

	Base Treatment	**Fairness Priming**	**P-Value**
w	7.50 (0.63)	8.26 (1.08)	**0.02**≠
Q	95.94 (14.00)	92.26 (9.02)	0.16≠
Q/Q*	1.09 (0.15)	1.14 (0.14)	0.38≠
Retailer's Profit	306.63 (61.26)	244.45 (104.44)	**0.04**≠
Manufacturer's Profit	416.99 (75.24)	468.72 (93.41)	**0.06**≠
Total Chain Profit	723.62 (57.61)	713.16 (43.75)	0.22≠
Contract Efficiency	0.91 (0.07)	0.90 (0.06)	0.22≠
Manufacturer's Profit Share	0.62 (0.09)	0.70 (0.15)	**0.04**
Underorder Quantity	5.95 (5.23)	3.53 (3.95)	**0.02**
Overorder Quantity	14.41 (8.83)	14.62 (7.29)	0.45
#Underorders	11.73 (7.36)	8.73 (5.77)	0.11
#Overorders	20.73 (9.2)	22.45 (7.3)	0.38

Comparison is conducted for periods 6-40. Standard deviations are reported in parenthesis. P-values marked with ≠ are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

A closer look at what happened in the transition period (period 6) may shed some light into this observation. In period 6, 17 out of 22 manufacturers increased the wholesale price to at least 8, from the value of 7 in the fair contract. Only five manufacturers offered a wholesale price of 7 or lower. What is more, the manufacturers in the first group obtained a higher average profit in periods 6-40 than the second group (15,701 versus 13,063). Hence, being aggressive in increasing the wholesale price seems to have worked well for the manufacturers. The retailers, on the other hand, did respond to the manufacturer's offer at period 6: Those who were offered a wholesale price of 7 or lower increased their order quantity compared to their first-five period average. Out of the 17 retailers who were offered a wholesale price of 8 or more, 12 reduced their order quantity compared to the first-five period average. So the retailers did respond to the price increases of the manufacturers but the response was not sufficient to change the profit allocation to their benefit.

5.5.3 Comparison of Rejected Contracts

Table 39 compares the rejected contracts between the Base and the FP treatments.

Table 39: Comparison of rejected contracts (Base and FP treatments)

	# Rejections	# Retailers with at least one rejection	Retailer's average newsvendor-predicted profit for the rejected contracts	Manufacturer' average newsvendor-predicted profit for the rejected contracts
Base Treatment	75	20	200.76	438.08
Fairness Priming	69	15	202.92	432.72

Fewer number of retailers exercised the option of contract rejection in the FP treatment. We don't observe any significant difference in average newsvendor-predicted profit shares in the rejected contracts between the two treatments. (The median of retailer's predicted profit share is 0.30 for both treatments and p-value is 0.74). The cumulative distribution of newsvendor-predicted profit share of the retailer for accepted and rejected contracts, presented in Figure 13 can shed more light onto the subjects' behavior. FP manufacturers were able to make the retailers accept less favorable contracts more frequently than base treatment manufacturers did. In particular, as seen in
Table 40, the rejection rates for high wholesale prices ($w>7$) are much lower in the FP treatment than in the Base treatment. Hence, once again we fail to observe the fairness priming effect on the retailers.

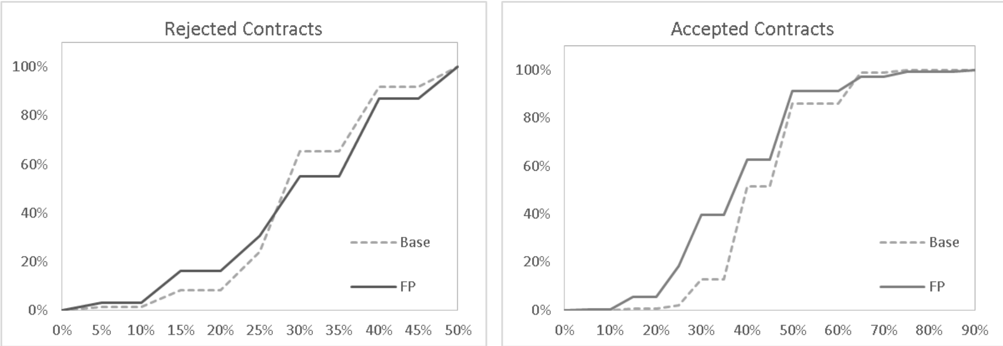


Figure 13: Cumulative distribution of newsvendor-predicted profit share of the retailer (Base and FP treatments)

Table 40: Rejection rates (Base and FP treatments)

w	**7**	**8**	**9**	**10**	**11**	**12**
Base Treatment	0.02	0.06	0.26	0.50	0.50	1.00
Fairness Priming	0.04	0.11	0.09	0.09	0.18	0.50

Next, we compare the post-rejection behavior of the manufacturers as presented in Figure 14. Here, w_t indicates the rejected wholesale price. For w_t >=9, we observe FP manufacturers to be less inclined to reduce their prices after a rejection than Base treatment manufacturers. FP manufacturers' price reduction frequency is 58% while WSP manufacturers' reduction frequency is 70%. For w_t 7 or 8, the difference is reversed with FP price reduction frequency being 32% and WSP reduction frequency being 19%. Thus, our explanation about manufacturers being primed towards more aggressive pricing behavior is supported only for prices above or equal to 9.

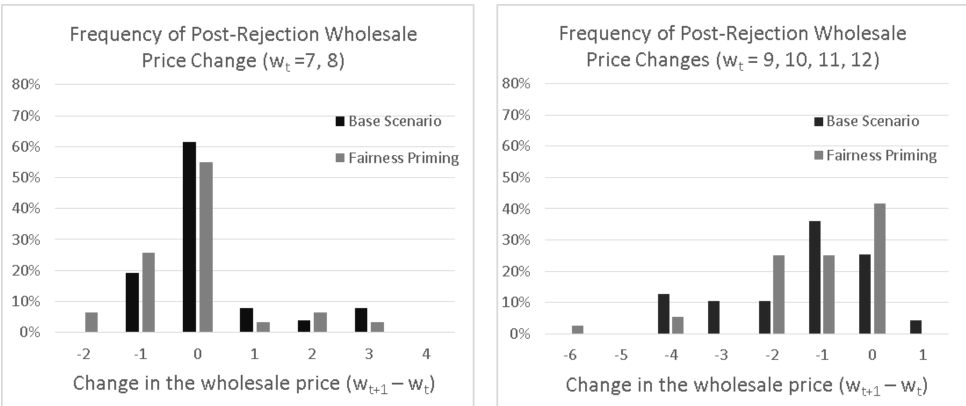


Figure 14: Post-rejection changes in wholesale price (Base and FP treatments)

5.5.4 Modelling Retailer's Behavior

We use the same random-effects ordered logit model used in Section 5.4.5 to analyze the retailer's response to fairness priming.

$$Y_{it} = \alpha + \beta_1 w_{it} + \beta_2 \Delta w_{it} + \beta_3 D_{t-1} + \beta_4 t + \eta_i + \varepsilon_{it} \tag{6}$$

Because the first five periods are used for priming, only periods 6-40 of the FP treatment are included in the analysis. The results of the estimation are shown in Table 41.

Table 41: Regression results for the retailer's behavior (Fairness Priming)

Variables	Dependent Category	Base Treatment	Fairness Priming (FP)
# Observations	Y=1	72	68
	Y=2	223	156
	Y=3	93	107
	Y=4	470	439
Log likelihood		-758.238	-770.72
Wald's Chi^2		39.85	23.68
Intercept	Y=1	-15.757 (2.487)***	-7.718 (3.636)*
	Y=2	-13.057 (2.417)***	-5.934 (3.604)
	Y=3	-12.32 (2.412)***	-5.124 (3.595)
w		-1.633 (0.3)***	-0.624 (0.378)*
Δw		-1.571 (0.267)***	-0.837 (0.295)**
D_{t-1}		0.002 (0.003)	0.001 (0.002)
t		0.008 (0.01)	0.018 (0.017)

Standard deviations are reported in parenthesis.
*** p-value<0.001, ** p-value<0.01, * p-value<0.1

We observe the impact of the wholesale price and the change in the wholesale price on the FP retailers' behavior to be weaker than the impact on the base treatment retailers' behavior. This finding also indicates fairness priming not to operate as expected on the retailers.

Similar to Section 5.4.5, an example can help grasp the implications of the above analysis. Imagine in the 20th period the most recent demand realization to be $D_{19}=100$, and the increase in the wholesale price from period 19 to be $\Delta w = 1$. Consider two cases, with the current wholesale price being w_{20}= 9 and 10. The estimated regression model suggests the probabilities of each possible order behavior category as shown in Table 42.

Table 42: Sample estimation of probabilities for results in Table 15

	W_{19}=8, W_{20}=9		W_{19}=9, W_{20}=10	
Order Behavior Category	**Base**	**FP**	**Base**	**FP**
Rejection (Y=1)	0.538	0.250	0.856	0.151
Underorder (Y=2)	0.408	0.415	0.133	0.364
Near-Optimal (Y=3)	0.028	0.152	0.006	0.190
Overorder (Y=4)	0.027	0.183	0.005	0.295

FP retailers exhibit a smaller likelihood of reacting to the contracts by rejection or by placing underorders than Base treatment retailers. Hence the hypothesis on fairness priming of retailers is observed to fail yet again.

5.6 Conclusions

In this paper we investigate the effects of power of commitment and fairness priming on the subjects' behavior in a simple supply chain experiment.

In our SP treatment, the manufacturer commits to the contract price for five periods while the retailer can make a different decision at every period. Intuition suggests that the manufacturer, being unable to react to the retailer's decisions immediately, would be worse off in this treatment compared to the base treatment. However, we observe the manufacturers to actually benefit from the commitment. They offer higher prices and earn higher profits. However much restricted the manufacturer might feel, the retailer seems to be more restricted as her impact on the manufacturer's contract decisions is reduced. In fact, in our post-experiment survey, for the SP treatment 70% of the retailer subjects, but only about 35% of the manufacturer subjects expressed that they felt restricted with commitment.

In the SPQ treatment, both the manufacturer and the retailer commit to their respective decisions for five periods. Under this treatment, we expected the retailers to claim some of the commitment power to themselves. Indeed, we observe the commitment to the order quantity decisions to earn the retailers more profit, and higher profit share compared to the case where only the manufacturer commits to his decision.

Here, although the number of periods is equal for the base, SP and SPQ treatments, number decisions per subject is not equal since base manufacturers make 40 decisions while SP and SPQ manufacturers make only 8 decisions. For (10-period) standing decisions treatment, Bolton and Katok (2008) extend the decision horizon 10 fold, in order to have the subjects make same number of ordering decisions. An extension for our study can be to extend the duration of the SP and SPQ treatments to 200 periods (40 decisions for the manufacturers) and compare the results with the findings of the current study.

To study the effect of fairness priming, we conduct the FP treatment where the interaction starts with a predetermined contract that distributes the expected profits equally. Contrary to our expectation, fairness priming backfired or was not strong enough. Here we primed the subjects with just the experience of a fair contract for five periods, no other stimuli was present. As an extension study, the impact of priming can be enhanced in two ways. 1) By extending the duration of the fair contract. 2) By encouraging the subjects to contemplate about fairness before the experiment. For this subjects can be given a tutorial about the profit allocations for each wholesale price and can be asked to think about what would be a fair price, how much of the total profit they deserve, etc. This way, we might obtain different results for the impact of fairness priming.

Acknowledgements

This research was funded by TÜBİTAK Grant 111K454. The authors would like to thank Elena Katok and Diana Wu for sharing their demand stream.

5.7 References

Akbay, Ü., Kaya, M. (2016). Experiments on Supply Chain Contracting: Effects of contract type. Working paper.

Bargh, J. A., Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social and personality psychology (pp. 253–285). New York: Cambridge University Press

Benartzi, S., R. H. Thaler. (1995). Myopic loss aversion and the equity premium puzzle. Quart. J. Econ. 110(1), 73–92.

Bendoly, E., Van Wezel, W., Bachrach, D. G. (Eds.). (2015). The handbook of behavioral operations management: Social and psychological dynamics in production and service settings. New York: Oxford University Press.

Bolton, G. E., Katok, E (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback. Manufacturing and Service Operations Management, 10(3), 519-538.

Bostian, AJ. A, Holt, C. A., Smith, A. M. (2008). Newsvendor "Pull-to-Center Effect": Adaptive Learning in a Laboratory Experiment. Manufacturing and Service Operations Management, 10(4), 590-608.

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

Fellner, G., M. Sutter. (2009). Causes, consequences and cures of myopic loss aversion: An experimental examination. The Economic Journal. 119(537), 900–916.

Güth, W., R. Schmittberger, B. Schwarze. (1982). An Experimental Analysis of Ultimatum Bargaining. Journal of Economic Behavior and Organization 3 (4), 367–388.

Ho, T., Zhang, J (2008). Designing Pricing Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter? Management Science, 54(4), 686-700.

Kagel, J.H., Roth, A. E. (1995). Introduction to experimental economics. ed: Kagel. J. H. and Roth, A. E., The Handbook of Experimental Economics, Princeton University Press, Princeton, 3-109.

Katok, E., Pavlov, V. (2013). Fairness in Supply Chain Contracts: A Laboratory Study, Journal of Operations Management, 31, 129-137.

Keser, C., Paleologo, G. A. 2004. Experimental Investigation of Supplier-Retailer Contracts: The Wholesale Price Contract. Working paper, CIRANO Scientific Series

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

Maxwell S, Nye P, Maxwell N. (1999) Less pain, same gain: the effects of priming fairness in price negotiations. Psychology and Marketing. 16(7):3 – 26.

Maxwell, S., Nye, P., Maxwell, N. (2003) The wrath of the fairness-primed negotiator when the reciprocity norm is violated. The Behavioral Aspect of Pricing. 46(5), 399-409

Morgan, J., Várdy, F. (2007) The value of commitment in contest and tournaments when observation is costly. Games and Economic Behavior. 60, 326-338.

Ockenfels, A., Selten, R. (2015) Impulse Balance and Multiple-Period Feedback in the Newsvendor Game. Production and Operations Management, 24(12), 1901-1906.

Roth, A. E. (1995) Bargaining Experiments. The Handbook of Experimental Economics, eds: John H. Kagel, A. E. Roth. Princeton University Press, 1995, 253-348.

Sanfey, A. G. (2009). Expectations and social decision-making: Biasing effects of prior knowledge on Ultimatum responses. Mind & Society. 8, 93–107

Schmidt, K. (1993) Commitment through incomplete information in a simple repeated bargaining model. Journal of Economic Theory. 60, 114-139.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Tversky, A.; Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science 185 (4157), 1124–1131.

Van Huyck, J. B., Battalio, R. C., Walters, M. F. (1995) Commitment versus Discretion in the Peasant-Dictator Game. Games and Economic Behavior, 10, 143-170.

Várdy, F. (2004). The value of commitment in Stackelberg games with observation costs. Games Economic. Behavior. 49, 374–400.

Villa, S., Castañeda J. A., Gonçalves P. (2014). The Effect of Power on Biased Newsvendor Order Behavior: An Experimental Study. POMS 25th Annual Conference.

Wu, D. Y. (2013). The Impact of Repeated Interactions on Supply Chain Contracts: A Laboratory Study. International Journal of Production Economics, 142, 3-15.

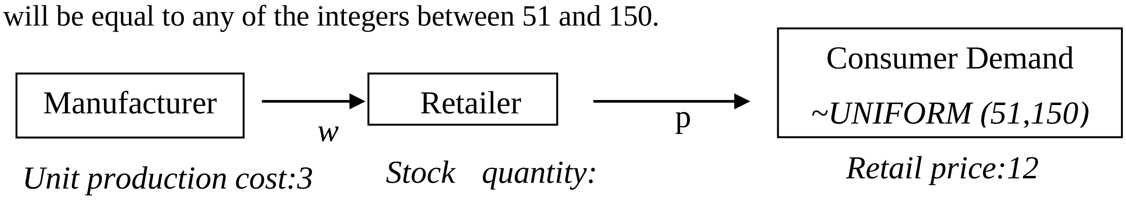
Wu, Y. (2006). Social Preferences in Coordination and Delegation Decisions- Models and Empirical Tests, (Doctoral Dissertation), INSEAD.

5.8 Appendix

5.8.1 Sample Experiment Instructions (Standing Price Scenario)

Scenario

We consider a manufacturer who produces a certain product, and a retailer who buys the product from the manufacturer and sells it to consumers. Consumer demand is uncertain. It is a random number distributed uniformly *between 51 and 150*. That is, there is a 1/100 chance that demand will be equal to any of the integers between 51 and 150.



We consider a three-stage game between the manufacturer and the retailer:

Stage-1: The manufacturer determines the ***wholesale price, w***. This is the price at which the manufacturer sells his product to the retailer. The wholesale price has to be an integer between 0 and the ***retail price 12***. This is the price at which the retailer sells the product to consumers.

Stage-2: The retailer observes the wholesale price offer of the manufacturer, and determines his ***stock quantity, Q*** for the product. The retailer may ***reject*** the manufacturer's offer by setting ***Q=0***. In this case, both firms earn zero profit. Otherwise, the retailer orders Q products from the manufacturer. The manufacturer produces this order by incurring a ***unit production cost 3*** per product, and delivers them to the retailer. The retailer stocks these products prior to the selling season. Because consumer demand can be between 51 and 150, the retailer's stock quantity Q decision also has to be between these values (if it is not equal to zero).

Stage-3: Random consumer demand is determined as "*d*". Using his stock of product, the retailer satisfies this demand as much as possible. The ***sales quantity*** of the retailer is the minimum of his stock quantity and the realized demand. Two cases are possible:

- If demand is higher than retailer's stock quantity (i.e., *d>Q)*, then retailer will sell all *Q* units, and *(d-Q)* units of demand will be unsatisfied (***unsatisfied demand***)
- If demand is less than the retailer's stock quantity, (i.e., *d<Q)*, then the retailer will sell *d* units, and *(Q-d)* products will be unsold (***leftover products***). These products have zero value.

Pricing Decision Effectiveness: **The manufacturer's pricing decisions are effective for 5 periods, and therefore the manufacturer makes a pricing decision once every 5 periods.** As such, the manufacturer will be asked to make pricing decisions at 1st, 6th, 11th, 16th etc. periods.

For instance, manufacturer's pricing decision at period 1 will be effective in periods 1, 2, 3, 4 and 5. The retailer will make a stock quantity decision at each period.

Each firm aims to maximize its *payoff* (or, profit) in the game.
The retailer's payoff is calculated as the sales price times the sales quantity, minus the payment to the manufacturer. That is $\mathbf{12 * sales - w * Q}$.
The manufacturer's payoff is calculated as the payment received from the retailer, minus the production cost. That is, $\mathbf{w * Q - 3 * Q}$

Note that there are two decisions in the game: The manufacturer determines the wholesale price *w* and afterwards, the retailer determines his stock quantity *Q*. Both of these decisions affect the payoff of both firms.

<u>Preparation for Our Experiments</u>

- The experiments will take place at the CAFÉ 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>Our Experiment</u>

- You will be randomly assigned to the role of either a manufacturer or a retailer. Manufacturer and retailers will be randomly matched with each other. You will play with the same partner throughout the experiment.
- The experiment repeats for a number of ***periods***, which are independent of each other. That is, a large or small demand realization in a period does not affect the demand in the later periods. Also, inventory is not carried from one period to the next. Leftover products from a period are assumed to be thrown away, and cannot be used to satisfy demand in following periods. Only your payoff will accumulate over periods.

<u>A Sample Screenshot</u>

The following figure on the next page 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 and the wholesale price that the manufacturer set at stage 1.
- The blue box in the upper right presents information on the last period.

- The red cell at the bottom is where you ***submit*** your decision to the server. You enter your decision value into the red cell and hit “enter” and then click on the green “Submit” button at the bottom (that will be visible during experiment). Note that the submit button will be activated only after you enter a valid decision and hit enter (or, click somewhere in the screen). Invalid entries will cause warnings.
- The cells in which you can enter values are labeled with “gray” or “red” background.

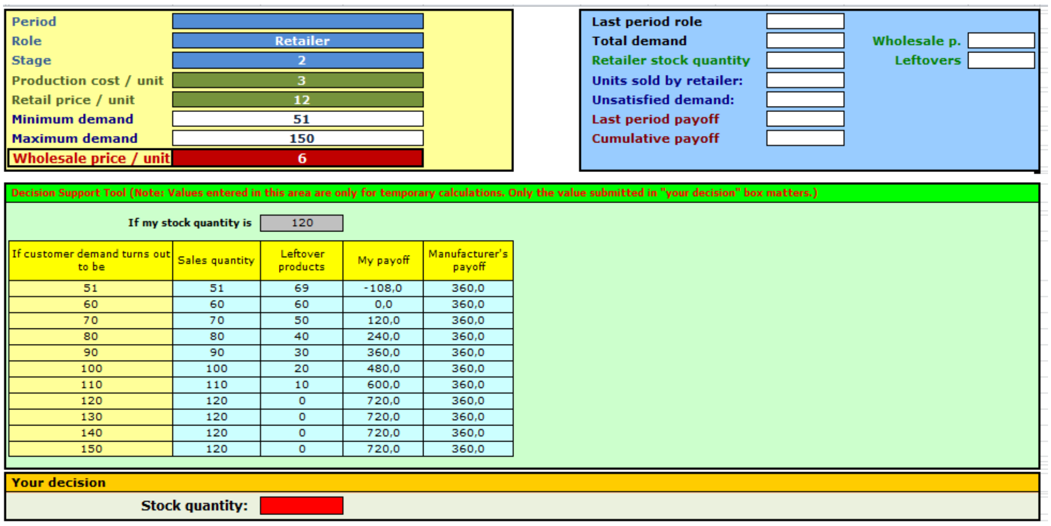


Figure 15: Retailer’s screen at stage 2

- You can check the results of previous periods by clicking the ***Historical Results*** tab in the bottom of the screen. This will open a second worksheet with the titles seen below:

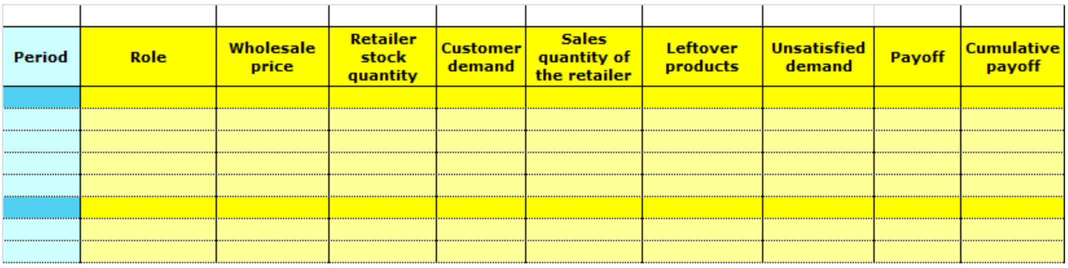


Figure 2: Historical results table

In the Historical Results sheet, the periods where the manufacturer will submit a wholesale price decision are highlighted bolder than the other periods.

<u>The Decision Support Tool</u>

Before you submit a decision, you can use the *decision support tool* that is in the middle of the screen. This tool allows you to calculate the outcome for certain values of your decision, the other firm’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 we record is the one you submit in the "stock quantity" box at the bottom of the screen.

Retailer's decision support tool at stage-1

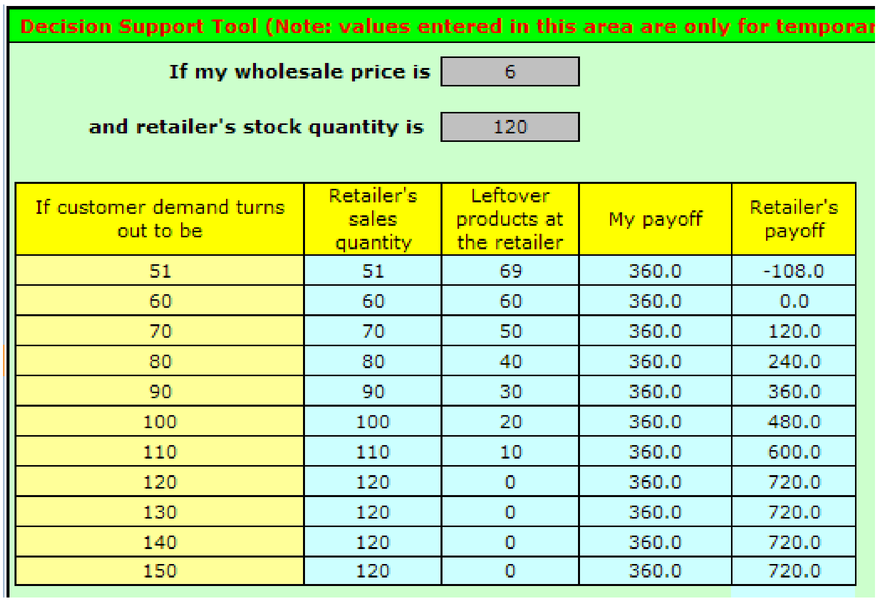
You may enter a "stock quantity" value in the top gray cell. To help you visualize the possible outcomes if you really set this stock quantity, the table summarizes the outcome for different consumer demand realizations (*d=51, 60, …, 150*) each in a row.

In the example in Figure 1, the retailer's stock quantity is entered as 120. We observe from the table that if consumer demand turns out to be, for example, 80, you (retailer) will sell 80 units because the demand is smaller than the stock quantity. Your leftover inventory will be 120-80=40 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 140. In this case, you (the retailer) will sell all of your 120 units, and there will be zero leftover inventory. Unsatisfied demand will be 140-120=20. The last two columns provide your payoff and the manufacturer's payoff.

Manufacturer's decision support tool at stage-1

At stage-1, you (the manufacturer) will only submit your wholesale 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. Figure 3 below illustrates what the outcome will be if you set 6 as your wholesale price, and if the retailer sets 120 as his stock quantity (i.e., if he orders 120 products from you).


Decision Support Tool (Note: values entered in this area are only for temporar

If my wholesale price is 6

and retailer's stock quantity is 120

If customer demand turns out to be	Retailer's sales quantity	Leftover products at the retailer	My payoff	Retailer's payoff
51	51	69	360.0	-108.0
60	60	60	360.0	0.0
70	70	50	360.0	120.0
80	80	40	360.0	240.0
90	90	30	360.0	360.0
100	100	20	360.0	480.0
110	110	10	360.0	600.0
120	120	0	360.0	720.0
130	120	0	360.0	720.0
140	120	0	360.0	720.0
150	120	0	360.0	720.0

Figure 3: Manufacturer's decision support tool at Stage 1.

5.8.2 Comparison Results when Rejected Contracts Included

4. Comparison of experiment results with theory. (WSP manipulations)

	Theory	EC	SC	SCO
w	10	8.37 (1)***	8.18 (0.78)***	8.09 (1.07)***
Q	67	83.97 (11.17)***	83.55 (17.79)**	84.23 (12.5)***
Retailer's Profit	117.68	218.22 (88.8)***	229.99 (76.09)***	254.73 (97.33)***
Manufacturer's Profit	469	431.46 (115.8)	415.32 (110.24)*	414.88 (114.49)*
Total Chain Profit	586.68	649.68 (77)**	645.32 (118.1)**	669.6 (76.29)***
Contract Efficiency	0.74	0.82 (0.1)**	0.81 (0.15)*	0.85 (0.1)**

Standard deviations are reported in parenthesis. P values are from a two tailed Mann Whitney U test.
**** p-value<0.001, ** p-value<0.01, * p-value<0.1*

5. Comparison of experiment results with for base scenario and SC scenarios

	Base Scenario	SC	P-Value
w	7.62 (0.55)	8.18 (0.78)	**0.01**≠
Q*	87.87 (15.77)	83.55 (17.79)	0.45≠
Q/Q	1 (0.18)	1.01 (0.23)	0.97≠
Retailer's Profit	278.7 (58.05)	229.99 (76.09)	**0.01**≠
Manufacturer's Profit	382.52 (85.19)	415.32 (110.24)	0.42≠
Total Chain Profit	661.22 (92.83)	645.32 (118.1)	0.95≠
Contract Efficiency	0.83 (0.12)	0.81 (0.15)	0.95≠
Underorder Quantity	11.65 (9.18)	10.59 (10.23)	0.19
Overorder Quantity	13.06 (7.86)	12.35 (8.54)	0.36
#Underorders	16.55 (9.72)	15.95 (9.78)	0.50
#Overorders	22.73 (9.98)	22.64 (10.47)	0.50

Standard deviations are reported in parenthesis. P-values marked with ≠, are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

6. Comparison of experiment results with for SC and SCO scenarios

	SC	SCO	P-Value
w	8.18 (0.78)	8.09 (1.07)	0.55≠
Q	83.55 (17.79)	84.23 (12.5)	0.84≠
Q/Q*	1.01 (0.23)	1.02 (0.2)	0.92≠
Retailer's Profit	229.99 (76.09)	254.73 (97.33)	0.20≠
Manufacturer's Profit	415.32 (110.24)	414.88 (114.49)	0.93≠
Total Chain Profit	645.32 (118.1)	669.6 (76.29)	0.72≠
Contract Efficiency	0.81 (0.15)	0.85 (0.1)	0.72≠
Underorder Quantity	10.59 (10.23)	8.53 (7.96)	0.21
Overorder Quantity	12.35 (8.54)	10.23 (8.8)	0.17

#Underorders	15.95 (9.78)	14.09 (11.71)	0.18
#Overorders	22.64 (10.47)	25 (12.15)	0.18

Standard deviations are reported in parenthesis. P-values marked with ≠, are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

7. Comparison of experiment results with for base scenario and EC scenarios

	Base Scenario	**EC**	**P-Value**
w	7.61 (0.61)	8.37 (1)	**0.01**≠
Q	88.76 (17.2)	83.97 (11.17)	**0.07**≠
Q/Q*	1.01 (0.2)	1.05 (0.2)	0.92≠
Retailer's Profit	281.96 (65.05)	218.22 (88.8)	**0.02**≠
Manufacturer's Profit	386.18 (89.57)	431.46 (115.8)	0.34≠
Total Chain Profit	668.14 (99.32)	649.68 (77)	0.18≠
Contract Efficiency	0.84 (0.13)	0.82 (0.1)	0.18≠
Underorder Quantity	14.41 (8.83)	14.62 (7.29)	0.45
Overorder Quantity	11.31 (9.71)	9.75 (9.48)	0.20
#Underorders	14.64 (8.99)	11.82 (7.25)	0.22
#Overorders	20.73 (9.2)	22.45 (7.3)	0.38

Comparison is done over periods 6-40. Standard deviations are reported in parenthesis. P-values marked with ≠, are from a two tailed Mann Whitney U test, while unmarked ones are from a one tailed test.

Chapter 6

6. GENDER DIFFERENCES ON NEWSVENDOR BEHAVIOR

Gender Differences on Newsvendor Behavior

6.1 Introduction

The newsvendor model is one of the most popular models in supply chain literature. The model offers a compact solution to the single-period inventory management of a product with stochastic demand. Here “newsvendor” is a gender-neutral term introduced to the literature to replace the original term “newsboy”. However experimental and empirical studies indicate significant differences between the two genders’ decision-making behavior. The aim of this paper is to analyze the gender differences in newsvendor ordering behavior and to identify various manifestations of this difference through an experimental study.

We present the results of an experiment in which human subjects undertake the role of a newsvendor and make ordering decisions in the face of stochastic demand. Similar to the literature we consider two profit margin treatments; high profit margin and low profit margin.

We aim to answer the following questions: First, do female and male decision-makers differ in terms of newsvendor ordering behavior? Second, how does gender affect the usage of decision heuristics?

We show that female subjects on average place smaller orders under both profit margins. The difference is significant for the high profit margin. Under both profit conditions, females make lower profit than males. We also demonstrate that female subjects have higher order variability and they resort to demand chasing heuristic more often than the male subjects do. We find no conclusive result regarding the use of the mean anchor heuristic.

The rest of the paper is organized as follows: Section 2 summarizes the related literature, Section 3 explains the analytical background and the experimental design. Section 4 presents the experiment results and analyses. Section 5 concludes the study and discusses further research directions.

6.2 Related Literature

Experimental studies on newsvendor behavior started with Schweitzer and Cachon (2000). The authors observe that order decisions exhibit regular deviations from the optimal quantity given by the classical newsvendor equation. Specifically, the average order quantity falls in between the optimal and the demand mean, as if it is being pulled towards the mean. The authors show that this phenomenon which they refer to as “the pull-to-center effect”, cannot be explained by the standard models such as risk preferences, waste aversion or stock-out aversion. Among their suggestions for possible explanations are the “mean anchor” and “demand chasing” heuristics. These two are special cases of anchoring-and-adjustment type heuristics which are introduced by Kahneman et al. (1982). Anchoring-and-adjustment type heuristics assume the decision-maker considers the first piece of available information as an anchor point regardless of its relevance to the task at hand. As more information becomes available, the decision-maker adjusts her decision towards the optimal.

Following Schweitzer and Cachon, the pull-to-center effect has been observed in various other studies. Bolton and Katok (2008) and Benzion et al. (2008) show that the pull-to-center effect weakens with time due to learning, and the orders get closer to the optimal. Bostian et al. (2008) explain the pull-to-center effect with an adaptive learning model based on probabilistic choice, recency and reinforcement biases. Kremer et al. (2010) demonstrate that random errors do not explain the newsvendor behavior. These authors find strong evidence for a matching-order-with-demand type heuristic, namely the mean anchor, demand chasing or minimizing ex-post inventory error heuristics.

The mean anchor heuristic assumes that the decision maker takes the demand mean as the anchor point, and adjusts order quantities towards the optimal. As a result, order quantities are expected to lie between the mean demand and the optimal quantity. In this sense the definition of the mean anchor heuristic coincides with the definition of the pull-to-center effect. As a matter of fact, Lau et al. (2013) use “pull-to-center effect” and the “mean anchor heuristic” interchangeably. Lau et al. make a review of various definitions and measures of the mean anchor heuristic used by several preceding behavioral operations studies. The authors conduct an individual level analysis on the experiment data of Bolton and Katok (2008) as well as the data from their own experiment. Their analysis reveal that ordering behavior of individuals are highly heterogeneous, and the pull-to-center effect is not necessarily supported at an individual level. In our study, we embrace Lau et al. (2013)'s notion and treat pull-to-center effect and the mean-anchor heuristic as the same.

The demand chasing heuristic assumes the decision maker to anchor at her most recent order decision, and make adjustments towards the most recent demand realization. This heuristic results from the random nature of demand realizations. Bolton and Katok (2008) document that by constraining order decisions to be effective for longer periods, demand chasing tendency can be suppressed and newsvendor performance can be improved. While Lurie and Swaminathan (2009) also find support for the improvement in newsvendor performance with reduced feedback frequency, Bostian et al. (2008) report no significant improvement. Lau and Bearden (2014) evaluate various methods used in the literature to check for existence of the demand chasing heuristic, and conclude inspecting the correlation of order decisions with the previous period's order decisions to be the best method.

In this study we analyze the use of mean anchor and demand chasing heuristics at an individual level in order to identify how gender difference manifests itself on newsvendor ordering decisions.

Gender differences in decision-making has been studied extensively in various fields. Byrnes et al. (1999) conduct an extensive literature search on gender differences in risk taking, and perform a meta-analysis on 150 studies from psychology, medicine, economics and some other fields. The authors conclude that females make less risky choices than males. On the other hand, in an experimental economics study Schubert et al. (1999) demonstrate that females don't always make less risky choices, and the decision frame is an important factor on gender differences. Croson and Gneezy (2009) present a review of gender differences in experimental economics focusing on risk preferences, social preferences and competition attitude. The authors conclude that females are more risk-averse than males.

Most related to our work is a study conducted by De Vericourt et al. (2013) which addresses gender differences in newsvendor ordering behavior. Under high profit margin, females are found to place significantly smaller orders than males; however, the difference is not significant under low profit margin. Using a mediation method, the authors demonstrate females to place smaller orders due to lower risk appetite. There are a few other behavioral operations studies which report a gender comparison of experiment results. In order to improve newsvendor performance, Lee and Siemsen's (2016) experimental study decomposes ordering decisions into forecasting and service level determination phases. The authors find no gender differences. Wu and Seidman (2014) conduct a tournament-type newsvendor experiment to address the effect of competition on newsvendor performance. The authors base their study on a high profit margin setting, and obtain parallel results to de Vericourt et al.'s. Specifically, in the high-competition environment females place smaller orders and earn lower profit. In the low-competition environment, the two studies report conflicting results.

6.3 Analytical Background and Experimental Design

Here we provide the theoretical solution of the model and experimental design.

6.3.1 Newsvendor Model and the Theoretical Solution

We consider a standard single-period, single-item newsvendor model. Right before each selling season, a newsvendor stocks Q units of a perishable product at a unit cost of *c*. Then, stochastic consumer demand *D* realizes and each product, if demanded, is sold at unit selling price *p*. At the end of the selling season, any unsold product is disposed of and any unsatisfied demand is lost. The sales quantity is found to be the minimum of *Q* and *D*, and the newsvendor's profit is expressed as

$$\pi(Q,D) = p\min\{Q,D\} - cQ\ .$$

Assuming that the demand has distribution function *F(.)* and density function *f(.)*, and taking expectation over *D* yields the expected profit as

$$E[\pi(Q,D)] = \int_0^Q (px - cQ) f(x)dx + (p-c)Q\int_Q^\infty f(x)dx.$$

This concave function is maximized at the well-known critical fractile solution $Q^* = F^{-1}\left(\frac{p-c}{p}\right)$.

6.3.2 Experimental Design

Previous behavioral operations studies have demonstrated that the newsvendor behavior is affected by the position of the optimal order quantity relative to the mean of the demand distribution. In other words, the profit margin being high or low is shown to alter the newsvendor subjects' ordering behavior. As explained earlier, this phenomenon is known as the pull-to-center effect. Literature on newsvendor ordering behavior has observed the pull-to-center effect to be asymmetrical under high and low profit margins: Under high profit margin, the difference between the average order quantity and the optimal is shown to be greater than the difference under low profit margin. Moreover, this asymmetry is carried over to the gender differences in de Vericourt et al.'s (2013) study. While the gender gap under high profit margin is significant in that study, the gap under low profit margin is not. To analyze the effects of profit margin, we also conduct separate high profit margin and low profit margin treatments in this study.

The selling price *p* is \$90 for both treatments. The purchasing cost *c* is \$35 and \$55 for the high and low profit margin treatments respectively. Consumer demand has a discrete uniform distribution between 50 and 150. Given these parameters, the optimal order quantities and the resulting optimal expected profit values for the two treatments are shown in Table 43.

Table 43: Experiment parameters and theoretical solution

Treatment	***c***	***p***	**Critical Fractile**	***Q****	**Optimal Expected Profit**
High Profit Margin	35	90	61%	111	4420
Low Profit Margin	55	90	39%	89	2420

Our experiments were conducted at Sabanci University with junior and senior undergraduate students who received formal training on the newsvendor model prior to the study. The experiments consisted of 40 consecutive and independent selling seasons (periods). Before the experiment, a pilot run of three periods was played in order to familiarize the subjects with the experiment interface. Out of 257 subjects, 11, all male, were able to compute the optimal order quantity and placed that order quantity in all periods of the experiment. The data of these subjects are included in our analyses. The sample sizes along with the gender breakdown are shown in Table 44.

Table 44: Sample sizes of the treatments

	Female Subjects	**Male Subjects**	**Total**
High Profit Margin	54	51	105
Low Profit Margin	67	85	152

The experiment was coded with MS Excel and VBA. Subjects entered their order decisions through an input-dialog box, after which a macro computed and displayed the demand realization, sales amount and profit value. A sample screenshot of the experiment software is presented in the Appendix. Prior to each session, written instructions explaining the experiment scenario and the use of the software were provided to the subjects. Additionally a brief tutorial was given before the experiment. Each session including the tutorial lasted about 35-40 minutes. A copy of the experiment instructions is presented in the Appendix.

We use nonparametric hypothesis tests in our analysis. For single-sample comparisons Wilcoxon test and for two-sample comparisons Mann Whitney U tests were used. Significance level is taken as 10%.

6.4 Results

In this section, we first compare the experiment results with theoretical predictions. Next, we compare the experiment performance of the genders with each other. Then, we compare the genders in decision heuristic use. Finally we present an aggregate-level analysis using a fixed-effects regression model.

6.4.1 Descriptive Statistics

6.4.1.1 Comparison with Theory

We first compare the experiment results with theoretical benchmarks. As previous studies of the field have established, we expect to have the experiment data to differ from theoretical predictions.

Hypothesis 1: *Under high profit margin, the average order quantity will be below the optimal; and under low profit margin, the average order quantity will be above the optimal.*

Table 45 presents results separately for the two genders. We observe the average order quantity in the high profit margin (HPM) treatment to be below the optimal, while in the low profit margin treatment (LPM) it is above the optimal. All deviations are significant for both genders and thus Hypothesis 1 is supported.

Table 45: Comparison of the order quantity and profit with theory

		Theory	Female Subjects				Male Subjects			
			Mean	**S. Dev.**	**Median**	**P**	**Mean**	**S. Dev.**	**Median**	**P**
HPM	Order Quant.	111	99.27	9.57	99.31	**0.00**	103.39	7.47	103.10	**0.00**
	Profit	4420	4210.70	150.00	4202.10	**0.00**	4318.50	177.50	4337.50	**0.00**
LPM	Order Quant.	89	94.00	9.57	95.05	**0.00**	96.56	7.43	95.83	**0.00**
	Profit	2420	2238.40	150.80	2242.10	**0.00**	2288.00	172.40	2334.5	**0.00**

P-values are obtained from two tailed Wilcoxon tests. Sample size is the number of subjects.

In the newsvendor setting, realized profit incorporates a factor of chance as it depends on demand realization. By excluding the effect of the random demand, the expected profit for a given order quantity decision offers a better performance measure. Figure 16 illustrates how the expected profit changes as a function of the order quantity.

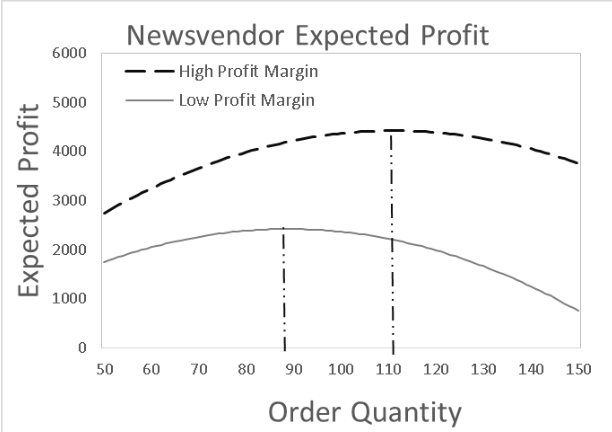


Figure 16: Expected profit as a function of the order quantity

Table 46 compares the average expected profit values based on subjects' order quantities with the optimal expected profit. Similar to the results presented above, the difference between theory and experiment data is seen to be significant for both genders.

Table 46: Comparison of the expected profit with theory

		Theory	Female Subjects				Male Subjects			
			Mean	**Std. Dev.**	**Med.**	**P**	**Mean**	**Std. Dev.**	**Med.**	**P**
HPM	Expected Profit	4420	4171.8	134.0	4185.7	**0.00**	4244.5	133.1	4264.1	**0.00**
LPM	Expected Profit	2420	2202.9	126.3	2239.6	**0.00**	2234.0	146.2	2261.6	**0.00**

P-values are obtained from two tailed Wilcoxon tests. Sample size is the number of subjects.

6.4.1.2 Gender Comparison

Here, we compare the experiment performances of the female and male subjects. Croson and Gneezy (2009), who review a number of experimental economics studies from risk, social and competitive preferences perspectives, conclude that females are more risk averse than males. To test this result on our subject pool, we conducted a survey analysis to which 92 female and 110 male subjects participated. For this survey, we asked the subjects to answer the Domain-Specific-Risk-Taking (DOSPERT) scale (Weber et al. 2002) questions before the experiment. DOSPERT scale has 30 questions designed to measure appetite for financial, social, ethical, health related and recreational risks. We find that females score significantly higher risk aversion than males for both financial questions (p-value = 0.004), and for all questions (p-value = 0.029).

Additionally, in order to check the relationship between risk-aversion and order quantity decisions, we regress the average order quantity decision of each subject with their DOSPERT financial score. We use the equation $\bar{Q}_i = \beta D_i^f + \varepsilon$ for this analysis. The regression results shown in Table 47 show that the order quantity decision increases with increases in DOSPERT financial score. A higher score in the DOSPERT scale translates to a lower risk aversion. Thus, as risk aversion increases, the average order quantity decision is found to decrease.

Table 47: Order quantity and risk-aversion regression results

Variables	HPM	LPM
Coefficient (Std. Dev)	4.23 (0.13)	3.75 (0.08)
P-value	0.00	0.00
R^2	95%	94%

The riskiness of an order quantity decision in the newsvendor setting can be measured with the variance of the profit value to be earned given that order quantity. As the order quantity increases, the variability of the profit increases. It follows that a risk-averse decision maker would avoid placing high orders. Hence we expect female subjects to place smaller orders than male subjects.

Hypothesis 2: *Female order quantities will be lower than male order quantities under both profit margins.*

Comparison results are shown in Table 48. As expected, female average order quantities are smaller than male average order quantities. The difference is significant only under high profit margin. Thus, Hypothesis 2 is partially supported. Female expected profit is lower than males, with significant difference under HPM. These results are in parallel to de Vericourt et al.'s (2013) findings. Under both profit margins male subjects make significantly more realized profit than females. Although their expected profit values are also higher, the gap between the realized profit values is wider, which indicates random demand realizations are more in favor of the male subjects.

Table 48: Gender comparison of the experiment performances

		Female Subjects			Male Subjects			P Value
		Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	
HPM	Order Quantity	99.27	9.57	99.31	103.39	7.47	103.10	**0.006**
	Expected Profit	4171.80	134.00	4185.70	4244.50	133.10	4264.10	**0.004**
	Realized Profit	4210.70	150.00	4202.10	4318.50	177.50	4337.50	**0.000**
LPM	Order Quantity	94.00	9.57	95.05	96.56	7.44	95.83	0.248
	Expected Profit	2202.90	126.30	2239.60	2234.00	146.20	2261.60	**0.068**
	Realized Profit	2238.40	150.80	2242.10	2288.00	172.40	2334.50	**0.006**

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

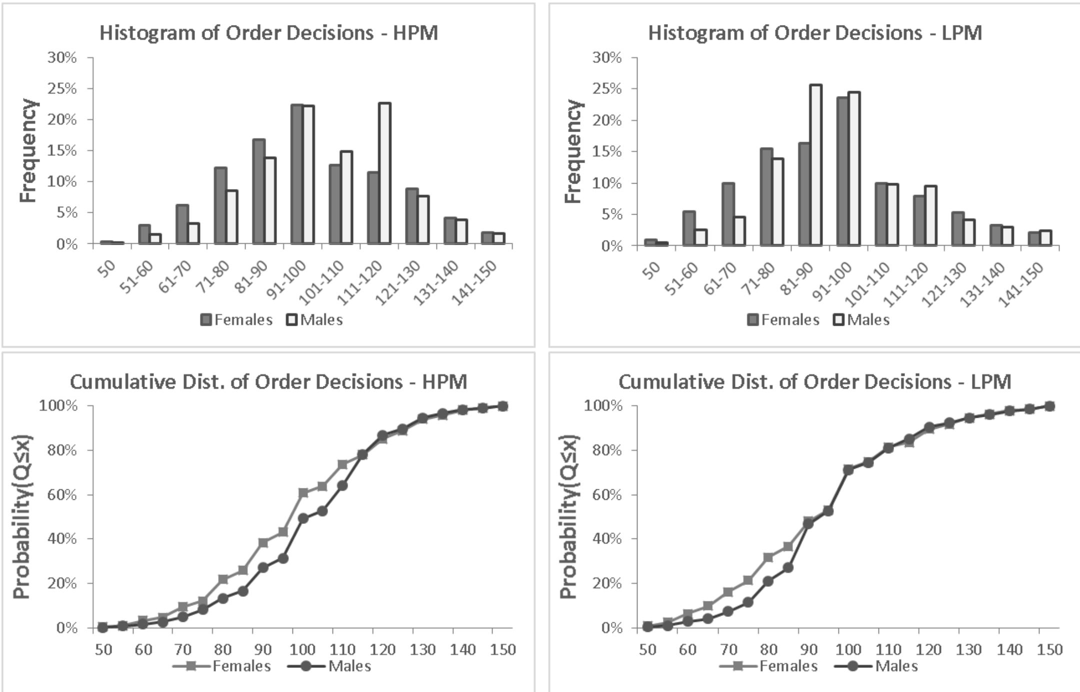


Figure 17: Histogram and cumulative distribution of order decisions (Pooled data)

The histogram and cumulative distribution of the pooled (over all subjects) order quantity decisions, which are presented in Figure 17, provide a visual and more illuminating depiction of the ordering difference between the two genders. We observe that female subjects place smaller orders with higher frequencies, and larger orders with lower frequencies than male subjects.

Next we compare the two genders in terms of order variability.

6.4.2 Order Variability

In this section we conduct an individual level analysis on order quantity variability. We compute order standard deviation for each decision maker, then compare the female and male samples of standard deviations under both profit margins. Table 49 presents the mean, median and standard deviation of these samples. We observe female orders to have higher standard deviation than male orders, with significant difference under LPM.

Table 49: Gender comparison of order standard deviation

	Female Subjects			**Male Subjects**			**P Value**
	Mean	**Std. Dev.**	**Median**	**Mean**	**Std. Dev.**	**Median**	
HPM	17.11	6.37	18.17	14.88	8.05	15.78	0.189
LPM	18.38	6.48	18.02	15.95	7.50	15.48	**0.021**

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

Next, we compare the number of unique (distinct) order decisions by each subject. Table 50 shows a similar result to Table 49: Female subjects have higher unique order counts than males, and the difference is significant under LPM.

Table 50: Gender comparison of unique order counts

	Female Subjects			**Male Subjects**			**P Value**
	Mean	**Std. Dev.**	**Median**	**Mean**	**Std. Dev.**	**Median**	
HPM	18.69	9.21	21.50	13.75	9.06	12.00	**0.010**
LPM	18.36	8.56	18.00	15.11	9.35	13.00	**0.011**

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

Given the observed variability in the order decisions, one might ask if there is significant increase or decrease in the order quantities over time. Table 51 shows percentage of subjects whose order decisions changed significantly from the first 10 periods to the last 10 periods, ignoring the data of the 20 periods in between. The results indicate that a significant change in order decisions is observed only for a small percentage of all subjects.

Table 51: Comparison of subjects whose order quantities changed from first 10 to last 10 periods

In the last 10 periods compared to the first 10	HPM Females	HPM Males	LPM Females	LPM Males
Increased	17% (#9)	16% (#8)	12% (#8)	13% (#11)
Decreased	6% (#3)	10% (#5)	7% (#5)	12% (#10)

Next, we compare the two genders based on their use of decision heuristics.

6.4.3 Use of Decision Heuristics

Here we compare the female and male subjects in terms of mean-anchor and demand chasing heuristic use.

6.4.3.1 Pull-to-Center Effect (Mean-Anchor Heuristic)

One of the theories explaining gender differences in information processing is the selectivity hypothesis (Meyers-Levy, Loken 2015). According to this hypothesis, female decision makers pursue a more comprehensive approach and take into account all available information, while males only consider selected information which they regard as more important. Additionally, females have lower threshold for information which means that they can perceive details and cues more than males can. Building on these findings, Kudryavtsev and Cohen (2011) hypothesize that women will be more prone to anchoring and hindsight biases. Conducting an experiment where the subjects are required to remember some recent financial indicators, the authors confirm women subjects to exhibit higher tendency for both biases. In another study conducted by Hügelschäfer and Achtziger (2014), which is focused on the effect of mindset on decision making, female subjects' decisions are found to be significantly closer to the anchor value than that of the male subjects, indicating that females are more affected by the anchoring behavior. An earlier study by Cervone and Peake (1986) reports no significant gender difference in anchoring behavior. However that study measures the anchoring effects on the self-reported self-efficacy of the subjects, which is fundamentally different in nature from the two studies mentioned above or from our study. We base our hypothesis on the two more recent studies, and expect female subjects to be affected more by the pull-to-center effect than male subjects are affected.

Hypothesis 3: *Females will be affected by the pull-to-center effect more than males will be.*

In the behavioral operations literature, various studies have used different methods to check for support for the pull-to-center effect. One method utilized by Lau and Bearden (2013) is to find the proportion of orders that fall in the interval between the optimal order quantity and the demand mean, which is referred to as the "pull-to-center region". For our high profit margin treatment, this region is [100,111] and for the low profit margin treatment it is [89,100]. Given this definition, Figure 18 presents the breakdown of the order decisions that fall into and outside the PTC region.

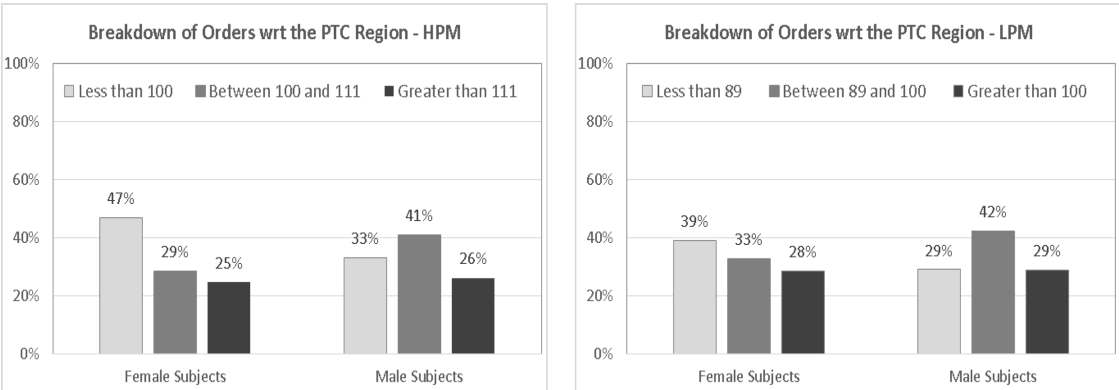


Figure 18: Percentage of orders in and out of the PTC region

We observe most of the order decisions to fall outside the PTC region under both profit margins and for both genders, indicating that the pull-to-center effect is not strongly supported on the individual decision level. Nevertheless, the classification results shown in Table 52 point to a significant difference between genders: Female subjects have more orders below the PTC region and less orders above the PTC region under both profit margins than male subjects do. This result is consistent with females undertaking less risk by placing smaller orders. We also observe that males place more orders in the PTC region than females, which suggests that males might be using the mean-anchor heuristic more than females do.

Table 52: Gender comparison of orders in and out of the PTC region

		Female Subjects			Male Subjects			P Value
		Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	
HPM	Less than 100	47%	24%	49%	33%	23%	38%	**0.003**
	Between 100 and 111	29%	22%	23%	41%	28%	30%	0.011
	Greater than 111	25%	21%	21%	26%	19%	25%	0.498
LPM	Less than 89	39%	21%	35%	29%	19%	30%	**0.004**
	Between 89 and 100	33%	18%	33%	42%	24%	38%	**0.017**
	Greater than 100	28%	18%	25%	29%	21%	28%	0.878

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

Bostian et al. (2009) suggests measuring the mean anchor heuristic by evaluating the regression equation $Q = \mu + \alpha(Q^* - \mu) + \varepsilon$. In this equation, α being between 0 and 1 translates to the expectation (with respect to ε) of the order quantity being between the demand mean and the optimal order quantity. If α is less than 0, the expected order quantity falls to the other side of the demand mean, further away from the optimal. Table 53 tabulates the descriptive statistics of the estimated α coefficients for our experiments.

Table 53: Gender comparison of estimated mean anchor heuristic α coefficients

	Female Subjects			Male Subjects			P Value
	Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	
HPM	-0.07	0.87	-0.06	0.31	0.68	0.28	**0.058**
LPM	0.55	0.87	0.45	0.31	0.68	0.38	0.248

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

Table 53 shows female α coefficients in our data to be significantly smaller than male coefficients under HPM. Thus, the demand mean has stronger pull over the female order decisions than male decisions, and female orders are smaller. Under LPM, the optimal order quantity has stronger effect on female order quantities as coefficients are observed to be higher. However, due to higher variance in the female coefficients, the difference is not significant.

The above results are not sufficient to draw a conclusion on the use of the mean-anchor heuristic. Table 54 exhibits the individual level comparison of the mean anchor α coefficients at 10% significance level. According to these results, there are more male subjects that can be said to follow the mean-anchor heuristic (i.e., α being between 0 and 1) than female subjects.

Table 54: Individual level comparison of mean anchor α coefficients

		Female Subjects			**Male Subjects**		
		$0 \le \alpha < 1$	$\alpha < 0$	$\alpha \ge 1$	$0 \le \alpha < 1$	$\alpha < 0$	$\alpha \ge 1$
HPM	All	39%	56%	6%	47%	35%	18%
	Significant	22%	0%	6%	29%	0%	18%
LPM	All	51%	27%	22%	59%	26%	15%
	Significant	33%	0%	22%	38%	0%	15%

Another individual level analysis utilized by Lau et al. (2014) is presented in Table 55. This analysis compares the percentage of female and male subjects whose modal order is at the demand mean or at the optimal quantity. Because the optimal orders are 111 and 89 for the high and low profit margins respectively, to account for any rounding behavior in the ordering decisions, we consider 110 and 90 as optimal as well.

Table 55: Percentage of subjects with modal order of 100 and $Q*$

	HPM Females	HPM Males	LPM Females	LPM Males
Modal Order 100	41% (#22)	35% (#18)	33% (#22)	24% (#20)
Modal Order Q*	15% (#8)	20% (#10)	16% (#11)	26% (#22)

Under both profit margins, there are more female subjects than male subjects whose modal order quantity is the mean demand. Additionally, the optimal order quantity $Q*$ seem to be the modal order for more male subjects than female subjects under both profit margins.

All in all, these results do not show a clear indication that female subjects resort to the mean-anchor heuristic more or less than male subjects do. What is clear is that experimental results need to be analyzed carefully, as different types of analysis can lead to opposing conclusions. Hence the support for Hypothesis 3 is inconclusive.

6.4.3.2 Demand Chasing Heuristic

The demand chasing heuristic is a special type of anchoring-and-adjustment heuristics, where the individuals anchor at their most recent order quantity and make adjustments towards the most recent demand realization. This heuristic, where the random nature of the demand plays a key role, is closely related to decision biases pertinent to probability and random events such as gambler's fallacy and hot hand fallacy. Gambler's fallacy is the bias where the decision maker assumes a series of random events to be negatively correlated (Rabin and Vayanos 2010). In our context, this translates to the decision maker expecting a low (high) demand following a high (low) demand realization. Hot hand fallacy is the opposite of gambler's fallacy, where the decision maker assumes that the correlation between consecutive random events is positive (Rabin and Vayanos 2010). Again in our context, this translates to expecting another a high (low) demand realization following a high (low) demand realization.

No study in literature reports gender differences in demand chasing behavior. However, several studies report gender differences for gambler's fallacy and hot hand fallacy. Following Kudryavtsev and Cohen (2011)'s study on gender differences in anchoring and hindsight bias, Kudryavtsev et al. (2013) conduct a survey-type study focusing on five decision heuristics, namely availability heuristic, disposition effect, gambler's fallacy, herd behavior and hot hand fallacy. These authors find all five decision heuristics to be stronger in women. Huber et al. (2010) obtain conflicting results in their financial decision making experiment where no significant difference between genders is reported Stöckl et al. (2014), who repeat Huber et al. (2010)'s experiment with individuals and teams of two, find that females exhibit higher tendency to fall victim to hot hand fallacy both as individuals and as groups. The authors find no significant difference for the gambler's fallacy. Suetens and Tyran (2012)'s study on Danish state lottery indicates that males exhibit a significant gambler's fallacy while females don't.

Despite the above conflicting results we expect females to chase the demand more than males do. Our reasoning is that the context used in Kudryavtsev et al. (2013)'s study is the most similar to ours. Huber et al. (2010) and Stöckl et al. (2014)'s studies use coin tosses as random events. Besides, while Huber et al. find no gender difference, Stöckl et al. report significant gender differences. Suetens and Tyran (2012)'s study is based on a lottery for which the individuals pick some numbers out of 36, which is similar to trying to point estimate the winning numbers. However, Kudryavtsev et al. (2013) ask the subjects to evaluate the likelihood of stock market indices increasing or decreasing given that these indices have been increasing for the last six months. The movements of stock market indices resemble the random movement of the demand realizations in our context.

Hypothesis 4: *Females will chase demand more than males do.*

Figure 19 displays the evolution of the average order quantity in our experiment over time. We observe the average order quantity decisions to follow the rise and falls of the demand realization. This behavior can be quantified and measured by several methods proposed by various studies. The first method, which is used by Schweitzer and Cachon (2000), Kremer et al. (2010) and Benzion et al. (2008) is to find the proportion of the order decisions that are changed towards or away from the previous period's demand realization.

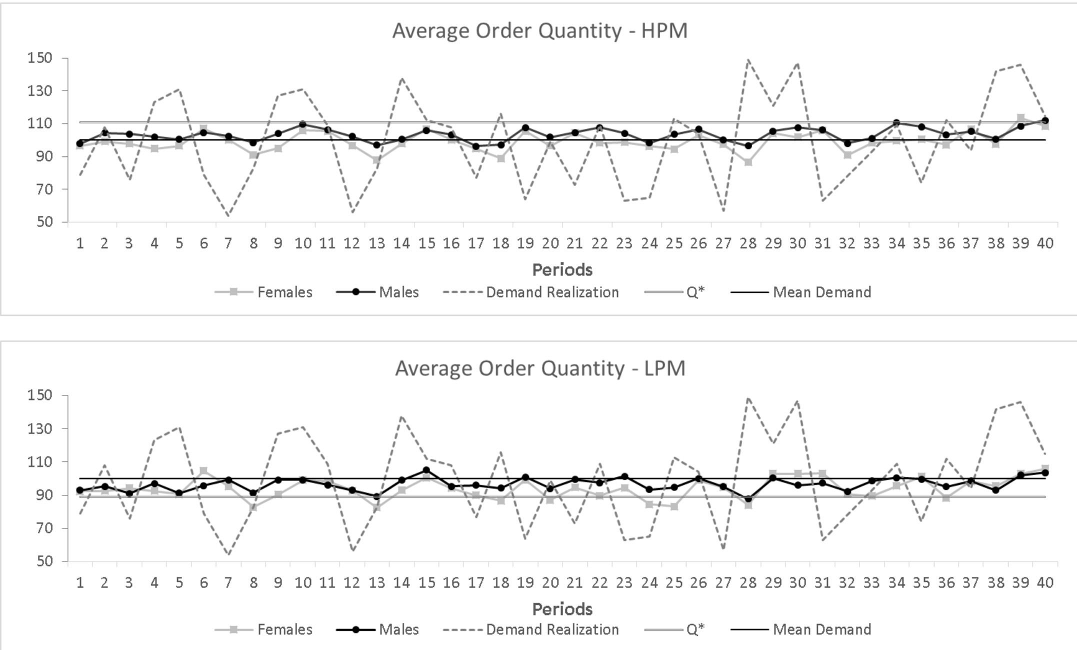


Figure 19: Evolution of the average order quantity over time

Table 13 presents the descriptive statistics and comparison results of the changes in the order decisions in our data. Under both profit margins and for both genders, majority of orders were changed towards the most recent demand realization. This tendency is stronger in the female subjects, and the difference is significant in the LPM treatment. On the other hand, there is not much gender difference in the orders that were changed away from the last period's demand. Lastly, we observe male subjects' repeat order tendency to be higher than female subjects, indicating male order decisions to be more stable. This result is in parallel with our finding in Section 4.2 that male subjects have smaller order variability than females.

Table 56: Percentage of orders changed towards and away from the previous demand

		Female Subjects			Male Subjects			P Value
		Mean	**Std. Dev.**	**Median**	**Mean**	**Std. Dev.**	**Median**	
HPM	Towards the Demand	61%	21%	65%	49%	25%	59%	**0.010**
	Away from the Demand	24%	12%	22%	22%	14%	23%	0.719
	No Change	15%	24%	5%	28%	35%	13%	**0.012**
LPM	Towards the Demand	64%	16%	64%	52%	20%	54%	**0.000**
	Away from the Demand	25%	12%	23%	26%	13%	26%	0.515
	No Change	11%	17%	5%	22%	26%	13%	**0.001**

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

Another method to quantify and measure the demand chasing behavior (Bostian et al. 2008) is to evaluate the regression equation $Q_t = Q_{t-1} + \beta(D_{t-1} - Q_{t-1}) + \varepsilon$. Here, the closer β is to 1, the stronger the effect of the previous period's demand realization on the order decision is. Table 57 presents the estimation results with our data. The higher β coefficients imply demand realization to have more effect on the female subjects' orders than males' under both HPM and LPM.

Table 57: Gender comparison of demand chasing regression estimations

	Female Subjects			Male Subjects			P Value
	Mean	**Std. Dev.**	**Median**	**Mean**	**Std. Dev.**	**Median**	
HPM	0.35	0.22	0.35	0.25	0.19	0.21	**0.022**
LPM	0.35	0.20	0.34	0.24	0.18	0.23	**0.000**

P-values are obtained from two tailed Mann Whitney U tests. Sample size is the number of subjects.

These aggregate level results indicate that the demand chasing heuristic has more impact on the female subjects than on males.

The third method involves the analysis of correlation of the individual order quantities with the previous period's demand realization (Bolton and Katok 2008, Lau et al. 2013). For our data, this method yields results in parallel with the second method. Table 58 presents the percentage of female and male subjects whose order decisions are correlated with the demand at 10% significance level.

Table 58: Percentage of subjects with orders correlated with the demand

	HPM Females	HPM Males	LPM Females	LPM Males
Positively correlated	54% (#29)	39% (#20)	57% (#38)	39% (#33)
Negatively correlated	2% (#1)	4% (#2)	1% (#1)	8% (#7)

Based on the aforementioned results from three alternative methods, female subjects seem to chase demand more than male subjects do in both the LPM and HPM treatments and at both aggregate and individual levels. Thus, Hypothesis 4 is supported.

Next we develop a fixed-effects regression model to quantify and compare the effects of previous period's demand realization and order quantity on the ordering decisions at an aggregate level.

6.4.4 Aggregate Level Analysis

In this section, in order to explain the gender differences in ordering behavior we develop the following fixed-effects regression model:

$$Q_{it} = \alpha + \beta_1 D_{t-1} + \beta_2 \mathrm{Q}_{i,t-1} + \beta_3 (t-1) + \eta_i + \varepsilon_{it}$$

Here the dependent variable is the order quantity at period *t*. The explanatory variables are the previous period's demand realization and order quantity. Additionally, the time variable is included into the model to check for any learning effect. The term η_i denotes the variability due to personal variations in the data, and ε_{it} is the common error term. The analysis is run separately for each profit margin and each gender over pooled data. Estimation results are presented in Table 59.

Table 59: Aggregate regression results

	HPM		LPM	
	Females	**Males**	**Females**	**Males**
Variables	**Coeff. (Std. Dev)**	**Coeff. (Std. Dev)**	**Coeff. (Std. Dev)**	**Coeff. (Std. Dev)**
Intercept	73.91 (4.91)***	80.18 (4.67) ***	51.59 (4.79)***	64.66 (4.48)***
Previous Demand	0.17 (0.03)***	0.08 (0.03)**	0.20 (0.02)***	0.07 (0.02)***
Previous Order	0.07 (0.04)*	0.14 (0.04)**	0.23 (0.05)***	0.24 (0.04)***
Time	0.05 (0.04)	0.05 (0.04)	0.04 (0.04)	0.05 (0.03)

**** P-value <0.001, ** P-value <0.01, * P-value <0.1*

Previous period's demand realization is found to be positively correlated with the order decisions of all groups, yet the coefficient is greater for female subjects under both profit margins. This implies, again in agreement with our previous results, that demand realization has stronger effect on female orders. Previous period's order quantity also appears to be significant for all groups. However the impact is much larger under low profit margin. This relatively high weight on the most recent order might be the reason why order decisions are usually higher than the optimal. Lastly, subjects' decisions do not seem to change significantly with time.

6.5 Conclusions

In this paper we investigate gender differences in newsvendor ordering behavior through an experimental study. In addition to comparing the order quantities and achieved profits, we also compare the two genders in their use of decisions heuristics. Our results show that female subjects place smaller orders, and base their decisions on the demand realization and on their most recent order quantity more than male subjects do.

Interestingly, 11 out of 257 subjects were able to identify the setting of the experiment as the newsvendor problem and compute the optimal order quantity. Moreover these subjects were not affected by the random demand realizations and placed the same, optimal, order quantity for 40 periods. These subjects were all male. This may be explained by the selectivity hypothesis. While female subjects focused on the details and tried to take into account every single bit of information, these male subjects may have selectively regarded the information pertaining to the model and disregarded all later information.

These results definitely do not imply males to be better newsvendor decision-makers than females. In fact, the smaller order quantities of females would become an advantage under a sufficiently low profit margin setting. Rather, the findings may provide guidance into the design of decision support tools for inventory and planning managers. For instance, to encourage higher order decisions, when necessary, female decision makers can be provided with stronger incentives, such as sales-based awards. Again, when needed, females can be encouraged to undertake more risk through self-improvement seminars. Or, to prevent potential demand chasing behavior, the update frequency of order decisions can be reduced.

We conduct a single-player experiment to address gender differences. Game theory literature has demonstrated that two-player games in which the players are informed of their opponent's gender reveal interesting gender differences. For instance, Solnick (2001) show that in an ultimatum game, while the gender of the first player (the proposer) doesn't affect the offer quantity, female second players are offered smaller values than males. Additionally, females are found to have higher minimum acceptance level than males. Hence as a further research opportunity, a two-player supply chain contracting experiment focusing on the effect of gender on the pricing and ordering decisions can be conducted.

6.6 References

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

Bolton, G. E., Katok, E. (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback. Manufacturing and Service Operations Management, 10(3), 519-538.

Bostian, AJ. A, Holt, C. A., Smith, A. M. (2008). Newsvendor Pull-to-Center Effect: Adaptive Learning in a Laboratory Experiment. Manufacturing and Service Operations Management, 10(4), 590-608.

Byrnes, J.P., Miller, D.C., Schafer, W.D. (1999). Gender differences in risk taking: a meta-analysis. Psychological Bulletin, 125, 367–383.

Cervone, D., Peake, P. K. (1986). Anchoring, Efficacy and Action: The Influence of Judgmental Heuristics on Self-Efficacy Judgments and Behavior. Journal of Personality and Social Psychology, 50 (3), 492-501.

Croson, R., Gneezy, U. (2009). Gender differences in preferences. Journal of Economic Literature, 47, 1–27.

De Vericourt, F. D., Bearden, J. N., Filipowicz, A. (2013). Sex, Risk and the Newsvendor. Journal of Operations Management, 31, 86-92.

Huber, J., Kirchler, M., Stöckl, T. (2010). The hot hand belief and the gambler’s fallacy in investment decisions under risk. Theory and Decision, 68, 445-462.

Hügelschäfer, S., Achtziger, A. (2014) On confident men and rational women: It’s all on your mind(set). Journal of Economic Psychology, 41, 31-44.

Kahneman, D., Slovic, P., Tversky, A. (1982). Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge.

Kremer, M., Minner, S., Van Wassenhove, L. N. (2010). Do Random Errors Explain Newsvendor Behavior? Manufacturing and Service Operations Management, 12(4), 673-681.

Kudryavtsev, A., Cohen, G. (2011). Behavioral biases in economic and financial knowledge: Are they the same for men and women? Advances in Management & Applied Economics, 1(1),15-52.

Kudryavtsev, A., Cohen, G., Hon-Snir, S. (2013). 'Rational' or 'Intuitive': Are Behavioral Biases Correlated Across Stock Market Investors? Contemporary Economics, 7(2), 31-53.

Lau, N., Bearden J. N. (2013). Newsvendor Demand Chasing Revisited. Management Science, 59(5), 1245-1249.

Lau, N., Bearden, J. N., Hasija, S. (2014). Newsvendor Pull-to-Center Reconsidered. Decision Support Systems, 58, 68-73.

Lee, Y. S., Siemsen E. (2016). Task Decomposition and Newsvendor Decision Making. Working Paper.

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

Meyers-Levy, J., Loken, B. (2015). Revisiting gender differences: What we know and what lies ahead. Journal of Consumer Psychology, 25(1), 129-149.

Rabin, M., Vayanos, D. (2010). The Gambler's and Hot-Hand Fallacies: Theory and Applications. The Review of Economic Studies, 77(2), 730-778.

Schubert, R., Brown M., Gysler, M., Brachinger, H. W. (1999). Financial Decision-Making: Are Women Really More Risk-Averse? American Economic Review, 89(2), 381-385.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Solnick, S. J. (2001) Gender Differences in the Ultimatum Game. Economic Inquiry, 39 (2), 189-200.

Suetens, S., Tyran, J.R. (2012). The gambler’s fallacy and gender. Journal of Economic Behavior and Organization, 83, 118-124.

Weber, E.U., Blais, A., Betz, E.N. (2002). A domain-specific risk attitude scale: measuring risk perceptions and risk behaviors. Journal of Behavioral Decision Making, 15, 262–290.

Wu, T., Seidman, A. (2014). Do Competitive Environments have an Effect on Managerial Decision Making? An Empirical Investigation of the Newsvendor Problem. 47th Hawaii International Conference on System Science.

6.7 Appendix: Instructions for Experiment (High Profit Margin)

Thanks for participating in this experiment. Our objective is to understand how human beings make decisions under uncertainty.

When you arrive at the recitation, we will ask you to:

1) Go to http://myweb.sabanciuniv.edu/ummuhana/experiment and follow the instructions.
2) After downloading the MS Excel file, open it and enable macros. (Select "Enable content" on the warning bar options)
3) Save the MS Excel file in the following format "YourStudentNumber.xlsm".
4) Enter your name and ID to related cells (cells C15 and C16).

Figure 1 displays a screen shot of the Excel file.

Imagine you are a retailer that orders swimsuits from a manufacturer and sells them during the summer sales season to consumers. You need to determine how much to order from the manufacturer before the sales season. You pay the manufacturer **$35 wholesale price** per unit of order.

You sell the swimsuits to consumers for **$90 per unit**. You do not know how many swimsuits will be demanded during the season, but you know that the demand will be **uniformly distributed between 50 and 150**. This means that there is an equal probability that your demand will be any integer between 50 and 150. If demand turns out to be larger than your **order quantity**, you will run out of swimsuits during the sales season and you will not be able to meet the demand of some consumers. If demand turns out to be lower than your order quantity, you will have unsold inventory at the end of the season. These units have zero value and they are disposed of.

The experiment consists of **40 seasons** played in an MS Excel sheet. Your objective is to maximize your total profit over the 40 seasons. At each season, you will determine your order quantity (an integer between 50 and 150) and enter it using the "Place Order" button. You will not be able to change your order once you enter it. After you click OK, the demand realization of that period will appear on the corresponding cell. The Excel sheet will calculate and display the outcome of that season in related yellow cells. Then, the experiment will proceed to the next season. The seasons are independent of each other. That is, once you are done with one season, the next season does not have any interaction with the previous one. Recall that unsold inventory at the end of a season has no value and it cannot be used to meet demand in a subsequent season.

Please pay attention to how we use colors in the Excel sheet: The input cells (where your order quantities and demand realizations will appear) are blue, parameter cells are green and output cells are yellow (See figure 1).

Your activities on the Excel sheet are restricted by a macro. Our goal is not to judge your mathematical skills, but to understand how you make decisions under uncertainty.

You will go through 3 warm-up seasons (Periods -2, -1 and 0) to make sure that you understand how the experiment works. Your profit during these seasons will not be taken into account. Note that your profit at each session depends on your order quantity and the realization of demand. The two examples below illustrate this calculation:

- Assume that you ordered 80 units in a season, for which you pay $35 * 80 = $2800 to the manufacturer. Suppose that the demand realization in that season is 60 units. In this case, you will sell 60 units and 20 units will be left over. Your sales revenue will be $90 * 60 = $5400, and your profit will be $5400 - $2800 = $2600.

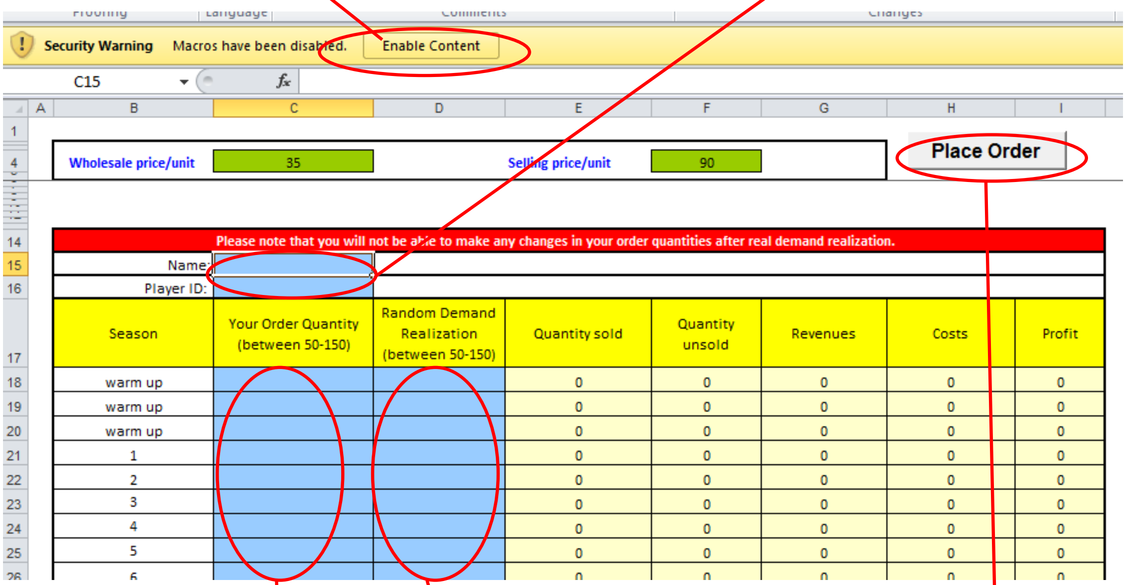
- Assume that you ordered 60 units in a season, for which you pay $35 * 60 = $2100 to the manufacturer. Suppose that the demand realization in that season is 135 units. In this case, you will sell all of the 60 units, and you will not be able to meet the remaining demand. Your sales revenue will be $90 * 60 = $5400 and your profit will be $5400 -$2100 = $3300. In this case, your profit would have been higher if you had ordered more units.

Below is the screenshot from the Excel sheet.

Please save the file with your student ID in the filename enter your name and student ID here.

Click to enable macros.

Figure 20 Screenshot of Excel sheet



To place order click the "Place Order" button, and a window will pop-up, where you will enter your order quantity.

(Blue) Your order quantities will appear here.

(Blue) The realizations of the demand will appear here.

Chapter 7

7. EFFECTS OF PERSONALITY TRAITS ON SUPPLY CHAIN DECISIONS

Effects of Personality Traits on Supply Chain Decisions

7.1 Introduction

In this chapter we study the interaction between experiment performance and subjects' personality traits which we measure through out-of-experiment surveys. The goal of this study is to investigate if the experiment behavior of individual decision makers can be estimated through personality traits measures. This way, if suboptimal behavior can be associated with certain traits, corrective or preventive measures can be taken in order to promote more efficient behavior.

We present results from two separate experimental studies. The first one is the newsvendor experiment study of Chapter 6. The second one is the contracting study of Chapters 3 and 5, which is based on a one-manufacturer-one-retailer supply chain where the retailer faces probabilistic consumer demand.

The sample size for each treatment is 21-22, and this number is divided into two or three according to the answers given to the survey questions. Despite this reduction in sample size, we were able to find meaningful differences in subject decisions.

The rest of the chapter is organized as follows: In Section 2 we summarize the related literature. In Section 3 we briefly explain the experimental design. Sections 4 through 7 present the research hypotheses and analysis results for self-esteem, regret tendency, risk-loss aversion and inequity aversion traits.

7.2 Related Literature

Earlier behavioral operations studies, such as Schweitzer and Cachon (2000), Keser and Paleologo (2004), and Bolton and Katok (2008) focused on aggregate level analysis. Later, individual heterogeneity of the subjects came into attention and behavioral operations studies started to include more individual level analyses. Lau et al. (2014), for instance, show that even though aggregate data may seem to support the pull-to-center effect, data from individual decision makers exhibit high heterogeneity and do not necessarily follow the pull-to-center effect.

Moritz (2010) and Moritz et al. (2013) suggest that individual heterogeneity of the subjects may be resulting from their cognitive reflection, major, years of experience, and managerial position. The authors show that cognitive reflection is closely related to the newsvendor performance. Individuals with high cognitive reflection test scores place better orders decisions and earn higher profit values.

Strohhecker and Grössler (2013) study the effect of intelligence, knowledge, personality and interest on inventory decisions. The authors base their study on a complex inventory management scenario and show that intelligence is the strongest factor affecting the quality of the decisions, and there is weaker but significant correlation between other traits and inventory performance.

To the best of our knowledge, no behavioral operations article other than the current one studies the effect of self-esteem, regret-aversion, risk and loss aversion, and inequity aversion on newsvendor decisions

7.3 Experimental Design

In this chapter, we use the data from the newsvendor experiment of **Chapter 6**, and the contracting experiments of **Chapter 3** and **Chapter 5**. Detailed information can be found in the mentioned chapters, here we briefly list the treatments.

Newsvendor Experiment

Subjects playing against a computer, facing a fixed wholesale price and random consumer demand. Two treatments: high profit margin (HPM) and low profit margin (LPM). Relevant surveys are self-esteem scale, regret tendency scale, risk and loss aversion scales and DOSPERT (domain specific risk taking) scale. Number of subjects answering each survey varies and is presented in Table 60.

Table 60: Number of subjects participating in the survey study

	Self-Esteem	Regret Tendency	Risk/Loss Aversion	DOSPERT
HPM	57	57	57	57
LPM	148	148	125	145

Contracting Experiment

Pair of subjects playing against each other. Manufacturer subjects determine the contract parameters and retailer subjects determine order quantity. Six treatments, three of which focus on contract type. These are wholesale price (WSP), buyback (BB) and revenue sharing (RS) contract treatments. Remaining three treatments focus on manipulations of the wholesale price contract, namely standing contract (SC), standing contract and order (SCO), and fairness priming (FP). Number of manufacturer-retailer pairs answering each survey is presented in Table 61.

Table 61: Number of pairs participating in the survey study for each treatment

	Self-Esteem	Regret Tendency	Risk/Loss Aversion	Inequity Aversion
WSP	22	22	22	22
BB	22	22	22	22
RS	21	21	21	21
SC	22	22	22	22
SCQ	22	22	22	22
FP	22	22	22	22

7.4 Self-Esteem Analysis

In psychology self-esteem is an individual's subjective and emotional assessment and appreciation of their self-worth. Self-esteem has been shown to be closely related to various behavioral aspects. Josephs et al. (1992) examine the connection between self-esteem and risky decision making. In an experiment where the subjects are faced with risky gambles, the authors observe that subjects with low self-esteem avoid risk and ambiguity in an effort to protect themselves from the threat of a negative outcome more than subjects with high self-esteem do. Building on these results, we expect subjects with higher self-esteem scores to undertake higher risk and place larger orders.

Hypothesis 1a: *Under both profit margins, newsvendors with high self-esteem scores will place larger orders.*

We expect subjects with high self-esteem to be more confident in their order decisions and to be more assertive. This way, they will change their order decisions less than subjects with low self-esteem.

Hypothesis 1b: *Newsvendors with high self-esteem scores will have smaller order variance.*

Furthermore, as we expect the high self-esteemed subjects to be more confident in their ordering decisions, they shall attach less importance to random demand realizations. Consequently, their order decisions shall be less affected by demand compared to the orders of subjects with low self-esteem.

Hypothesis 1c: *Newsvendors with high self-esteem scores will be less affected by the demand chasing behavior.*

We expect similar more confident, assertive and risk-taking behavior from the high-esteem retailers in the contracting experiments. Because the contract parameters (and as a result, the optimal order quantity) are determined by the manufacturer, and can change from period to period, we compare subjects with respect to average ratio of the order quantity decision to the optimal order quantity.

Hypothesis 1d: *Retailers with high self-esteem scores will have higher Q/Q* ratio.*

It is not straightforward to identify how much of the change in the order quantity from one period to the next is caused by the change in the contract, random demand realization or self-esteem of the retailer. Hence, we won't compare retailers' order variability or demand chasing behavior. However, we expect retailers with high self-esteem scores to demand and extract more of the total chain profits.

Hypothesis 1e: *Retailers with high self-esteem scores will have higher expected profit.*

Similarly, we expect a higher self-esteem score to translate to a more assertive, more risk-taking manufacturer who will offer higher prices, and claim higher profit and profit share.

Here the newsvendor-predicted profit and profit share are better measures as they only depend on the manufacturer's decisions.

Hypothesis 1f: *Manufacturers with high self-esteem scores will have higher newsvendor predicted profit and profit shares.*

To test these hypotheses we asked the subjects to answer Rosenberg's self-esteem scale questions presented in Table 62. This scale, a well-known tool for measuring self-esteem is developed by Rosenberg (1965). The scale consists of five positive and five negative questions. Evaluation of the answers is displayed in Table 63.

Table 62: Rosenberg's self-esteem scale

Below is a list of statements related with your general feelings about yourself. Please read each statement carefully and select the option which describes how you feel about yourself best.
1. I feel that I am a person of worth, at least on an equal plane with others. 2. I feel that I have a number of good qualities. 3. All in all, I am inclined to feel that I am a failure. 4. I am able to do things as well as most other people. 5. I feel I do not have much to be proud of. 6. I take a positive attitude toward myself. 7. On the whole, I am satisfied with myself. 8. I wish I could have more respect for myself. 9. I certainly feel useless at times. 10. At times I think I am no good at all.

Table 63: Evaluation of Rosenberg's self-esteem scale

Questions 1, 2, 4, 6, 7 (positive questions)	**Questions 3, 5, 8, 9, 10 (negative questions)**
Strongly Agree: 3	Strongly Agree: 0
Agree: 2	Agree: 1
Disagree: 1	Disagree: 2
Strongly Disagree: 0	Strongly Disagree: 3

Scores from all questions are summed and averaged to find the overall self-esteem score. In the standard test, this score is compared with the threshold values of 1.5 and 2.5 to categorize each subject into low, normal or high self-esteem categories. However, in our study, we divide subjects into two groups as low and high, corresponding to those that have lower and higher than average self-esteem score in their respective treatments.

In order to check the validity of our survey results, we compare our findings with the findings of two other studies measuring Rosenberg's self-esteem scores of Turkish subjects. The average self-esteem score of our whole subject pool (n=459) is 2.24 and the median is 2.3. In Schmitt and Allik (2005), the average self-esteem score of 409 Turkish participants is found to be 2.214. Another study by Akpınar et al. (2014) reports the average self-esteem score of participants of ages from 19 to 24 as 2.24. These findings support the findings of our self-esteem surveys.

7.4.1 Newsvendor Experiment Results

For each newsvendor, we compute the average order quantity, average profit, average expected profit, and order standard deviation over 40 periods. To check for the demand chasing behavior, we calculate the correlation of order decisions with the previous periods' demand realization. These computations constitute the unit of analysis in our comparison of high and low self-esteem subjects.

Comparison results are shown in Table 64. Under HPM, high self-esteem subjects are observed to have larger orders and higher profit values, however the differences under LPM are not significant. In addition, under HPM, high self-esteem subjects have lower order variability, indicating they are more confident in their order decisions. Under LPM, high self-esteem subjects have lower order correlation with the demand, suggesting that there may be a negative relation between demand chasing behavior and self-esteem.

Table 64: Self-esteem comparison results of the newsvendor experiment

	Self Esteem Score	Sample Size		Avg. Order	Avg. Profit	Avg. Exp. Profit	Order Std. Dev.	Corr. w/ Demand
HPM	SE<2.25	30	Mean	99.34	4204.00	4158.61	17.83	0.23
			Median	99.33	4197.56	4179.65	17.98	0.22
			Std. Dev.	9.93	156.36	148.56	6.52	0.32
	SE≥2.25	27	Mean	103.00	4278.93	4229.61	13.79	0.33

			Median	103.35	4330.25	4290.16	14.22	0.32
			Std. Dev.	10.02	219.96	148.83	9.08	0.30
	P-values			**0.03**	**0.01**	**0.03**	**0.02**	0.15
LPM	**SE<2.25**	73	**Mean**	95.55	2269.30	2224.21	17.19	0.28
			Median	95.75	2308.21	2263.71	16.49	0.29
			Std. Dev.	7.67	163.16	147.82	7.03	0.26
	SE≥2.25	75	**Mean**	95.03	2265.52	2217.82	16.78	0.17
			Median	95.05	2279.87	2229.73	17.55	0.18
			Std. Dev.	9.41	168.35	131.58	7.40	0.30
		P-values		0.36	0.43	0.22	0.43	**0.02**

Hence, *Hypotheses 1a – 1c* are partially supported. Additionally, the design of the experiment (i.e. high-low profit margin) is seen to be a factor affecting the impact of self-esteem on the ordering behavior.

7.4.2 Contracting Experiment Results

For each retailer we compute the average *Q/Q** ratio, expected profit, realized profit and realized profit share. The comparison results are shown in Table 65. Our analysis reveals that retailers with high self-esteem have higher average *Q/Q** ratio under RS, FP and SP treatments. In addition, high self-esteem subjects are found to have higher expected and realized profit values and higher realized profit shares under WSP, SP and SPQ treatments.

Hence *Hypotheses 1d and 1e* are partially supported. And we observe that contract type and design of the experiment are also affective on the impact of self-esteem.

Table 65: Self-esteem comparison results of the contracting experiment retailers

Self Esteem Score		**WSP**					**BB**				
		#	**Q/Q***	**Exp. Profit**	**Real. Profit**	**Real. Profit Share**	**#**	**Q/Q***	**Exp. Profit**	**Real. Profit**	**Real. Profit Share**
SE<2.25	**Mean**	9	1.12	257.84	281.01	0.33	9	1.00	250.46	264.23	0.33
	Median		1.10	270.02	290.08	0.33		1.01	223.40	246.80	0.31
	Std. Dev.		0.13	46.20	50.43	0.09		0.13	84.81	87.46	0.13
SE≥2.25	**Mean**	13	1.07	298.42	321.30	0.40	13	1.02	224.34	237.26	0.29
	Median		1.09	296.75	322.63	0.42		0.97	212.84	225.64	0.30
	Std. Dev.		0.16	55.31	55.03	0.08		0.16	78.10	79.39	0.12
	P-value		0.29	**0.05**	**0.08**	**0.04**		0.46	0.29	0.24	0.24

Self Esteem Score		RS					FP				
		#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share	#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share
SE<2.25	**Mean**	13	0.94	276.81	286.60	0.38	10	1.12	262.68	287.59	0.35
	Median		0.94	299.54	335.88	0.44		1.12	285.64	306.58	0.37
	Std. Dev.		0.07	96.52	97.79	0.12		0.15	99.61	99.69	0.14
SE≥2.25	**Mean**	8	1.08	219.43	236.59	0.28	12	1.14	219.84	237.02	0.28
	Median		1.01	230.08	245.31	0.30		1.14	235.02	253.45	0.31
	Std. Dev.		0.17	107.88	118.55	0.16		0.10	71.69	79.71	0.11
	P-value		**0.01**	0.14	0.21	0.11		0.25	0.11	**0.08**	**0.09**

Self Esteem Score		SP					SPQ				
		#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share	#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share
SE<2.25	**Mean**	8	1.09	225.68	243.04	0.31	13	1.14	216.79	235.43	0.29
	Median		1.07	236.26	252.36	0.35		1.16	219.85	238.45	0.30
	Std. Dev.		0.18	59.73	62.13	0.12		0.20	94.85	94.82	0.16
SE≥2.25	**Mean**	14	1.13	247.86	263.83	0.31	9	1.02	309.03	326.54	0.42
	Median		1.09	253.99	283.03	0.30		1.01	279.13	303.35	0.41
	Std. Dev.		0.12	82.02	84.15	0.11		0.07	86.31	92.17	0.11
	P-value		0.29	0.29	0.21	0.39		**0.08**	**0.04**	**0.04**	**0.03**

For the manufacturers, we compute and compare the average predicted profit and predicted profit shares. From the comparison results presented in Table 66, we see that manufacturers with high self-esteem scores have higher predicted profit and predicted profit share under WSP, BB and RS contracts. Thus *Hypothesis 1f* is partially supported.

Table 66: Self-esteem comparison results of the contracting experiment manufacturers

Self Esteem Score		WSP			BB			RS		
		#	Pred. Profit	Pred. Profit Share	#	Pred. Profit	Pred. Profit Share	#	Pred. Profit	Pred. Profit Share
SE<2.25	**Mean**	12	383.94	0.55	8	473.75	0.66	14	427.93	0.59
	Median		391.96	0.57		464.70	0.68		425.66	0.56
	Std. Dev.		25.83	0.05		57.56	0.08		53.26	0.09
SE≥2.25	**Mean**	10	396.61	0.58	14	489.95	0.68	7	506.73	0.75
	Median		403.63	0.59		496.01	0.69		506.72	0.71
	Std. Dev.		31.07	0.06		53.70	0.09		44.63	0.07
	P-value		0.12	0.13		0.29	0.37		**0.00**	**0.00**

		FP			SP			SPQ		

Self Esteem Score		#	Pred. Profit	Pred. Profit Share	#	Pred. Profit	Pred. Profit Share	#	Pred. Profit	Pred. Profit Share
SE<2.25	**Mean**	16	417.33	0.65	10	413.99	0.63	15	409.84	0.62
	Median		416.27	0.64		410.13	0.63		412.00	0.61
	Std. Dev.		30.96	0.09		26.35	0.07		38.79	0.10
SE≥2.25	**Mean**	6	402.13	0.62	12	409.47	0.62	7	393.04	0.60
	Median		420.53	0.63		418.75	0.63		404.13	0.59
	Std. Dev.		54.91	0.13		40.63	0.08		54.57	0.14
	P-value		0.33	0.30		0.47	0.50		0.27	0.32

7.5 Regret Aversion Analysis

Researchers have been studying effects of emotions on decision making process. One of these emotions is regret. Regret is the conscious and emotional state of feeling sad and disappointed for losses or mistakes. Researchers have shown that regret can shape the human behavior in two ways; before and after the decision is made. After the decision is made and resulted in an undesirable outcome, individuals may act as to undo their action (Gilovich and Medvec 1995). Before the decision is made, individuals may anticipate regret and try to minimize the regret they may feel later (Bell 1982). Some experimental studies show that anticipated regret leads to risk-aversion (Josephs et al. 1992). Contrary to those studies, Zeelenberg et al. (1996) show that in the face of anticipated regret, individuals make regret-minimizing choices rather than risk minimizing choices.

For newsvendors decisions regret is related to inventory error, which is the difference between the demand realization and order quantity. Schweitzer and Cachon (2000) propose the “minimization of ex-post inventory error” heuristic as one alternative to explain the pull-to-center effect observed in newsvendor decisions. That is, subjects’ order quantity decisions may be pulled towards the center of the demand distribution because of their avoidance of inventory error.

Here we assume that all subjects have regret aversion, and what differentiates them is the degree of regret they might feel after making a decision. We refer to this as the regret tendency of a subject. We expect subjects with high regret tendency to act in a more regret averse fashion and be more prone to the pull-to-center effect.

Hypothesis 2a*: Newsvendors with high regret tendency will be more affected by the pull-to-center effect.*

We expect similar behavior from the retailers in the supply chain contracting experiment.

Hypothesis 2b*: Retailers with high regret tendency will be more affected by the pull-to-center effect.*

Zeelenberg and Beattie (1997) discuss that in an ultimatum game, there are two types of regrets associated with the proposer's decisions, namely, offering too little causing the offer being rejected; and offering too much. The authors argue that in such a setting the regret from offering too little and getting rejected dominates the regret from offering too much. Hence they hypothesize and show that regret aversion leads to higher offers.

Recall that our contracting setting is similar to an ultimatum game, in the sense that the manufacturer offers a split of supply chain profits to the retailer and the retailer decides whether to accept this offer or not,. Despite this similarity, comparison of regrets associated with offering too little and too much is different. The main reason behind this is that the responder in our case, the retailer, doesn't make a simple accept or reject decision. Instead, she decides on the order quantity, which affects the eventual allocation of the profits. And the retailers most of the time choose to accept the contract offers. As a matter of fact, the highest number of rejections per retailer is 22 out of 40 decisions under the RS contract, and this pair is removed from the analysis as an outlier. In other words, the chances of a contract offer getting rejected is low. Additionally, the retailer may accept the contract offer yet decide to protest it by placing a smaller order quantity. However as a result of our experimental design, the minimum order quantity the retailer can order is 51. Thus, if the retailer accepts the contract, even with the lowest order quantity, the manufacturer will earn a considerable amount of profit. Therefore the regret associated with offering too little is not as severe as in an ultimatum game and the dominant regret here is the one associated with offering too much.

Here again we assume all subjects have regret aversion, but they differ in their regret tendency. Hence, we expect the manufacturers with high regret tendency to offer lower predicted profit and profit share to the retailer.

***Hypothesis 2c**: Manufacturers with high regret tendency will have higher newsvendor predicted profit and profit shares.*

To test this hypothesis, we use the regret scale proposed by Schwartz et al. (2002) provided below, consisting of one negative and four positive questions.

Table 67: Schwartz et al.'s regret scale

Below is a list of statements related with your general feelings about your decision making. Please read each statement carefully and select the option which describes how you feel about your decisions best.
1. Once I make a decision, I don't look back. 2. Whenever I make a choice, I'm curious about what would have happened if I had chosen differently. 3. Whenever I make a choice, I try to get information about how the other alternatives turned out. 4. If I make a choice and it turns out well, I still feel like something of a failure if I find out that another choice would have turned out better. 5. When I think about how I'm doing in life, I often assess opportunities I have passed up.

The score from each question is evaluated and averaged to find the overall score. The literature doesn't offer specific threshold values for comparison, however a higher score implies higher tendency to regret after a decision. The evaluation of each question is given in Table 68.

Table 68: Evaluation of Schwarz et al.'s regret scale

Question 1 (negative question)	**Questions 2-5 (positive questions)**
Completely disagree: 7	Completely disagree: 1
Moderately disagree: 6	Moderately disagree: 2
Somewhat disagree: 5	Somewhat disagree: 3
Not sure: 4	Not sure: 4
Somewhat agree: 3	Somewhat agree: 5
Moderately agree: 2	Moderately agree: 6
Completely agree: 1	Completely agree: 7

We compare the subjects regret tendency to the average score of the whole subject pool, which is found to be 4.49.

7.5.1 Newsvendor Experiment Results

Here we compare average order quantity and the average distance of order quantity decisions from the demand mean in terms of the newsvendors' regret tendency. Results presented in Table 69 show that under LPM, orders of newsvendors with high regret tendency are closer to the mean. *Hence Hypothesis 2a* is partially supported.

Table 69: Regret-tendency comparison of the newsvendor experiment

	Regret Tendecy	# Obs.		Avg. Order	Distance from Mean
HPM	**Regret <4.49**	27	**Mean**	100.48	15.89
			Median	99.75	13.95
			Std. Dev.	10.84	6.58
	Regret ≥4.49	30	**Mean**	101.60	16.02
			Median	100.43	15.96
			Std. Dev.	9.45	6.58
	P-values			0.30	0.44
LPM	**Regret <4.49**	67	**Mean**	94.56	18.06
			Median	95.35	16.98
			Std. Dev.	10.08	6.14
	Regret ≥4.49	81	**Mean**	95.89	14.43
			Median	95.35	13.43
			Std. Dev.	7.10	5.34
	P-values			0.35	**0.00**

7.5.2 Contracting Experiment Results

Here we first compare retailers in terms of the pull-to-center effect. Because the contract parameters and hence the optimal order changes from period to period, we compare the retailers in terms of percentage of orders that fall onto the pull-to-center region.

Table 70 shows the comparison results. Under WSP, BB, SP and SPQ treatments, retailers with high regret tendency have more orders in the PTC region; the difference is significant under only WSP. Under RS the difference is reverse and significant, which might be a result of contract design. Hence *Hypothesis 2b* is partially supported.

Table 70: Regret tendency comparison of the pull-to-center effect for contracting experiment retailers

Regret Tendency		WSP #	WSP % orders in PTC region	BB #	BB % orders in PTC region
Regret <4.49	**Mean**	11	15%	7	27%
	Median		13%		22%
	Std. Dev.		14%		19%
Regret ≥4.49	**Mean**	11	31%	15	31%
	Median		26%		24%
	Std. Dev.		21%		23%
	P-value		**0.04**		0.43

Regret Tendency		RS #	RS % orders in PTC region	FP #	FP % orders in PTC region
Regret <4.49	**Mean**	12	29%	10	39%
	Median		24%		36%
	Std. Dev.		17%		23%
Regret ≥4.49	**Mean**	9	16%	12	37%
	Median		13%		34%
	Std. Dev.		13%		15%
	P-value		**0.02**		0.49

Regret Tendency		SP #	SP % orders in PTC region	SPQ #	SPQ % orders in PTC region
Regret <4.49	**Mean**	11	28%	8	38%
	Median		28%		36%
	Std. Dev.		17%		20%
Regret ≥4.49	**Mean**	11	36%	14	45%
	Median		31%		40%
	Std. Dev.		20%		25%
	P-value		0.20		0.29

In addition to the pull-to-center effect, here we compare the retailers in terms of order quantity, expected and realized profit and realized profit share in Table 71. Under WSP, BB and RS retailers with high regret tendency place smaller orders, yet under FP, SP and SPQ the difference is reversed. Under WSP and SPQ, retailers with high regret tendency have lower profit and profit shares, under BB and RS they have higher profit and profit shares. These comparisons may be due to other factors that are unaccounted for here, such as the behavior of the manufacturer these retailers are matched with.

Table 71: Regret-tendency comparison of the contracting experiment retailers

Regret Tendency		WSP					BB				
		#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share	#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share
Regret <4.49	**Mean**	11	1.08	295.20	318.31	0.39	7	1.02	226.41	237.13	0.30
	Median		1.10	295.78	316.24	0.41		1.05	209.07	222.30	0.27
	Std. Dev.		0.16	50.86	50.34	0.07		0.15	77.44	79.69	0.12
Regret ≥4.49	**Mean**	11	1.09	268.44	291.32	0.36	15	1.01	239.05	253.51	0.31
	Median		1.09	260.59	282.55	0.34		0.97	222.19	229.86	0.30
	Std. Dev.		0.14	57.22	59.98	0.10		0.15	83.50	85.08	0.13
	P-value		0.49	**0.08**	**0.09**	0.21		0.35	0.46	0.38	0.38

Regret Tendency		RS					FP				
		#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share	#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share
Regret <4.49	**Mean**	12	1.05	218.80	228.59	0.29	10	1.09	242.75	258.76	0.32
	Median		0.99	218.04	220.20	0.30		1.05	236.95	266.42	0.34
	Std. Dev.		0.15	102.87	105.27	0.14		0.09	67.29	70.67	0.10
Regret ≥4.49	**Mean**	9	0.93	294.81	309.75	0.41	12	1.16	236.44	261.05	0.30
	Median		0.92	337.61	351.51	0.45		1.17	244.78	267.34	0.34
	Std. Dev.		0.05	92.28	97.18	0.11		0.14	102.19	107.99	0.15
	P-value		**0.01**	**0.04**	**0.06**	**0.03**		**0.06**	0.47	0.40	0.42

Regret Tendency		SP					SPQ				
		#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share	#	Q/Q*	Exp. Profit	Real. Profit	Real. Profit Share
Regret <4.49	**Mean**	11	1.07	248.16	262.45	0.33	8	1.09	253.32	273.34	0.34
	Median		1.05	249.04	281.75	0.36		1.04	279.13	303.35	0.41
	Std. Dev.		0.10	89.36	102.34	0.13		0.20	122.45	124.21	0.18
Regret ≥4.49	**Mean**	11	1.16	231.42	250.09	0.29	14	1.09	255.22	272.33	0.34
	Median		1.20	263.64	272.49	0.32		1.07	254.61	263.41	0.33
	Std. Dev.		0.16	63.11	48.77	0.10		0.16	91.08	92.82	0.14
	P-value		0.14	0.41	0.38	0.25		0.45	0.50	0.45	0.37

The comparison results for the manufacturers' predicted profit and profit share are shown in Table 72. Manufacturers with high regret tendency are observed to obtain higher predicted profit and profit shares under BB, RS, FP, SP and SPQ treatments. The differences are significant for RS and SP. Only in the WSP treatment, manufacturers with high regret tendency have lower predicted profit and profit shares. These subjects may be more concerned about the regret associated with contract rejections. Additionally, once again we observe the impact of experimental design on the impact of personality traits. All in all we conclude that *Hypothesis 2c* is partially supported.

Table 72: Regret-tendency comparison of the contracting experiment manufacturers

Regret Tendency		WSP			BB			RS		
		#	**Pred. Profit**	**Pred. Profit Share**	**#**	**Pred. Profit**	**Pred. Profit Share**	**#**	**Pred. Profit**	**Pred. Profit Share**
Regret <4.49	**Mean**	10	402.23	0.59	7	474.85	0.67	12	439.18	0.62
	Median		403.14	0.59		485.61	0.69		430.30	0.61
	Std. Dev.		20.84	0.05		58.35	0.09		53.57	0.11
Regret ≥4.49	**Mean**	12	379.26	0.55	15	493.27	0.68	9	476.72	0.66
	Median		389.83	0.57		495.92	0.68		499.03	0.74
	Std. Dev.		30.33	0.06		51.10	0.10		71.79	0.12
	P-value		**0.05**	**0.07**		0.30	0.42		**0.09**	**0.10**

Regret Tendency		FP			SP			SPQ		
		#	**Pred. Profit**	**Pred. Profit Share**	**#**	**Pred. Profit**	**Pred. Profit Share**	**#**	**Pred. Profit**	**Pred. Profit Share**
Regret <4.49	**Mean**	13	404.37	0.63	12	397.90	0.59	11	405.44	0.61
	Median		412.33	0.62		401.75	0.60		408.06	0.60
	Std. Dev.		45.41	0.12		37.15	0.08		42.17	0.11
Regret ≥4.49	**Mean**	9	423.75	0.65	10	427.88	0.66	11	403.55	0.61
	Median		433.17	0.67		436.31	0.67		416.38	0.63
	Std. Dev.		25.22	0.07		22.21	0.06		47.41	0.11
	P-value		0.22	0.22		**0.02**	**0.04**		0.49	0.49

7.6 Risk and Loss Aversion Analysis

Risk aversion is defined as the human behavior of avoiding uncertainty when faced with risky choices. In economic theory, risk-aversion is modelled using expected utility functions. One of the definitions of risk-averse preferences is that the utility of an outcome reduces with the variability of the outcome. A common way to characterize risk attitude of an individual is to characterize the behavior when offered a risky lottery. For a risk-neutral individual the risky lottery has equal value to its certainty equivalent (expected value). A risk-averse individual values the risky lottery less than the certainty equivalent.

When the lottery includes possibility of losses, loss-aversion is also associated with the decision. Loss-aversion is the preference of avoiding losses over obtaining of the same size gains (Kahneman and Tversky 1984). For a loss-averse individual, a certain amount of money lost has higher value than that amount of money gained.

For a newsvendor, the riskiness of an ordering decision increases with the order quantity. With the lowest possible order the risk will be minimized, and with a high order quantity, the chances of a loss is higher. Hence we expect subjects with higher risk and loss aversion to act more cautiously and avoid inventory risk and losses. Thus, newsvendors with higher risk and loss aversion shall place smaller orders. This translates into lower profits under HPM and higher profits under LPM.

Hypothesis 3a*: Newsvendors with high risk and loss aversion will place smaller orders.*

Similarly, for the contracting experiment, we expect the retailers to place smaller orders and have smaller *Q/Q** ratios.

Hypothesis 3b*: Retailers with high risk and loss aversion will have smaller Q/Q* ratio.*

We expect risk-averse manufacturers to offer lower prices in order to avoid rejection or the retailer protesting with a small order quantity. This will result in the profit value and profit shares of these manufacturers being lower.

Hypothesis 3c*: Manufacturers with high risk and loss aversion will offer lower prices and have lower newsvendor predicted profit and profit shares.*

To test these hypotheses, we use two different sets of questions to measure risk and loss appetite of the subjects. The first method is a survey consisting of four questions. The questions are based on the studies by Hartog et al. (2000) and Gachter et al. (2010). Here, the risk and loss aversion of the subject is measured by comparing their answer to the certainty equivalent of the risky gambles presented in the questions.

Table 73: Risk and loss aversion survey questions

Question 1: There is a lottery draw in which only 10 people will participate and the prize is 1000TL. How much would you be willing to pay for a ticket to the lottery?
Question 2: You are going to be given as a gift either 100 TL or a ticket to the above lottery. Which one would you choose?
Question 3: You are given as a gift a ticket to this lottery, i.e., you didn't pay anything for the ticket. There is someone who is interested in buying the ticket from you, and this person is very talented at negotiation. What is the minimum price you would be willing to sell the ticket?
Question 4: In each of the below situations, there is a tossing the coin game, in which if the coin turns up HEADS, you will win 6 TL. How much you will lose if the coin turns up TAILS is indicated in each situation. If you don't accept the game, nothing will happen. Please indicate if you would accept the game or not. Game 1: TAILS: you will lose 2 TL, HEADS: you will win 6 TL Game 2: TAILS: you will lose 3 TL, HEADS: you will win 6 TL Game 3: TAILS: you will lose 4 TL, HEADS: you will win 6 TL Game 4: TAILS: you will lose 5 TL, HEADS: you will win 6 TL Game 5: TAILS: you will lose 6 TL, HEADS: you will win 6 TL Game 6: TAILS: you will lose 7 TL, HEADS: you will win 6 TL Game 7: TAILS: you will lose 8 TL, HEADS: you will win 6 TL

The certainty equivalent of the risky lottery presented in the first three questions is 100 TL. Hence for the first question, an answer less than 100 TL signals risk-aversion and greater than 100 TL signals risk-seeking behavior. For the second question, selection of the cash implies risk aversion while selection of the lottery ticket implies risk-seeking. The third question is related to both risk and loss aversion behavior. An answer greater than 100 TL indicates risk and loss aversion. For question 4, until game 5, the games have positive certainty equivalents, game 5 has 0 certainty equivalent, and games 6 and 7 have negative certainty equivalent. Hence a risk and loss averse individual is expected to avoid accepting games 5, 6, and 7. The degree of risk and loss aversion can be measured by the number of games accepted by the individual. The lower the number of games accepted, the higher risk and loss aversion.

The second method we use to evaluate the risk attitude of the subjects is the domain-specific-risk-taking (DOSPERT) scale which consists of 30 questions from 5 different domains. Questions from financial domain are highlighted in Table 74.

Table 74: DOSPERT scale questions

For each of the following statements, please indicate the likelihood that you would engage in the described activity or behavior if you were to find yourself in that situation. Provide a rating from Extremely Unlikely to Extremely Likely, using the following scale.	Domain
1. Admitting that your tastes are different from those of a friend.	Social
2. Going camping in the wilderness.	Recreational
3. Betting a day’s income at the horse races.	Financial
4. Investing 10% of your annual income in a moderate growth mutual fund.	Financial
5. Drinking heavily at a social function.	Health/Safety
6. Taking some questionable deductions on your income tax return.	Ethical
7. Disagreeing with an authority figure on a major issue.	Social
8. Betting a day’s income at a high-stake poker game.	Financial
9. Having an affair with a married man/woman.	Ethical
10. Passing off somebody else’s work as your own.	Ethical
11. Going down a ski run that is beyond your ability.	Recreational
12. Investing 5% of your annual income in a very speculative stock.	Financial
13. Going whitewater rafting at high water in the spring.	Recreational
14. Betting a day’s income on the outcome of a sporting event	Financial
15. Engaging in unprotected sex.	Health/Safety
16. Revealing a friend’s secret to someone else.	Ethical
17. Driving a car without wearing a seat belt.	Health/Safety
18. Investing 10% of your annual income in a new business venture.	Financial
19. Taking a skydiving class.	Recreational
20. Riding a motorcycle without a helmet.	Health/Safety
21. Choosing a career that you truly enjoy over a more secure one.	Social
22. Speaking your mind about an unpopular issue in a meeting at work.	Social
23. Sunbathing without sunscreen.	Health/Safety
24. Bungee jumping off a tall bridge.	Recreational
25. Piloting a small plane.	Recreational
26. Walking home alone at night in an unsafe area of town.	Health/Safety
27. Moving to a city far away from your extended family.	Social
28. Starting a new career in your mid-thirties.	Social
29. Leaving your young children alone at home while running an errand.	Ethical
30. Not returning a wallet you found that contains $200.	Ethical

7.6.1 Newsvendor Experiment Results

Comparisons presented in Table 75 yield conflicting results for each of the four questions. Under HPM, risk averse newsvendors place larger orders according to questions 1, 3 and 4. Same applies for under LPM questions 3 and 4. These results can be interpreted in three ways, 1) Risk and loss aversion may not be very effective in determination of the order decisions or 2) These survey questions may not be effective in measuring the risk and loss aversion of the decision makers or 3) Number of observations is not enough.

Table 75: Risk and loss aversion comparison of the newsvendor experiment

			Question 1		Question 2		Question 3		Question 4	
			#	**Avg.**	**#**	**Avg.**	**#**	**Avg.**	**#**	**Avg.**
HPM	**Risk-Averse**	**Mean**	36	102.7	29	100.0	14	100.8	32	103.1
		Median		100.8		99.50		102.6		101.8
		Std.		9.99		11.08		8.17		10.06
	Risk Seeking	**Mean**	2	99.61	28	102.1	28	100.3	2	94.38
		Median		99.61		100.1		99.88		94.38
		Std.		0.19		8.95		9.62		5.83
	P-values			0.41		0.29		**0.00**		0.35
LPM	**Risk-Averse**	**Mean**	74	95.62	65	95.37	27	97.19	65	95.02
		Median		95.14		95.35		98.10		95.00
		Std.		7.57		7.38		6.64		9.15
	Risk Seeking	**Mean**	11	98.05	60	96.37	65	95.93	7	92.97
		Median		96.58		95.94		94.95		91.83
		Std.		8.70		9.65		8.13		4.34
	P-values			0.18		0.40		0.20		**0.00**

DOSPERT scale comparisons are presented in Table 76. Under HPM, risk averse newsvendors make smaller order quantity decisions, under LPM risk averse newsvendors make higher order decisions.

All in all, our analyses results are inconclusive about *Hypothesis 3a.*

Table 76 DOSPERT scale risk aversion comparison of newsvendor experiment

			All Questions		Financial Questions	
			# Obs.	**Avg. Q**	**# Obs.**	**Avg. Q**
HPM	**Risk-Averse**	**Mean**	33	99.50	31	101.16
		Median		99.48		99.50
		Std. Dev.		8.98		9.52
	Risk Seeking	**Mean**	24	103.23	26	100.96
		Median		101.74		100.19
		Std. Dev.		11.21		10.85
	P-values			**0.08**		0.42
LPM	**Risk-Averse**	**Mean**	62	95.59	61	95.31
		Median		96.50		95.60
		Std. Dev.		9.69		10.61

Risk Seeking	Mean Median Std. Dev.	83	94.92 93.63 7.80	84	95.13 94.98 6.91
P-values			0.12		0.30

7.6.2 Contracting Experiment Results

Comparison results for the retailers are shown in Table 77 and Table 78. Retailers who choose 100 TL over the ticket, ie., retailers with risk averse preferences seem to place higher order quantity decisions, for WSP, BB, FP, SP and SPQ, however none of the differences is significant. Under RS risk averse retailers place smaller order decisions and the difference is significant. As for question 4, risk averse retailers place smaller orders under WSP, FP and SP. However under BB, RS and SPQ the difference is reversed and significant under RS and SPQ.

These results combines, our analyses are inconclusive about *Hypothesis 3b.*

Table 77: Question 2 Risk aversion comparison of the contracting experiment retailers

		WSP		BB		RS	
		# Obs.	**Q/Q***	**# Obs.**	**Q/Q***	**# Obs.**	**Q/Q***
100 TL	**Mean** **Median** **Std. Dev.**	5	1.04 1.10 0.20	10	1.04 1.03 0.13	5	0.92 0.92 0.04
Ticket	**Mean** **Median** **Std. Dev.**	17	1.10 1.09 0.13	12	0.99 0.98 0.16	16	1.01 0.98 0.14
P-values			0.37		0.20		**0.05**
		FP		**SP**		**SPQ**	
		# Obs.	**Q/Q***	**# Obs.**	**Q/Q***	**# Obs.**	**Q/Q***
100 TL	**Mean** **Median** **Std. Dev.**	6	1.13 1.15 0.12	12	1.14 1.09 0.14	6	1.13 1.07 0.24
Ticket	**Mean** **Median** **Std. Dev.**	16	1.13 1.12 0.13	10	1.09 1.08 0.15	16	1.07 1.04 0.14
P-values			0.41		0.30		0.33

Table 78: Question 4 Risk and loss aversion comparison of the contracting experiment retailers

		WSP		BB		RS	
		# Obs.	Q/Q*	# Obs.	Q/Q*	# Obs.	Q/Q*
Accept 3 or less games	**Mean**	11	1.11	10	1.09	10	1.01
	Median		1.09		1.09		0.97
	Std. Dev.		0.12		0.11		0.18
Accept 4 or more games	**Mean**	11	1.07	12	0.95	11	0.98
	Median		1.10		0.96		0.96
	Std. Dev.		0.18		0.15		0.07
P-values			0.38		**0.01**		0.31
		FP		**SP**		**SPQ**	
		# Obs.	**Q/Q***	**# Obs.**	**Q/Q***	**# Obs.**	**Q/Q***
Accept 3 or less games	**Mean**	11	1.11	9	1.12	10	1.14
	Median		1.12		1.07		1.12
	Std. Dev.		0.12		0.16		0.19
Accept 4 or more games	**Mean**	11	1.15	13	1.11	12	1.05
	Median		1.16		1.09		1.02
	Std. Dev.		0.13		0.14		0.14
P-values			0.23		0.33		**0.06**

Comparison results of the manufacturers are presented in Table 79 and Table 80. According to question 2 answers, manufacturers who select 100 TL over the lotter ticket have higher predicted profit and profit share for all treatments except the SP treatment, which contradicts our hypothesis.

Table 79: Risk aversion comparison of the contracting experiment manufacturers – Question 2

		WSP			BB		
		# Obs.	Pred. Profit	Pred. Profit Share	# Obs.	Pred. Profit	Pred. Profit Share
100 TL	**Mean**	12	397.01	0.59	10	488.88	0.71
	Median		397.95	0.58		490.76	0.74
	Std. Dev.		25.13	0.05		60.97	0.10
Ticket	**Mean**	10	380.93	0.55	12	480.04	0.64
	Median		386.30	0.55		479.73	0.66
	Std. Dev.		30.82	0.06		50.57	0.06
P-values			0.11	**0.06**		0.30	**0.03**
		RS			**FP**		
		# Obs.	**Pred. Profit**	**Pred. Profit Share**	**# Obs.**	**Pred. Profit**	**Pred. Profit Share**

100 TL	**Mean**	10	468.74	0.68	9	414.77	0.65
	Median		465.41	0.70		413.29	0.61
	Std. Dev.		74.89	0.14		43.13	0.12
Ticket	**Mean**	11	440.98	0.61	13	412.09	0.63
	Median		429.14	0.61		419.26	0.64
	Std. Dev.		48.25	0.09		35.98	0.09
P-values			0.16	**0.07**		0.36	0.33
		SP			**SPQ**		
		# Obs.	**Pred. Profit**	**Pred. Profit Share**	**# Obs.**	**Pred. Profit**	**Pred. Profit Share**
100 TL	**Mean**	12	405.10	0.60	11	403.00	0.61
	Median		410.69	0.61		412.00	0.61
	Std. Dev.		40.25	0.09		38.71	0.10
Ticket	**Mean**	10	419.23	0.64	11	405.99	0.62
	Median		420.06	0.65		395.63	0.58
	Std. Dev.		24.96	0.06		50.24	0.12
P-values			0.22	0.14		0.42	0.42

According to question 4 comparison of the manufacturers, presented in Table 80, risk averse manufacturers have smaller predicted profit and profit share under WSP, RS, SP and SPQ treatments, the difference is significant under SPQ.

Hence we conclude that *Hypothesis 3c* is partially supported.

Table 80: Risk and loss aversion comparison of the contracting experiment manufacturers – Question 4

		WSP			**BB**		
		# Obs.	**Pred.**	**Pred.**	**# Obs.**	**Pred.**	**Pred.**
Accept 3 or less games	**Mean**	12	387.34	0.56	11	488.38	0.67
	Median		391.96	0.57		501.37	0.69
	Std. Dev.		32.67	0.07		48.74	0.07
Accept 4 or more games	**Mean**	10	392.53	0.57	11	479.73	0.67
	Median		397.95	0.58		469.19	0.68
	Std. Dev.		23.60	0.04		61.52	0.11
P-values			0.32	0.21		0.36	0.49
		RS			**FP**		
		# Obs.	**Pred.**	**Pred.**	**# Obs.**	**Pred.**	**Pred.**
Accept 3 or less games	**Mean**	12	449.37	0.64	10	410.35	0.64
	Median		441.96	0.64		420.06	0.65
	Std. Dev.		65.41	0.13		43.02	0.10
Accept 4 or more games	**Mean**	9	460.64	0.64	12	415.54	0.64
	Median		447.82	0.69		416.27	0.62
	Std. Dev.		61.32	0.10		35.23	0.10

P-values			0.5	0.36		0.47	0.47
		SP			**SPQ**		
		# Obs.	**Pred.**	**Pred.**	**# Obs.**	**Pred.**	**Pred.**
Accept 3 or less games	**Mean**	9	407.54	0.61	11	384.25	0.58
	Median		423.75	0.63		388.63	0.56
	Std. Dev.		47.46	0.10		49.15	0.12
Accept 4 or more games	**Mean**	13	414.28	0.63	11	424.74	0.65
	Median		412.00	0.62		423.75	0.63
	Std. Dev.		22.84	0.06		26.65	0.08
P-values			0.47	0.42		**0.02**	**0.04**

7.7 Inequity Aversion Analysis

Inequity aversion is the human behavior of preferring fairness and avoiding inequitable allocations. Many researchers such as Fehr and Schmidt (1999) have shown that individual decision makers have strong aversion towards inequity. These authors illustrate that inequity aversion is not symmetrical. That is, decision maker's response to inequity when their payoff is higher or lower than their opponent's is different. We adopt this asymmetrical inequity aversion notion. We refer to the case where the decision maker receives higher payoff than his opponent as "advantageous inequity" and the case where he receives a lower payoff as "disadvantageous inequity".

In our experiments the manufacturer has the first-mover advantage and often receives a higher share of the total supply chain profit, leaving the retailers at a disadvantage. Thus we expect the retailers' decisions to be affected by disadvantageous inequity aversion. Retailers with high aversion towards disadvantageous inequity shall act more assertive and ask for higher profit shares.

Hypothesis 4a: Retailers with high disadvantageous inequity aversion will reject more contracts and have higher profit shares.

As the manufacturers have the profit advantage in our setting, we expect them to be affected by the advantageous inequity aversion and offer more favorable contracts to the retailer to achieve a more equitable distribution of profits.

Hypothesis 4b: Manufacturers with high advantageous inequity aversion will offer lower prices and will have lower profit shares.

To measure inequity aversion we ask the subjects three questions that are presented in Table 81. These questions are based on the ultimatum and dictator games developed by Forsythe et al. (1994).

Table 81: Inequity aversion survey questions

Question 1: You are given 100 TL, to distribute it between yourself and a complete stranger. The money is not given as a compensation for some work, it is just given. You are to decide who gets how much, and the other person cannot reject your distribution. How much of the 100 TL would you take for yourself?
Question 2: Again you are given 100 TL to distribute between yourself and this complete stranger. But now the stranger can reject your distribution and in that case you both will receive 0 TL. Now, how much of the 100 TL will you take for yourself?
Question 3: This time, the 100 TL is given to the other guy to distribute between the two of you. Below are possible splitting scenarios. You can reject the distribution and in that case you both will receive 0TL. Please indicate if you would accept or reject each scenario. 1. You: 10 TL -- Other guy: 90 TL 2. You: 20 TL -- Other guy: 80 TL 3. You: 30 TL -- Other guy: 70 TL 4. You: 40 TL -- Other guy: 60 TL 5. You: 50 TL -- Other guy: 50 TL

The first two questions are related to advantageous inequity, for which an answer different from 50 TL results in an inequity. Subjects who take more than 50 TL to themselves have lower aversion towards advantageous inequity. A smaller number of accepted offers in the third question indicate higher disadvantageous inequity aversion.

7.7.1 Experiment Results

Comparison results of the retailers are presented in Table 82 and Table 83. Here we conduct two comparisons differentiating the sensitivity levels of the retailers towards inequity. Table 82 compares highly sensitive retailers with the remaining ones. We see that under all contracts highly sensitive retailers have higher contract rejections and the difference is significant under SP treatment. . In addition, highly sensitive retailers have higher expected profit under BB, RS, FP, SP and SPQ treatments and the difference is significant under SP.

Table 82: Inequity aversion comparison of the contracting experiment retailers – high sensitivity

		WSP			BB		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept only 50-50	**Mean**	7	3.71	247.19	5	5.80	256.16
	Median		3.00	241.33		7.00	230.76
	Std. Dev.		4.19	43.23		4.60	89.49
Other	**Mean**	15	3.27	297.98	17	4.12	228.81
	Median		2.00	295.78		4.00	222.19
	Std. Dev.		3.10	52.79		2.67	78.83
P-values			0.21	0.14		0.25	0.13
		RS			**FP**		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept only 50-50	**Mean**	5	5.40	300.51	7	3.29	231.94
	Median		2.00	333.22		3.00	255.71
	Std. Dev.		6.11	102.27		2.56	81.50
Other	**Mean**	16	2.75	240.71	15	3.07	242.75
	Median		2.00	230.08		1.00	236.95
	Std. Dev.		2.70	101.36		3.83	90.80
P-values			0.31	0.37		0.27	0.14
		SP			**SPQ**		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept only 50-50	**Mean**	8	6.13	251.52	8	3.13	293.29
	Median		3.00	272.67		5.00	289.68
	Std. Dev.		6.56	48.30		2.59	79.57
Other	**Mean**	14	2.71	233.09	14	2.14	232.38
	Median		1.50	232.65		0.00	241.31
	Std. Dev.		3.00	86.33		2.57	107.27
P-values			**0.07**	**0.09**		0.29	0.27

The comparison of moderately sensitive retailers with the remaining ones is presented in Table 83. We see that moderately inequity averse retailers have higher contract rejections under BB, FP, SP and SPQ treatments and higher expected profit under RS and SPQ treatment.

Hence we conclude that *Hypothesis 4a* is partially supported.

Table 83: Inequity aversion comparison of the contracting experiment retailers – moderate sensitivity

		WSP			BB		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept 50-50 and 40-60	**Mean**	14	3.29	260.60	15	4.80	218.06
	Median		2.00	266.44		4.00	212.84
	Std. Dev.		3.24	45.70		3.57	83.89
Other	**Mean**	8	3.63	318.96	7	3.86	271.39
	Median		2.50	319.91		4.00	300.36
	Std. Dev.		3.85	50.79		2.12	60.91
P-values			0.34	**0.09**		0.11	0.49
		RS			**FP**		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept 50-50 and 40-60	**Mean**	11	3.27	268.11	16	3.44	231.59
	Median		2.00	293.89		2.50	235.34
	Std. Dev.		4.47	101.40		3.33	89.54
Other	**Mean**	10	3.50	240.47	6	2.33	259.90
	Median		2.50	230.08		0.00	261.86
	Std. Dev.		3.06	106.84		3.83	80.11
P-values			0.20	0.42		**0.08**	0.25
		SP			**SPQ**		
		# Obs.	**# Rejections**	**Expected Profit**	**# Obs.**	**# Rejections**	**Expected Profit**
Accept 50-50 and 40-60	**Mean**	15	4.13	228.51	15	2.67	263.32
	Median		1.00	216.27		5.00	260.14
	Std. Dev.		5.37	63.37		2.58	82.42
Other	**Mean**	7	3.57	263.97	7	2.14	235.68
	Median		5.00	258.94		0.00	219.85
	Std. Dev.		3.46	93.74		2.67	137.89
P-values			0.25	**0.01**		0.27	0.38

Inequity aversion comparison results of the manufacturers are presented in Table 84. We see that inequity averse manufacturers have smaller predicted profit and profit share under WSP, RS, FP and SP treatments. The difference is only significant for the FP treatment. We conclude that *Hypothesis 4b* is partially supported.

Table 84: Inequity aversion of contracting experiment manufacturers

		WSP			BB		
		# Obs.	**Pred. Profit**	**Pred. Profit Share**	**# Obs.**	**Pred. Profit**	**Pred. Profit Share**
Offer 50-50	**Mean**	6	390.94	0.57	11	484.91	0.69
	Median		389.83	0.57		495.92	0.71
	Std. Dev.		12.07	0.03		58.72	0.10
Take more than	**Mean**	16	389.23	0.57	11	483.20	0.65
	Median		397.95	0.57		463.34	0.66

50	Std. Dev.		32.83	0.07		52.45	0.07
P-values			0.30	0.41		0.46	0.13
		RS			FP		
		# Obs.	Pred. Profit	Pred. Profit Share	# Obs.	Pred. Profit	Pred. Profit Share
Offer 50-50	Mean		440.40	0.62		395.81	0.60
	Median	12	429.72	0.58	10	403.14	0.59
	Std. Dev.		52.87	0.12		44.77	0.11
Take more than 50	Mean		472.59	0.67		427.66	0.67
	Median	9	453.62	0.69	12	434.09	0.67
	Std. Dev.		72.22	0.11		25.01	0.08
P-values			0.14	0.14		**0.04**	**0.04**
		SP			SPQ		
		# Obs.	Pred. Profit	Pred. Profit Share	# Obs.	Pred. Profit	Pred. Profit Share
Offer 50-50	Mean		401.13	0.61		405.14	0.62
	Median	9	407.63	0.62	14	419.31	0.63
	Std. Dev.		45.51	0.10		51.99	0.13
Take more than 50	Mean		418.72	0.63		403.36	0.59
	Median	13	423.75	0.63	8	393.56	0.57
	Std. Dev.		22.93	0.06		27.07	0.06
P-values			0.19	0.32		0.25	0.21

7.8 Conclusion

In this chapter we investigate the effect of various personality traits on the newsvendor', the retailer's and the manufacturer's decisions. Despite the small sample sizes, we find results supporting our hypotheses, albeit weakly and partially, indicating a relation between personality traits and experiment performance. We used only one data point for each subject as we take averages of the performance measures over all accepted contracts of the subject. Representing each subject with just one average data point substantially truncates the effect of personality traits. As a further study, the analyses can be conducted on pooled data of all subjects in each treatment. Additionally, one can develop advanced models to test the joint effect of these personality traits.

7.9 References

Akbay, Ü., Kaya, M. (2016). Experiments on Supply Chain Contracting: Effects of contract type. Working paper.

Akbay, Ü., Kaya, M. (2016). Experiments on the Power of Commitment and Fairness Priming in Supply Chain Decisions. Working paper.

Akbay, Ü., Kaya, M. (2016). Gender Differences on Newsvendor Behavior. Working paper.

Akpınar, S., Bay, M., Şahan, S. B. (2014). "The examination of students' self-esteem and assertiveness levels who study at physical education and sport academy", Prime journal of Social Sciences, 3(3), 648-652.

Bell, D. E. (1982). Regret in decision making under uncertainty. Operations Research, 30, 961–981.

Bolton, G. E., Katok E. (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback, Manufacturing and Service Operations Management, 10, 519-538.

Gachter, S., Johnson, E. J., Herrmann, A. (2010). "Individual-level loss aversion in riskless and risky choices", CEDEX Discussion Paper No. 2010-20, University of Nottingham.

Gilovich, T., Medvec, V. H. (1995). The experience of regret: what, when and why. Psychological Review, 102, 379–395.

Hartog, J., Ferrer-I-Corbonell, A., Jonker, N. (2000). On a simple survey measure of individual risk aversion. CESifo Working Paper Series No. 363, CESifo, Munich, Germany.

Josephs, R. A., Larrick, R. P., Steele, C. M., Nisbett, R. E. (1992). Protecting the self from the negative consequences of risky decisions. Journal of Personality and Social Psychology, 62(1), 26-37.

Keser, C., Paleologo, G. A. 2004. Experimental Investigation of Supplier-Retailer Contracts: The Wholesale Price Contract. Working paper, CIRANO Scientific Series

Lau, N., Bearden, J. N., Hasija, S. (2014). Newsvendor Pull-to-Center Reconsidered. Decision Support Systems, 58, 68-73.

Moritz, B. (2010). Cognition and Heterogeneity in Supply Chain Planning: A Study of Inventory Decision Making, (Doctoral Dissertation), University of Minnesota.

Moritz, B., Hill, A. V., Donohue, K 2013. "Cognition and Individual Differences in the Newsvendor Problem: Behavior under Dual Process Theory", Journal of Operations Management, 31, 72-85.

Rosenberg, M. (1965). “Society and the adolescent self-image”, Princeton, NJ: Princeton University Press.

Schwartz, B., Ward, A., Monterosso, J., Lyubomisky, S., White, K., Lehman, D. R. (2002). “Maximizing versus satisficing: Happiness is a matter of choice”, Journal of Personality and Social Psychology, 83-5.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Strohhecker, J., Grössler A. (2013). “Do personal traits influence inventory management performance? The case of intelligence, personality, interest and knowledge”, Int. Journal of Production Economics 142, 37-50.

Zeelenberg, M., Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. Organizational Behavior and Human Decision Processes, 67, 63–78

Zeelenberg, M., Beattie, J., van der Pligt, J., de Vries, N. K. (1996). Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making. Organizational Behavior and Human Decision Processes, 65(2), 148-158.

Chapter 8

8. CONCLUSION

In this dissertation we conduct decision-making experiments with human subjects on simple supply chain settings.

In Chapter 3, we study the effect of contract type on supply chain contracting and compare three widely used supply chain contracts, namely, the wholesale price, buyback and revenue sharing contracts. Our experiments are based on a simple scenario where the manufacturer produces and sells a certain product to the retailer who faces the standard newsvendor problem. Our analyses reveal that in the human-to-human experiment, the total supply chain profits are shared more equitably than the theoretical predictions. We also observe that the wholesale price contract performs as well as the more sophisticated buyback and revenue sharing contracts.

In Chapter 4, using expected utility theory, we aim to incorporate behavioral factors that might be affecting the retailers' and the manufacturers' decision in the experiment presented in Chapter 3. We study risk-aversion, loss-aversion, inventory error aversion, and social preference models.

In Chapter 5, we investigate the effects of power of commitment and fairness priming on supply chain relationships. We use a simple wholesale price contract scenario and forcing the manufacturer's and the retailer's decisions to be effective for five periods we study the effect of power of commitment. Our results show that commitment to decision provides the firms with an advantage over the other partner and this helps them increase their profit values. Additionally we observe that the fairness priming isn't strong enough on the retailers and benefits the manufacturers instead.

In Chapter 6, we study gender differences in a newsvendor experiment. We find that female subjects place smaller orders than male subjects. Also female subjects are more affected by the stochastic nature of the consumer demand and are more prone to demand-chasing behavior.

In Chapter 7, we investigate the relation between experiment performance and various personality traits. We study, self-esteem, regret-aversion, risk and loss aversion, and inequity aversion.

8.1 Comparison with a Randomized System

Here, as an extension study, we compare the performance of human decision makers in our experiments with the performance of a system that makes random decisions.

8.1.1 Newsvendor Experiment

For the randomized system we generate 10000 newsvendors each of whom make 40 ordering decisions by choosing any order quantity between 50 and 150 with equal probability. Then we compute the expected profit values corresponding to these ordering decisions, and find averages for each random individual. Table 85 presents the comparison of the average performance measures for the human and randomized newsvendors. Naturally, the average order quantity for the randomized system is close to the mean of the demand distribution. Randomized ordering decisions lead to lower expected profit than the ordering decisions of the human newsvendors. However, ratio of the expected profit achieved by the randomized newsvendors to the theoretical optimal profit is 90% under HPM and 82% under LPM. So we conclude that although the random ordering decisions achieve quite a high level of efficiency, they cannot outperform a human decision maker.

Table 85: Comparison of the experiment results with randomized system – Newsvendor experiment

		Theory	Female Subjects	Male Subjects	Rnd. Newsvendor
HPM	Order Quant.	111	99.27 (9.57)	103.39 (7.47)	99.87 (4.57)
	Expected Profit	4420	4171.8 (134)	4244.5 (133.1)	3983.21 (68.56)
LPM	Order Quant.	89	94 (9.57)	96.56 (7.43)	99.89 (4.64)
	Expected Profit	2420	2202.9 (126.3)	2234 (146.2)	1988.75 (68.92)

Standard deviations are reported in parentheses.

8.1.2 Contracting Experiments

For the contracting experiment we consider three variations of the randomized system:

1. **Rational manufacturer, random retailer**
   The manufacturer is rational, follows the theory and offers the optimal contract parameters while the retailer makes a random order quantity decision by choosing each order quantity between 51 and 150 with equal probability.
2. **Random manufacturer, rational retailer**
   The retailer is rational, follows the standard theory and places the newsvendor optimal Q^* as the order decision. The manufacturer makes a random contract decision by choosing among feasible contract parameters with equal probability.
3. **Random manufacturer, random retailer**
   Both the manufacturer and the retailer make random decisions.

Table 86 - Table 88 compare the performances of human decision makers in the experiments with the performances of the randomized systems. We make the comparison on the expected level, hence for the experiment data only the accepted contracts are included in the analyses. For the experiment data, each data point is the average over all periods for each retailer/manufacturer, which constitute a sample of 22 for the WSP and BB contracts and 21 for RS contract. For the randomized system, simulation the sample size is 10,000. Each data point is an average of 40 periods.

For the WSP contract, all three randomized systems as well as the human decision makers outperform the theoretical efficiency expectation. System1 and System3 achieve slightly higher efficiency than human decision makers. In System2, when the retailer is rational, i.e., placing $Q^*(w)$ as order decisions, the efficiency of the system reduces to 85%. These two results can be interpreted as retailer's suboptimal, even random ordering decisions under WSP contract increase contract performance. This may be due to higher quantities (ie., orders above the demand mean) being ordered more frequently in these systems. When we compare the allocation of the total supply chain profits among the two firms, we observe that compared to the completely randomized system, the firm who makes rational (optimal) decisions increases their profit share. Compared to the completely random system, human retailers have higher profit share. Additionally, human retailers have higher profit share than the theoretical expectation. Hence we can conclude that the decision makers being human, rational or randomized has a strong effect on the profit allocation.

Table 86: Comparison of the experiment results with randomized system – WSP contract

	Theory	**Experiment**	**System 1: Rat. Mfg. Rnd. Ret.**	**System 2: Rnd. Mfg. Rat. Ret.**	**System 3: Rnd. Mfg. Rnd. Ret.**
w	10	7.5 (0.51)	10	7.51 (0.45)	7.50 (0.46)
Q	67	96.36 (9.85)	100.32 (4.43)	87.43 (3.79)	100.37 (4.52)
Ret.'s Exp. Profit	117.68	281.8 (51.7)	1.97 (19.00)	345.04 (40.36)	252.12 (48.08)
Mfg.'s Exp. Profit	469	418.94 (66.3)	702.23 (30.99)	325.78 (24.82)	452.05 (51.2)
Total Exp. Chain Profit	586.68	700.74 (38.58)	704.2 (15.06)	670.82 (17.88)	704.18 (15.29)
Exp. Contract Efficiency	0.74	0.88 (0.05)	0.89 (0.02)	0.85 (0.02)	0.89 (0.02)
Mfg.'s Exp. Profit Share	0.80	0.60 (0.07)	0.98 (0.03)	0.54 (0.05)	0.63 (0.07)

Standard deviations are reported in parentheses.
Experiment sample size n=22, simulation sample sizes n=10,000.

Table 87: Comparison of the experiment results with randomized system – BB contract

	Theory	**Experiment**	**System 1: Rat. Mfg. Rnd. Ret.**	**System 2: Rnd. Mfg. Rat. Ret.**	**System 3: Rnd. Mfg. Rnd. Ret.**
w	11	8.72 (0.71)	11	7.5 (0.45)	7.48 (0.45)
b	10	5.02 (1.48)	10	3.74 (0.47)	3.74 (0.48)
Q	100	100.02 (9.33)	100.48 (4.70)	104.65 (4.80)	100.61 (4.70)
Ret.'s Exp. Profit	75.5	232.84 (62.63)	67.22 (1.10)	385.4 (42.99)	316.76 (43.98)
Mfg.'s Exp. Profit	677.5	482.11 (75.36)	637.54 (15.42)	321.32 (32.00)	388.22 (46.02)
Total Exp. Chain Profit	753	714.96 (32.52)	704.77 (15.90)	706.72 (17.46)	704.98 (15.67)
Exp. Contract Efficiency	0.95	0.90 (0.04)	0.89 (0.02)	0.89 (0.02)	0.89 (0.02)
Mfg.'s Exp. Profit Share	0.90	0.68 (0.09)	0.90 (0.00)	0.49 (0.05)	0.55 (0.06)

Standard deviations are reported in parentheses.
Experiment sample size n=22; simulation sample sizes n=10,000.

Under the BB contract, none of the systems nor the experiment achieves the theoretical efficiency level of 95%. All three systems achieve an efficiency level of 89%, which is slightly less than the experiment efficiency level. Despite the efficiency levels being close, there is significant difference in the allocation of supply chain profits between the three systems and the experiment performance. Similar to the result under WSP contract, we see that compared to System 3, the firm who makes rational decisions increase its profit share. Human manufacturers have higher profit share than the completely random system manufacturers. This may be due to human manufacturers making contracting decisions by considering both contract parameters, while the randomized manufacturer determines them independently of each other.

Table 88: Comparison of the experiment results with randomized system – RS contract

	Theory	**Experiment**	**System 1: Rat. Mfg. Rnd. Ret.**	**System 2: Rnd. Mfg. Rat. Ret.**	**System 3: Rnd. Mfg. Rnd. Ret.**
w	1	4.33 (1.45)	1	5.96 (0.58)	6.01 (0.57)
r	10	4.1 (1.44)	10	3.03 (0.46)	2.97 (0.46)
Q	100	95.45 (12.17)	100.75 (4.61)	85.69 (5.11)	100.49 (4.58)
Ret.'s Exp. Profit	75.5	252.38 (89.81)	67.18 (1.17)	246.87 (43.44)	153.47 (47.89)
Mfg.'s Exp. Profit	677.5	444.19 (81.13)	638.14 (14.92)	387.84 (30.17)	551.02 (51.22)
Total Exp. Chain Profit	753	696.57 (40.71)	705.31 (15.4)	634.71 (20.06)	704.49 (15.30)
Exp. Contract Efficiency	0.95	0.88 (0.05)	0.89 (0.02)	0.80 (0.03)	0.89 (0.02)
Mfg.'s Exp. Profit Share	0.90	0.64 (0.12)	0.90 (0.00)	0.67 (0.06)	0.77 (0.07)

Standard deviations are reported in parentheses.
Experiment sample size n=21; simulation sample sizes n=10,000.

Under the RS contract, similar to the BB contract none of the randomized systems or the experiment achieves the theoretical efficiency level of 95%. Comparison of the experiment efficiency and the randomized system efficiencies is very similar to the one under WSP contract. System 1 and System 3 achieve slightly higher efficiency than the human decision makers. System 2 with the rational retailer achieves a lower efficiency level. Again the firm who makes rational decisions increases its profit share compared to the completely randomized system. Randomized manufacturers achieve higher profit share than human manufacturers. This may be due to again the two contract parameters being determined together by human manufacturers while they are independently determined in the randomized systems.

Overall, we observe that randomized systems achieve quite high efficiency levels, as high as (and even higher than) the efficiency level achieved with human decision makers. Hence from a centralized supply chain point of view, a randomized decision system may be as good as human decision makers. However the centralized supply chain point of view ignores the allocation of profit between the two firms, which is significantly important for decentralized systems.

From each firm's point of view, there is a significant difference between trying to make a good decision within personal limitations and making random decisions. In the newsvendor experiment, both female and male subjects make better decisions and achieve higher expected profit than the random ordering decisions. In the contracting experiment, human retailers achieve higher profit shares than randomized retailers.

Moreover we observe that the contract design has an impact on the performance of the randomized system. While the wholesale price and revenue sharing contracts affect the random system performance in a similar fashion, buyback contract differs from these two.

These results show that the subjects in our decision making experiments do not make random decisions, but rather try to achieve a good profit or profit share by making "good" decisions within the limitations of the strategic interaction in the experiment and their bounded rationalities and personal traits.

9. REFERENCES

Akpınar, S., Bay, M., Şahan, S. B. (2014). “The examination of students’ self esteem and assertiveness levels who study at physical education and sport academy”, Prime journal of Social Sciences, 3(3), 648-652.

Bargh, J. A., Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social and personality psychology (pp. 253–285). New York: Cambridge University Press

Becker-Peth, M., Katok E., Thonemann U. W. (2013). Designing Buyback Contracts for Irrational but Predictable Newsvendors. Management Science 59(8),1800-1816.

Bell, D. E. (1982). Regret in decision making under uncertainty. Operations Research, 30, 961–981.

Benartzi, S., R. H. Thaler. (1995). Myopic loss aversion and the equity premium puzzle. Quart. J. Econ. 110(1), 73–92.

Bendoly, E., K. Donohue, K. L. Schultz. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of Operations Management, 24, 737-752.

Bendoly, E., Van Wezel, W., Bachrach, D. G. (Eds.). (2015). The handbook of behavioral operations management: Social and psychological dynamics in production and service settings. New York: Oxford University Press.

Benzion, U., Cohen Y., Peled R., Shavit T. (2008). Decision-making and the newsvendor problem: an experimental study, Journal of Operations 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., Ockenfels, A. (2000). ERC: a theory of equity, reciprocity and competition. The American Economic Review, 90(1), 166-193.

Bolton, G. E., Katok E. (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback, Manufacturing and Service Operations Management, 10, 519-538.

Bolton, G. E., Ockenfels A., Thonemann U. W. (2012). Managers and Students as Newsvendors. Management Science, 58(12), 2225-2233.

Bostian, AJ. A, Holt C. A., Jain S., Ramdas, K. (2012). Is Transshipment a Behaviorally-Robust Risk-Pooling Strategy? Working Paper.

Bostian, AJ. A, Holt C. A., Smith A. M. (2008). Newsvendor "Pull-to-Center" Effect: Adaptive Learning in a Laboratory Experiment, Manufacturing and Service Operations Management,

Brown, A. O., Tang C. S. (2006). The Impact of Alternative Performance Measures on Single-Period Inventory Policy, Journal of Industrial and Management Optimization, 2, 297-318.

Byrnes, J.P., Miller, D.C., Schafer, W.D. (1999). Gender differences in risk taking: a meta-analysis. Psychological Bulletin, 125, 367–383.

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.

Caliskan-Demirag, O., Chen Y., Li J. (2010). Channel Coordination Under Fairness Concerns and Nonlinear Demand, European Journal of Operations Research, 207, 1321-1326.

Cervone, D., Peake, P. K. (1986). Anchoring, Efficacy and Action: The Influence of Judgmental Heuristics on Self-Efficacy Judgments and Behavior. Journal of Personality and Social Psychology, 50 (3), 492-501.

Chen, L., Kök A. G., Tong J. (2013). The Effect of Payment Schemes on Inventory Decisions: The Role of Mental Accounting. Management Science, 59(2), 436-451.

Corbett, C. J., Fransoo, J. C. (2007). Enterpreneurs and newsvendors: do small businesses follow the newsvendor logic when making inventory decisions? Working Paper.

Croson, D. C., R. Croson, Y. Ren. (2013). How to Manage an Overconfident Newsvendor. Management Science, 59(11), 2502-2517.

Croson, R., Gneezy, U. (2009). Gender differences in preferences. Journal of Economic Literature, 47, 1–27.

Çuhadaroğlu, F. (1986). "Adölesanlarda Benlik Saygısı",(Residency Thesis), Hacettepe Üniversitesi Tıp Fakültesi Psikiyatri ABD.

Cui, T. H., Malluci P. (2016). Bounded Rationality and Fairness Ideals in Distribution Channels. Working Paper, to appear in Journal of Marketing Research.

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

Cui, Y., Chen L. G., Chen J., Gavirneni S., Wang, Qi. (2013). Chinese Perspectives on Newsvendor Bias: An Exploratory Note. Journal of Operations Management, 31, 93-97.

Davis, A. M. (2015). An Experimental Investigation of Pull Contracts. Production and Operations Management, 24(2), 325-340.

Davis, A. M., Katok, E., Santamar, N. (2014). Push, Pull, or Both? A Behavioral Study of How the Allocation of Inventory Risk Affects Channel Efficiency. Management Science, 60(11), 2666-2683.

De Vericourt, F. D., Bearden, J. N., Filipowicz, A. (2013). Sex, Risk and the Newsvendor. Journal of Operations Management, 31, 86-92.

Dimitriou, S., Robinson S., Kotiadis K. (2009). The Impact of Human Decision Makers' Individualities on the wholesale price contract's efficiency, Winter Simulation Conference, Austin, TX, USA, 2009, 2353-2364.

Erhun, F. (2011). Contract Complexity and Performance, Tutorials in Operations Research, Fehr, E., Schmitdt, K. M. (1999). A Theory of Fairness, Competition and Cooperation. Quarterly Journal of Economics, 114(3).

Fellner, G., M. Sutter. (2009). Causes, consequences and cures of myopic loss aversion: An experimental examination. The Economic Journal. 119(537), 900–916.

Feng, T., Keller L. R., Zheng X. (2011). Decision making in the newsvendor problem: A cross-national laboratory study, Omega, 39, 41-50.

Fisher, M., Raman, A. (1996). Reducing the cost of demand uncertainty through accurate response to early sales. Special Issue on New Directions in Operations Management, 44(1), 87-99.

Gachter, S., Johnson, E. J., Herrmann, A. (2010). "Individual-level loss aversion in riskless and risky choices", CEDEX Discussion Paper No. 2010-20, University of Nottingham.

Gavirneni, S., Isen A. M. (2010). Anatomy of a Newsvendor Decision: Observations from a Verbal Protocol Analysis, Production and Operations Management, 19, 453-462.

Gavirneni, S., Xia Y. (2009). Anchor Selection and Group Dynamics in Newsvendor Decisions-A note, Decision Analysis, 6, 87-97.

Gigerenzer, G., Selten, R. (2002). Bounded Rationality. Cambridge: MIT Press

Gilovich, T., Medvec, V. H. (1995). The experience of regret: what, when and why. Psychological Review, 102, 379–395.

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. Technometrics 11(1), 1-21.

Güth, W., R. Schmittberger, B. Schwarze. (1982). An Experimental Analysis of Ultimatum Bargaining. Journal of Economic Behavior and Organization 3 (4), 367–388.

Hartog, J., Ferrer-I-Corbonell, A., Jonker, N. (2000). On a simple survey measure of individual risk aversion. CESifo Working Paper Series No. 363, CESifo, Munich, Germany.

Haruvy, E., Katok E., Pavlov, V. (2016). Can Structured Bargaining Improve Channel Efficiency? Working Paper.

Ho, T., Zhang, J (2008). Designing Pricing Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter? Management Science, 54(4), 686-700.

Hoskin, R. E. (1983). Opportunity Cost and Behavior, Journal of Accounting and Research, 21, 78-95.

Huber, J., Kirchler, M., Stöckl, T. (2010). The hot hand belief and the gambler's fallacy in investment decisions under risk. Theory and Decision, 68, 445-462.

Hügelschäfer, S., Achtziger, A. (2014) On confident men and rational women: It's all on your mind(set). Journal of Economic Psychology, 41, 31-44.

Hung, K., Ryu S. (2008). Changing risk preferences in supply chain inventory decisions, Production Planning and Control: The Management of Operations, 19, 770-780.

Hyndman, K., Kraiselburd S., Watson N. (2014). Coordination in Games with Strategic Complementarities: An Experiment on Fixed vs. Random Matching. Production and Operations Management, 23, 221–238.

Hyndman, K., Kraiselburd, S., Watson, N. (2013). Aligning Capacity Decisions in Supply Chains When Demand Forecasts Are Private Information: Theory and Experiment. Manufacturing & Service Operations Management, 15(1),102-117.

Josephs, R. A., Larrick, R. P., Steele, C. M., Nisbett, R. E. (1992). Protecting the self from the negative consequences of risky decisions. Journal of Personality and Social Psychology, 62(1), 26-37.

Kagel, J.H., A. E. Roth. (1995). Introduction to experimental economics. ed: Kagel. J. H. and Roth, A. E., The Handbook of Experimental Economics, Princeton University Press, Princeton, 3-109.

Kahneman, D., Slovic, P., Tversky, A. (1982). Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge.

Kahneman, D., Tversky, A. (1984). Choices, Values, and Frames. American Psychologist 39 (4): 341–350

Kalkanci, B., Chen, K., Erhun, F., (2014). Complexity as a Contract Design Factor: A Human-to-Human Experimental Study. Production and Operations Management, 23(2), 269–284.

Kalkanci, B., K. Chen, F. Erhun. (2011). Contract Complexity and Performance under Asymmetric Demand Information: An Experimental Evaluation. Management Science, 57(4), 689-704.

Katok, E., D. Y. Wu. (2009). Contracting in Supply Chains: A Laboratory Investigation. Management Science, 55(12), 1953-1968.

Katok, E., Olsen T., Pavlov V. (2014). Wholesale Pricing under Mild and Privately Known Concerns for Fairness. Production and Operations Management, 23(2). 285-302

Katok, E., Pavlov, V. (2013). Fairness in Supply Chain Contracts: A Laboratory Study, Journal of Operations Management, 31, 129-137.

Katok, E., Wu, D. Y. (2009). Contracting in Supply Chains: A Laboratory Investigation, Management Science, 55, 1953-1968.

Keser, C., Paleologo, G. A. 2004. Experimental Investigation of Supplier-Retailer Contracts: The Wholesale Price Contract. Working paper, CIRANO Scientific Series

Kremer, M. (2008). Behavioral Perspectives on Risk Sharing in Supply Chains, (Doctoral Dissertation), University of Mannheim.

Kremer, M., Minner, S., Van Wassenhove, L. N. (2010). Do Random Errors Explain Newsvendor Behavior? Manufacturing and Service Operations Management, 12(4), 673-681.

Kudryavtsev, A., Cohen, G. (2011). Behavioral biases in economic and financial knowledge: Are they the same for men and women? Advances in Management & Applied Economics, 1(1),15-52.

Kudryavtsev, A., Cohen, G., Hon-Snir, S. (2013). 'Rational' or 'Intuitive': Are Behavioral Biases Correlated Across Stock Market Investors? Contemporary Economics, 7(2), 31-53.

Lau, N., Bearden J. N. (2013). Newsvendor Demand Chasing Revisited. Management Science, 59(5), 1245-1249.

Lau, N., Bearden, J. N., Hasija, S. (2014). Newsvendor Pull-to-Center Reconsidered. Decision Support Systems, 58, 68-73.

Lee, Y. S., Siemsen E. (2016). Task Decomposition and Newsvendor Decision Making. Working Paper.

Lim, N., Ho T. (2007). Designing Price Contracts for Boundedly Rational Customers: Does the Number of Blocks Matter? Marketing Science, 26, 312-326.

Loch, C. H., Wu Y. (2008). Social Preferences and Supply Chain Performance: An Experimental Study, Management Science, 54(11), 1835-1849.

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

Maxwell S, Nye P, Maxwell N. (1999) Less pain, same gain: the effects of priming fairness in price negotiations. Psychology and Marketing.16(7):3 – 26.

Maxwell, S., Nye, P., Maxwell, N. (2003) The wrath of the fairness-primed negotiator when the reciprocity norm is violated. The Behavioral Aspect of Pricing. 46(5), 399-409

Meyers-Levy, J., Loken, B. (2015). Revisiting gender differences: What we know and what lies ahead. Journal of Consumer Psychology, 25(1), 129-149.

Morgan, J., Várdy, F. (2007). The value of commitment in contest and tournaments when observation is costly. Games and Economic Behavior. 60, 326-338.

Moritz, B. (2010). Cognition and Heterogeneity in Supply Chain Planning: A Study of Inventory Decision Making, (Doctoral Dissertation), University of Minnesota.

Moritz, B., Hill, A.V., Donohue, K. (2013). "Cognition and Individual Differences in the Newsvendor Problem: Behavior Under Dual Process Theory", Journal of Operations Management, 31, 72-85.

Ockenfels, A., Selten, R. (2015). Impulse Balance and Multiple-Period Feedback in the Newsvendor Game. Production and Operations Management, 24(12), 1901-1906.

Olivares, M., Terwiesch C., Cassorla L. (2009). Structural Estimation of the Newsvendor Model: An Application to Reserving Operating Room Time, Management Science, 54, 41-55.

Özalp, Ö. Zheng, Y. (2012). A Tale of Two Countries: Trust and Trustworthiness in China and

Özer, Ö., Zheng, Y., Chen, K. (2011). Trust in Forecast Information Sharing, Management Science, 57, 1111-1137.

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

Pavlov, V. (2009). Essays on Supply Chain Coordination, (Doctoral Dissertation), The Pennsylvania State University.

Pavlov, V., Katok E., Fairness and Coordination Failures in Supply Chain Contracts, Working Paper, (2011).

Pranoto, Y. (2005). Effects of Human Decision Bias on Supply Chain Performance, (Doctoral Dissertation), Georgia Institute of Technology.

Rabin, M., Vayanos, D. (2010). The Gambler's and Hot-Hand Fallacies: Theory and Applications. The Review of Economic Studies, 77(2), 730-778.

Ren, Y., Croson, R. (2013). Overconfidence in Newsvendor Orders: An Experimental Study. Management Science, 59(11), 2502-2517.

Rosenberg, M. (1965). “Society and the adolescent self-image”, Princeton, NJ: Princeton University Press.

Roth, A. E. (1995) Bargaining Experiments. The Handbook of Experimental Economics, eds: John H. Kagel, A. E. Roth. Princeton University Press, 1995, 253-348.

Rudi, N., Drake, D. (2014). Observation Bias: The Impact of Demand Censoring on Newsvendor Level and Adjustment Behavior. Management Science, 60(5),1334-1345.

Şahin, N. (2011). Behavioral Factors affecting the Newsvendor Problem, Masters Thesis, Sabancı University.

Sanfey, A. G. (2009). Expectations and social decision-making: Biasing effects of prior knowledge on Ultimatum responses. Mind & Society. 8, 93–107

Schmidt, K. (1993) Commitment through incomplete information in a simple repeated bargaining model. Journal of Economic Theory. 60, 114-139.

Schubert, R., Brown M., Gysler, M., Brachinger, H. W. (1999). Financial Decision-Making: Are Women Really More Risk-Averse? American Economic Review, 89(2), 381-385.

Schultz, K. L., J. O. McClain, L. Robinson, L. J. Thomas. (2007). The Use of Framing in Inventory Decisions. Working paper, Johnson School Research Papers #02, Cornell University.

Schwartz, B., Ward, A., Monterosso, J., Lyubomisky, S., White, K., Lehman, D. R. (2002). "Maximizing versus satisficing: Happiness is a matter of choice", Journal of Personality and Social Psychology, 83-5.

Schweitzer, M. E., Cachon, G. P. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. Management Science, 46(3), 404-420.

Shen, S., Wang H. (2011). Supplier's Pricing Decision When Facing Irrational Retailers, 8th International Conference on Service Systems and Service Management (ICSSSM), Tianjin, China, 2011, 1-5.

Siegel, S. (1956). Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

Solnick, S. J. (2001) Gender Differences in the Ultimatum Game. Economic Inquiry, 39 (2), 189-200.

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

Strohhecker, J., Grossler A. (2013). “Do personal traits influence inventory management performance? The case of intelligence, personality, interest and knowledge”, Int. Journal of Production Economics 142, 37-50.

Su, X. (2008). Bounded Rationality in Newsvendor Models. Manufacturing and Service Operations Management, 10(4), 566-589.

Suetens, S., Tyran, J.R. (2012). The gambler’s fallacy and gender. Journal of Economic Behavior and Organization, 83, 118-124.

Tokar, T., Aloysius J. A., Waller M. A. (2012). Supply Chain Inventory Replenishment: The Debiasing Effect of Declarative Knowledge, Decision Sciences, 43.

Tversky, A., D. Kahneman. (1981). The Framing of Decisions and the Psychology of Choice. Science, 211, 453-458.

Tversky, A.; Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science 185 (4157), 1124–1131.

Van Huyck, J. B., Battalio, R. C., Walters, M. F. (1995) Commitment versus Discretion in the Peasant-Dictator Game. Games and Economic Behavior, 10, 143-170.

Várdy, F. (2004). The value of commitment in Stackelberg games with observation costs. Games Economic. Behavior. 49, 374–400.

Villa, S., Castañeda J. A., Gonçalves P. (2014). The Effect of Power on Biased Newsvendor Order Behavior: An Experimental Study. POMS 25th Annual Conference.

Wachtel, R. E., Dexter, F. (2010). Review of Behavioral Operations Experimental Studies of Newsvendor Problems for Operating Room Management, Economics, Education, and Policy, 110, 1698-1710.

Wang, C. X., Webster, S. (2007). Channel Coordination for a Supply Chain with a Risk-Neutral Manufacturer and a Loss-Averse Retailer, Decision Sciences, 38, 361-389.

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

Weber, E.U., Blais, A., Betz, E.N. (2002). A domain-specific risk attitude scale: measuring risk perceptions and risk behaviors. Journal of Behavioral Decision Making, 15, 262–290.

Wu, D. Y. (2013). The Impact of Repeated Interactions on Supply Chain Contracts: A Laboratory Study. International Journal of Production Economics, 142, 3-15.

Wu, T., Seidman, A. (2014). Do Competitive Environments have an Effect on Managerial Decision Making? An Empirical Investigation of the Newsvendor Problem. 47th Hawaii International Conference on System Science.

Wu, Y. (2006). Social Preferences in Coordination and Delegation Decisions- Models and Empirical Tests, (Doctoral Dissertation), INSEAD.

Zeelenberg, M., Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. Organizational Behavior and Human Decision Processes, 67, 63–78

Zeelenberg, M., Beattie, J., van der Pligt, J., de Vries, N. K. (1996). Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making. Organizational Behavior and Human Decision Processes, 65(2), 148-158.

Zhang, Y., Donohue K., Cui H. T. (2016). Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing. Management Science, 62(6), 1734-1754.