RISK-AVERSE NEWSVENDOR MODELS

BY SUNGYONG CHOI

A dissertation submitted to the

Graduate School—Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Professor Andrzej Ruszczyński

and approved by

Newark, New Jersey

October, 2009

ABSTRACT OF THE DISSERTATION

Risk-Averse Newsvendor Models

by Sungyong Choi

Dissertation Director: Professor Andrzej Ruszczyński

I consider single- and multi-product risk-averse newsvendor models under two risk measures, coherent measures of risk and exponential utility function. Following from the typical format of a newsvendor model, I formulate the problems in the single- and multi-product cases and establish my models to take risk aversion into account. Thus, my models can capture the decision making of inventory managers at a different angle than most of literature in supply chain management. The key research questions are how the degree of risk aversion and product demand dependence structure interact with each other and affect jointly to the optimal decision of inventory managers. My models can find their applications in many manufacturing, distribution and retailing companies that handle short life-cycle products.

From my extensive literature review, I summarize and tabulate the literature of risk-averse inventory models and categorize typical approaches to risk-averse inventory models into four groups by the risk measures used. I discuss similarities and differences between the models. In particular, I provide clear axiomatic criteria to evaluate validity of risk measures in risk-averse newsvendor models. By the axiomatic criteria, coherent measures of risk are chosen to fit best for risk-averse newsvendor

models, but the exponential utility function is also studied for a comparison purpose. This axiomatic approach can be also applicable to other types of risk-averse inventory models.

In the main results, I study the impact of risk aversion on the optimal ordering quantity. For single-product models, I obtain closed-form optimal ordering quantity under coherent measures of risk and closed-form approximation under exponential utility function. For a large but finite number of products, I also obtain closed-form approximations under the both risk measures when product demands are independent. My approximations are as simple to compute as the risk-neutral newsvendor solutions and the gap between the optimal solutions and approximations quickly converges to zero as the number of products increases. Then I prove that the risk-neutral solution is asymptotically optimal under coherent measures of risk, as the number of products tends to be infinity. The same proposition is proved under exponential utility function, as the ratio of the degree of risk aversion to the number of products goes to zero. Thus, in both cases, risk aversion has no impact in the limit. Demand dependence significantly affects the optimal ordering quantity. I derive analytical and numerical insights for the interplay between demand correlation and risk aversion. All these results are consistent with our insights and confirmed by numerical examples from my computational study.

I conclude my dissertation by comparing risk-averse newsvendor models and financial portfolio optimization models.

Acknowledgements

I would like to first express my sincere gratitude to my advisor, Professor Andrzej Ruszczyński. I am very fortunate to be his student as a research and teaching assistant throughout my stay in the program. I have been educated to be a confident future professor by his patient directions from his rich and lifelong wisdoms and experiences. I am also very fortunate, under his guidance, to find so innovative and valuable research topic that I will dedicate myself utterly to work with. I can see his enthusiasm and passion as well as scientific expertise which has been my continuous source of my career motivation.

I am also very much indebted to him for his kind help and support during my dissertation work. I deeply appreciate his valuable advice and feedback on my research. I am grateful for his help and encouragement to develop my teaching and communication skills and to search for my next place. It has been a great honor to be a student of such a distinguished professor, who is widely recognized as a leading figure in this area.

I am grateful sincerely to Professor Yao Zhao, a member of my dissertation committee. His kind and active involvement in my dissertation work helped my work position better in this area. His valuable advice and feedback also helped me to develop my teaching and presentation skills. I am very much honorable to work with such a rising professor in this area from his proven scholarship.

I want to thank other members of my dissertation committee, Professors Jian Yang, Darinka Dentcheva, Ronald Armstrong and Michael Katehakis. Especially, I appreciate Professor Yang to accept my request to be a reference for my job search. His help and encouragement was substantial for me. I also appreciate the opportunity, given by Professor Dentcheva, to present my preliminary results of my dissertation at the 4^{th} Rutgers-Stevens Workshop Optimization of Stochastic Systems that she organized at the Stevens Institute of Technology in March 2007. I also thank Professor Armstrong's and Professor Katehakis' advices and feedback to me during my study in the program.

I have had an excellent opportunity at MSIS (Management Science and Information Systems) and SCMMS (Supply Chain Management and Marketing Sciences) departments, RUTCOR (RUTgers Center of Operations Research) and NJIT (New Jersey Institute of Technology) to work with other faculty members, students and staff members. I enjoyed interacting with all of them. I especially thank my colleagues, Adam Fleischhacker, Sitki Gulten, Rose Kiwanuka, Kathleen Martino, Mojisola Otegbeye, Dinesh Pai, Srinivasa Puranam, Junmin Shi and Gang Wang for their fruitful discussions and supports.

While at Korea University, I studied and worked with Professors S. Sung, K. Whang, M. Park, K. Park, D. K. Kim, D. S. Kim, H. Rhim, and H. Shin, who provoked my interests into research in general, and greatly encouraged me to pursue my Ph.D studies in America. I also send my gratitude to them.

I also thank all my friends of high school and university alumni living in Korea and locally at New York and New Jersey. I especially thank Jaehoon Kim for his generous and consistent support. His confidence in me has given me the courage to pursue my studies.

Last, but not the least, I would like to thank my parents and brother. With their help and encouragements, I was able to finish my Ph.D. study. I owe them everything. This Ph.D. degree is theirs.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 Motivations

Stochastic inventory management problems are at the center of interests in Operations Management. In these problems, inventory managers have to decide optimal ordering quantities before random demands are realized. The firm's objective is to determine the optimal solutions for products so as to optimize certain performance measures.

My dissertation research studies how risk-averse and rational inventory managers decide their optimal choices under risk. In particular, I focus on how risk aversion of inventory managers affects their optimal decision in the newsvendor model. My interests in this line of research originated when I noticed a gap on the impacts of risk aversion in the inventory management literature. That is, the inventory management literature mainly assumes that inventory managers aim to maximize the expected value of profits or minimize the expected value of costs, but does not consider variations of random outcomes. However, inventory managers may not be ready to possibly suffer significant losses with the hope of obtaining large profits. Instead, they may prefer worse *on average* but stable outcomes. Schweitzer and Cachon (2000) provide experimental evidence suggesting that inventory managers may have different attitude of risk preferences depending on product characteristics. More specifically, inventory managers can be risk-averse for high-value products.

By such reasons, risk aversion can capture the decision making of inventory managers at a different angle than most of literature. Risk aversion is consistent to rational inventory managers and it significantly affects the optimal strategy of inventory managers. These make risk aversion a very interesting and important factor to be considered when one analyzes the optimal choices of inventory managers. However, in the supply chain management literature, few attempts have been made so far.

In recent years, risk-averse inventory models have received increasing attentions from academia and industry. Risk-averse optimization was invited to the tutorial sessions in both of 2007 and 2008 INFORMS (The Institute for Operations Research and Management Sciences) annual conferences. An important and exogenous event is the on-going subprime mortgage financial crisis, which motivates a number of academicians and practitioners to recognize the importance of evaluating and managing risks in their decision making. With such increased attention from academia and industry, my research is very timely and interesting topic in Operations Management. My models can find their application in many manufacturing, distribution and retailing companies and also military applications, where inventory managers cannot be assumed to be risk-neutral.

1.2 Objectives

In my dissertation, I study risk-averse newsvendor problems. I divide the problems into three independent but closely related essays. The first essay is a single-product risk-averse newsvendor model with coherent measures of risk and the second and third essays are multi-product risk-averse newsvendor models with coherent measures of risk and exponential utility function, respectively.

Following from the typical format of the classical newsvendor model, there is a single product (the first essay) or multiple products (the second and third essays) with random demands to be sold in a single-selling season. Backordering is not allowed in all three models. The vendor has to order the items before the demand is realized and the demand is only known as the form of its probability distribution. On the one hand, when demand exceeds supply for any product, the excessive demand is lost. On the other hand, when supply exceeds demand, the excessive inventory is sold at a loss.

In the classical model, the objective function is to maximize the expected value of the total profit and does not consider risk. It has a well-known simple analytical solution. Due to its simplicity and versatility, there exist many applications, such as plant capacity or overbooking problems. However, the outcomes which are actually observed by inventory managers are random, and they cannot always rely on repeated similar chances. Especially for short product-life cycle products, the first few outcomes may turn out to be very bad due to the variability of the outcomes, and entail unacceptable losses. Therefore, I aim to replace the risk-neutral performance measure by new measures taking risk aversion into account.

1.3 Contributions to the Literature

From the three models to be studied in my dissertation, I aim to contribute to the literature as follows:

1. By my extensive literature review, I categorize typical approaches to risk-averse inventory models into four groups by the risk measures used and outline their similarities and differences. Then, I summarize and tabulate the literature of risk-averse inventory models by model types and the four groups of the risk measures in a matrix.

2. I provide an axiomatic approach to evaluate validity of risk measures in risk-averse newsvendor models. This approach can be applicable to other risk-averse inventory models.

3. I develop risk-averse inventory models for single- and multi-product cases under general law-invariant coherent measures of risk and exponential utility function.

4. For single-product models,

- I study the impact of the degree of risk aversion under general law-invariant coherent measures of risk and exponential utility function.
- I obtain a closed-form optimal solution under general law-invariant coherent measures of risk and closed-form approximation under exponential utility function.

5. For multi-product models with independent demands case,

    - I prove convexity of the objective functions under general law-invariant coherent measures of risk and exponential utility function.
    - I study the impact of the degree of risk aversion under general law-invariant coherent measures of risk and exponential utility function.
    - I obtain closed-form approximations under general law-invariant coherent measures of risk and exponential utility function.
    - I analyze the asymptotic behavior of the solution under general law-invariant coherent measures of risk with respect to the number of products. I also analyze the asymptotic behaviors of the solution under exponential utility function with respect to the ratio of the number of products to the degree of risk aversion.

6. For the multi-product models with dependent demands case,

    - I derive analytical and numerical insights on the interplay of demand correlation and risk aversion under general coherent measures of risk and exponential utility function.

7. I conduct an extensive computational study to confirm the analytical results and derive numerical insights.

8. I compare multi-product newsvendor models to financial portfolio optimization models.

1.4 Structure

The remainder of the dissertation is organized as follows: Chapter 2 categorizes typical approaches to risk-averse inventory models into four groups - Expected Utility Theory, Stochastic Dominance, Chance Constraints and Mean-Risk Analysis. I discuss their similarities and differences. Then, I also summarize and tabulate the literature of risk-averse inventory models by model types and risk measures in a 5-by-5 matrix. Finally, chapter 2 provides a clear axiomatic approach to evaluate validity of each risk measure in risk-averse newsvendor models.

Chapter 3 studies single-product newsvendor models. It focuses on the impact of the degree of risk aversion on the optimal ordering quantity and closed-form optimal solutions under mean-deviation from quantile and general law-invariant coherent measures of risk.

Chapter 4 is dedicated to the multi-product newsvendor model under general law-invariant coherent measures of risk. The key research questions are to analyze the impacts of risk aversion and demand dependence to the optimal order quantity.

Chapter 5 is dedicated to the multi-product newsvendor model under exponential utility function. The main results of this Chapter are well in harmony with Chapter 4.

Chapter 6 is dedicated to computational study to confirm the analytical results studied from Chapter 3 through 5, and derive insights for risk-averse newsvendors.

Chapter 7 contains concluding remarks and compares multi-product newsvendor models and financial portfolio optimization models.

Chapter 2
Inventory Models under Risk

2.1 Decision Making under Risk

Most of decision theory is *normative*. That is, it analyzes the optimal choice of an ideal decision maker with full rationality based on hypotheses of how people act. However, since we know that sometimes people do not make their decision in an optimal way, there is another approach to describe what decision makers actually do and we call it *descriptive* decision analysis. These two approaches are closely related to each other, *i.e.*, it is possible to relax normative assumptions to study what people actually do. On the other hand, it is helpful to construct norms from revealing the actual decision making.

In my dissertation, I assume that decision makers act in an optimal way. That is, they want to optimize certain performance measures. Inventory managers make their decisions on behalf of their organizations, but I exclude the possibility of any principal-agent problems. Typically they have historical data for their decisions and each outcome can be assumed easily to occur with a known probability. Under such situation, the chance of non-optimal behavior is minimized.

The term *risk* has different meanings to different people. In my dissertation, I used the word as uncertainty of random outcome. More specifically, it implies that only the probability for every possible random outcome is given in advance. Then decision makers have different risk preferences to random outcomes and the risk preferences can be categorized into three cases: risk-neutral, risk-averse and risk-seeking behaviors.

Risk-neutral decision makers focus only the expected value regardless of the risk of the outcome. From that sense, risk-neutral decision making is equivalent to expected-value optimization model. Risk-averse decision making comes from the natural dilemma of the decision makers who want risk premiums for uncertain outcomes. It is more relevant to conservative decision making, but the degree of risk aversion may be different for each risk-averse decision maker. Risk-seeking decision maker compares the best possible outcome from risky performance options and fixed outcomes from certain performance options. Then, they choose the option of the best possible outcome unless the guaranteed outcomes are better than the best possible outcome among the risky performance options.

Among these three risk preferences, inventory managers are generally assumed not to be risk-seeking. Actually, risk-seeking behaviors are more concerned with the psychological aspects of individuals to buy lotteries or gambling. However, these situations are very different from those for inventory managers.

The risk-neutral decision making provide the best decision *on average*. This may be justified by the Law of Large Numbers. When performance measure represents profit, higher measured value of the performance is always better than lower value of the performance measures. Then I will represent a risk-neutral optimization model.

Consider an optimization model where the decision vector x affects a random performance measure, ϕ_x. Here, for all $x \in \mathscr{X}$ with $\mathscr{X}$ being a vector space, $\phi_x : \Omega \to \mathbb{R}$ is a measurable function on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ where Ω is the sample space, $\mathscr{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on Ω. Then, the risk-neutral optimization model consider the expected-value optimization problem such as:

$$\max_{x \in X} \mathbb{E}[\phi_x]. \tag{2.1}$$

where $X \subseteq \mathscr{X}$ is a feasible set. All the considerations can be modified easily for the case of the reverse preference when lower measured value is preferred to higher measured value. However, if one is concerned with few (or just one) realizations and the

Law of Large Numbers cannot be invoked, this formulation is inappropriate. Consequently the attempts to overcome the drawbacks of the expected value optimization have a long history, and I will describe them in detail in the following sections.

2.2 Risk Measures in Inventory Models

In my dissertation, I focus on risk-averse inventory models and risk neutrality is considered as a reference for comparison purposes. By my extensive literature review, I classify typical approaches to risk-averse inventory models into four groups by the risk measures used. They are expected utility theory, stochastic dominance, chance constraints and mean-risk analysis. (Cumulative) Prospect theory by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) mainly focus on the descriptive model of individual decision makers. It can explain why a single individual decision maker buys insurance and lotteries simultaneously, which is not explained by expected utility theory. However, this situation can be explained by the third-order stochastic dominance relation and is very different from that in inventory models. Therefore, I do not separate it as one of the independent categories.

2.2.1 Expected Utility Theory

The expected utility model was initiated by Daniel Bernoulli as a solution of the well-known St. Petersburg paradox. Then, the modern theory of the expected utility by von Neumann and Morgenstern (1944) derives, from simple axioms, the existence of a nondecreasing utility function, which transforms (in a nonlinear way) the observed outcomes. That is, each rational decision maker has a nondecreasing utility function $u(\cdot)$ such that he prefers random outcome X over Y if and only if $\mathbb{E}[u(X)] > \mathbb{E}[u(Y)]$, and then he optimizes, instead of the expected outcome, the expected value of the utility function. Therefore, the decision maker solves the following optimization model.

$$\max_{x \in X} \mathbb{E}[u(\phi_x)]. \tag{2.2}$$

In the maximization context, when the outcome represents profit, risk-averse decision makers have nondecreasing and concave utility functions. An important example of nondecreasing and concave utility function is the exponential utility function and the optimization model can be represented such as

$$\max_{x \in X} \mathbb{E}\Big[-e^{-\phi_x} \Big]. \tag{2.3}$$

Expected utility theory and its properties are intuitive and elegant but practically ineffective construct. The critics of the expected utility theory raise the following issues: First, it is very difficult or even impossible to elicit the exact functional form of the utility function for each decision maker. Roy (1952) argued that *"A man who seeks advice about his actions will not be grateful for the suggestion that he maximize expected utility."* This problem can be mitigated in the stochastic dominance approach, which is an abstract generalization of expected utility theory. Second, there exist some counter-intuitive examples to the expected utility theory such as Allais Paradox (see Allais (1953)) and (Modified) Ellsberg Paradox (see Ellsberg (1961)).

2.2.2 Stochastic Dominance

The second category is based on the theory of *stochastic dominance*, developed in statistics and economics (see Lehmann (1955), Quirk and Saposnik (1962), Hadar and Russell (1969), Rothschild and Stiglitz (1969) and Hanoch and Levy (1969)). Stochastic dominance relations are sequence of partial orders defined on the space of random variables, and thus allow for pairwise comparison of different random variables (see Whitmore and Findlay (1978) and Levy (1992)). It originated from the majorization theory (see Hardy, Littlewood and Polya (1934) and Marshall and Olkin (1979)) for the discrete case and was later extended to general distributions, and is now widely used in economics and finance. Detailed and comprehensive discussion of the concept of stochastic dominance and its relation to downside risk measures is given at Ogryczak and Ruszczyński (1999 and 2001).

An important feature of the stochastic dominance theory is its universal character with respect to utility functions. More specifically, the distribution of a random outcome X is preferred to random outcome Y in terms of a stochastic dominance relation if and only if expected utility of X is preferred to expected utility of Y for all utility functions in a certain class, called the generator of the relation. Thus, stochastic dominance is an abstract generalization of utility function approach which eliminates the need to specify explicitly a certain inventory manager's utility function.

Among this sequence of relations of stochastic dominance, the second-order stochastic dominance relation, in particular, corresponds to all nondecreasing and concave utility functions, and is thus well suited to model risk-averse preferences. For an overview of these issues, see Fishburn (1964), Whitmore and Findlay (1978), Bawa (1982), Levy (1992 and 2006) and Müller and Stoyan (2002).

The strongest preference relation is the statewise stochastic dominance. Suppose that there are two random variables, X and Y defined on the same measurable probability space $(\Omega, \mathscr{F}, P)$. Then, the statewise dominance of X to Y is equivalent that X has at least as good outcome as Y in every possible state of nature. Under this statewise dominance, anyone who prefers better to worse will always choose X. Thus, one does not have to consider any performance functions and it is not exactly related to the axioms of risk preference. This relation is exemplified in the following table.

	ω_1	ω_2
X	-1	3
Y	-1	-1

Table 2.1: Example of the Statewise Dominance.

In Table 2.1, I set up $\Omega = \{\omega_1, \omega_2\}$ and $\mathbb{P}(\omega_1)$ and $\mathbb{P}(\omega_2) = 1 - \mathbb{P}(\omega_1)$ can be assigned arbitrarily from zero to one. I assume that larger value is always preferred to smaller value in this table. Each random variable X and Y has a value for any possible states of nature, ω_1 and ω_2, and $X(\omega)$ is always better than $Y(\omega)$ for ω_1 and ω_2. Thus, X dominates Y by the rule of statewise dominance.

Then, the second and third strongest preference relations are the first- and second-order stochastic dominance relations and then the $(k+1)^{th}$ strongest preference relation is the k^{th}-order stochastic dominance relation for $k \in \mathbb{N}$. In the each degree of stochastic dominance relations, random variables are compared pairwisely by the same degree of performance function constructed from their distributions.

The first performance function, $F_X^{(1)}(\eta)$, is defined as:

$$F_X^{(1)}(\eta) = F_X(\eta) = P(X \leq \eta), \quad \forall \eta \in \mathbb{R}. \tag{2.4}$$

where $F_X(\eta) = P(X \leq \eta)$ denotes the right-continuous cumulative distribution function of a random variable X itself. Then, the weak relation of the *first-order stochastic*

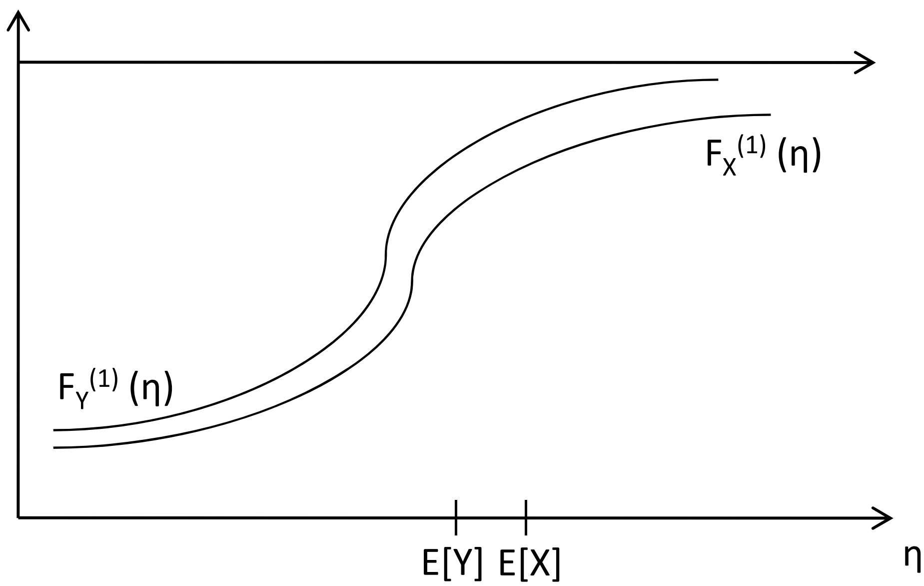

Figure 2.1: The First-Order Stochastic Dominance

dominance (FSD), which we denote by $X \succeq_{(1)} Y$, is illustrated at Figure 2.1 and defined as follows (see Lehmann (1955) and Quirk and Saposnik (1962)):

$$X \succeq_{(1)} Y \Leftrightarrow F_X^{(1)} \leq F_Y^{(1)}, \quad \forall \eta \in \mathbb{R}. \tag{2.5}$$

In this case, it is said that X dominates Y under the rule of the first stochastic dominance. It implies that for any outcome $\eta \in \mathbb{R}$, X gives a higher probability of receiving the outcome equal to or better than η under Y.

The second performance function, $F_X^{(2)}(\eta)$, is given by the areas below the distribution function F_X,

$$F_X^{(2)}(\eta) = \int_{-\infty}^{\eta} F_X(\alpha) d\alpha, \quad \eta \in \mathbb{R}.$$

It is well known that if $\mathbb{E}[|X|] < \infty$, the function $F_X^{(2)}(\eta)$ is well defined for all $\eta \in \mathbb{R}$.

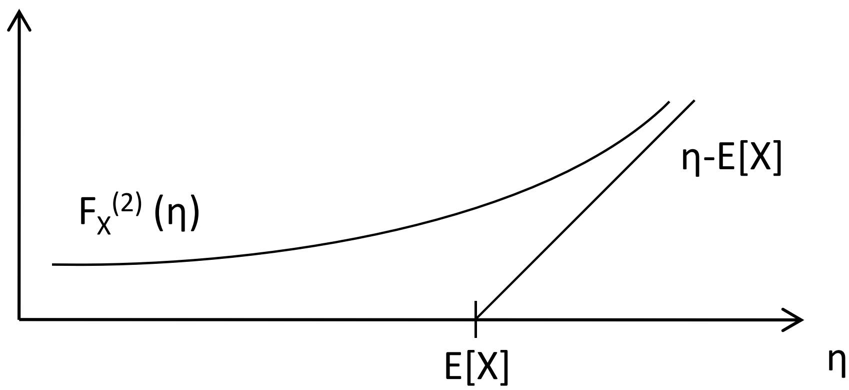

Figure 2.2: The Outcome-Risk Diagram

Then, $F_X^{(2)}(\eta)$ is continuous, convex and nonnegative and nondecreasing. In addition, if $F_X(\eta^0) > 0$, then $F_X^{(2)}(\eta)$ is strictly increasing, $\forall \eta \geq \eta^0$. The graph of the function $F_X^{(2)}$ can be illustrated at Figure 2.2 as the Outcome-Risk (O-R) diagram (refer to Ogryczak and Ruszczyński (2002)) where it has two asymptotes intersecting at the point $(\mu_X, 0)$. More specifically, the η-axis is the left asymptote and the line $\eta - \mu_X$ is the right asymptote.

Then, the weak relation of the *second-order stochastic dominance* (SSD), which I denote by $X \succeq_{(2)} Y$, is defined as follows (see Hadar and Russell (1969) and Hanoch and Levy (1969)):

$$X \succeq_{(2)} Y \Leftrightarrow F_X^{(2)} \leq F_Y^{(2)}, \quad \forall \eta \in \mathbb{R}. \tag{2.6}$$

The second-order stochastic dominance relation has a crucial role for decision making under risk. The function $F_X^{(2)}$ can also be expressed as the expected shortfall (see

Ogryczak and Ruszczyński (1999)). That is, for each target value η we have,

$$\begin{aligned} F_X^{(2)}(\eta) &= \int_{-\infty}^{\eta} (\eta - \xi) P_X(d\xi). \\ &= \mathbb{E}[\max\{\eta - X, 0\}] = \mathbb{P}\{X \leq \eta\}\mathbb{E}\{\eta - X | X \leq \eta\}. \\ &= \mathbb{E}[(\eta - X)_+]. \end{aligned}$$

Therefore, the second-order stochastic dominance relation can be represented by the following inequality equivalently (see Noyan (2006)) and can be regarded as a continuum of integrated chance constraints:

$$\mathbb{E}[(\eta - X)_+] \leq \mathbb{E}[(\eta - Y)_+], \quad \forall \eta \in \mathbb{R}.$$

In Tables 2.2 and 2.3, we set up $\Omega = \{\omega_1, \omega_2\}$ and $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = 0.5$, and then the

	ω_1	ω_2
X	-1	3
Y	1	-1

Table 2.2: Example of the First-Order, but not Statewise Dominance.

	ω_1	ω_2
X	4	2
Y	1	5

Table 2.3: Example of the Second-Order, but not First-Order Stochastic Dominance.

random variable X and Y have different values for each possible state of nature, ω_1 and ω_2. By the definitions of the first- and second-order stochastic dominance relations, Tables 2.2 and 2.3 clearly shows an example of the first and second-order stochastic dominance relations, respectively.

For higher-order stochastic dominance relations, the k^{th} performance function can be defined recursively with $X \in \mathscr{L}_m(\Omega, \mathscr{F}, \mathbb{P})$ such as:

$$F_X^{(k)}(\eta) = \int_{-\infty}^{\eta} F_X^{(k-1)}(\xi) d\xi, \quad \forall \eta \in \mathbb{R}, \quad k = 3, 4, \ldots, m+1.$$

Then, the weak relation of the k^{th}*-order stochastic dominance* (kSD), which I denote by $X \succeq_{(k)} Y$, is defined as follows (see Ogryczak and Ruszczyński (2001)):

$$X \succeq_{(k)} Y \Leftrightarrow F_X^{(k)}(\eta) \leq F_Y^{(k)}(\eta), \quad \forall \eta \in \mathbb{R}. \tag{2.7}$$

The corresponding strict dominance relation $\succ_{(k)}$ is also defined similarly as follows:

$$X \succ_{(k)} Y \Leftrightarrow X \succeq_{(k)} Y \text{ and } X \not\preceq_{(k)} Y. \tag{2.8}$$

Thus, it is said that X dominates Y *by the kSD rule*, if $F_X^{(k)}(\eta) \leq F_Y^{(k)}(\eta)$, $\forall \eta \in \mathbb{R}$, with strict inequality holding for at least one η. In the equations (2.7) and (2.8), one implicitly assumes that the functions $F_X^{(k)}$ and $F_Y^{(k)}$ are well defined; this is guaranteed when $\mathbb{E}(|X|^{k-1})$ and $\mathbb{E}(|Y|^{k-1})$ are defined finitely.

Clearly $X \succeq_{(k-1)} Y$ implies $X \succeq_{(k)} Y$ and $X \succ_{(k-1)} Y$ implies $X \succ_{(k)} Y$ when the k^{th} performance function $F_X^{(k)}$ is well defined. In the remainder of my dissertation, the term of *stochastic dominance relation* refers to the weak relation of the corresponding degree of stochastic dominance without mentioning explicitly.

As the special case of higher-order stochastic dominance relations, the third-order stochastic dominance (TSD) is consistent to *Ruin Aversion*. TSD assumes that decision makers prefer positive skewness to negative skewness as the third derivative of the utility is positive. It can explain that some decision makers buy insurance and lottery at the same time whose utility function is a S-shaped curve. It is also consistent with (Cumulative) Prospect Theory (see Kahneman and Tverski (1979) and Tverski and Kahneman (1992)).

Unfortunately, the stochastic dominance approach does not provide us with a simple computational recipe. In fact, it is a multiple criteria model with a continuum of criteria. Therefore, it has been used as a constraint (Dentcheva and Ruszczyński (2003) and Dentcheva and Ruszczyński (2004)), and also utilized as a reference standard whether a particular solution approach is appropriate (Ogryczak and Ruszczyński

(1999) and Ruszczyński and Vanderbei (2003)). Then I consider the following optimization model with the k^{th}-order stochastic dominance constraint:

$$\begin{aligned} \max \quad & \mathbb{E}[\phi_x] \\ \text{subject to} \quad & \phi_x \succeq_{(k)} Y. \\ & x \in X. \end{aligned}$$

Such models use random benchmark performance and find decision vectors x to maximize the expected outcome at which the random outcome ϕ_x is more preferred to the benchmark random performance. Because the second-order stochastic dominance is consistent to risk-averse preference, this condition gives an indirect way to formulate risk-aversion when k is equal to two.

2.2.3 Chance Constraints

The third category specifies certain constraints on the probabilities that measure the risk such as:

$$\mathbb{P}(\phi_x \geq \eta) \geq 1 - \alpha. \tag{2.9}$$

where η is a fixed target value and $\alpha \in (0, 1)$ is the maximum level of risk of violating the stochastic constraint, $\phi_x \geq \eta$. Then, I consider the following optimization model.

$$\begin{aligned} \max \quad & \mathbb{E}[\phi_x] \\ \text{subject to} \quad & \mathbb{P}(\phi_x \geq \eta) \geq 1 - \alpha. \\ & x \in X. \end{aligned}$$

Such models were initiated and developed in Charnes, Cooper and Symonds (1958), Miller and Wagner (1965) and Prekopa (1970). Recently Prekopa (2003) provides a thorough overview of the state of the art of the optimization theory with chance constraints. Theoretically, a chance constraint is a relaxed version of the first-order stochastic dominance relation, and thus it is related to the expected utility theory, but there is no equivalence. However, chance constraints sometimes lead to non-convex formulations of the resulting optimization problems.

In connection with the stochastic constraints models, some variations and extensions have been suggested. They are the collection of the conditional expectations and the conditional expectation constraints by Prekopa (1973), integrated chance constraints by Klein Haneveld (1986 and 2002), value-at-risk constraints by Dowd (1997), conditional value-at-risk constraints by Rockafellar and Uryasev (2000 and 2002) and Pflug (2000) and expected shortfall constraints by Acerbi and Tasche (2002).

In finance, chance constraints are very popular and have been actively used under the name of *Value-at-Risk* constraints. In finance literature, the α-quantile of the random performance X in a profit maximization context is defined as follows:

$$\inf\{\eta : F_X(\eta) \geq \alpha\}.$$

and the α-quantile is called the Value-at-Risk (VaR) *at the confidence level* α and *denoted by* $\mathrm{VaR}_\alpha(X)$, $\alpha \in (0, 1]$.

2.2.4 Mean-Risk Analysis

The last category, originating from finance, is the *mean-risk* analysis. It quantifies the problem in a lucid form of two criteria: the *mean* (the expected value of the outcome), and the *risk* (a scalar measure of the variability of the outcome). In mean-risk analysis, one uses a specified functional $r : \mathscr{X} \to \mathbb{R}$, where $\mathscr{X}$ is a certain space of measurable functions on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, to represent variability of the random outcomes, and then solves the problem:

$$\min_{x \in X}\{-\mathbb{E}[\phi_x] + \lambda r[\phi_x]\}. \tag{2.10}$$

Here, λ is a nonnegative trade-off constant between the expected outcome and the scalar-measured value of the variability of the outcome. This allows a simple trade-off analysis analytically and geometrically.

In the minimization context, one selects from the universe of all possible solutions those that are *efficient*: for a given value of the mean he minimizes the risk, or equivalently, for a given value of risk he maximizes the mean. Such an approach has many advantages: it allows one to formulate the problem as a parametric optimization problem, and it facilitates the trade-off analysis between mean and risk. However, for some popular dispersion statistics used as risk measures, the mean-risk analysis may lead to inferior conclusion. Thus, it is of primary importance to decide a good risk measure for each type of the decision models to be considered.

In the context of portfolio optimization, Markowitz (1952 and 1959) used the variance of the return as the risk functional, i.e.

$$r[\phi_x] = \mathbb{V}ar[\phi_x]] = \mathbb{E}[(\phi_x - \mathbb{E}[\phi_x])^2].$$

It is easy to compute, and it reduces the financial portfolio selection problem to a parametric quadratic programming problem. Since an approximation to the mean-variance model initiated by Sharpe (1971), many attempts have been made to linearize the portfolio optimization problem. This resulted in the consideration of various risk measures which were the multiple criteria LP computable model in the case of finite discrete random variable based on majorization theory (see Hardy *et al.* (1934) and Marshall and Olkin (1979)) and Lorenz-type orders (see Lorenz (1905) and Arnold (1980)).

However, one can construct simple counterexamples that show imperfection of the variance as the risk measure: it shows a symmetric property and treats over-performance equally as under-performance. More importantly, it may suggest a portfolio which is stochastically dominated by another portfolio. Again, Table 2.1 is a good example where an efficient solution (in the sense from mean-risk analysis) is dominated by another solution. Clearly, X may be preferred to Y by the rule of statewise dominance in Table 2.1. However, $\mathbb{E}(Y) - 1 \cdot \mathbb{V}ar(Y) = -1 > -3 = \mathbb{E}(X) - 1 \cdot \mathbb{V}ar(X)$. This implies that Y is more preferable to X under mean-variance criterion, which is contradictory to the concept of stochastic dominance.

To overcome the drawbacks of the mean-variance analysis, the general theory of *coherent measures of risk* was suggested by Artzner, Delbaen, Eber and Heath (1999) and extended to general probability spaces by Delbaen (2002). For further generalizations, see Föllmer and Schied (2002,2004), Kusuoka (2003) and Ruszczyński and Shapiro (2005, 2006a). Dynamic version for a multi-period case were analyzed, among others, by Riedel (2004), Kusuoka and Morimoto (2004), Ruszczyński and Shapiro (2006b).

In this theory, an integrated performance measure is considered, comprising both mean and certain variability measure, and several transparent axioms are imposed to guarantee consistency with intuition about rational risk-averse decision making. One need to note that the term *risk* means a scalar-valued measure of variability in mean-risk analysis. However, in coherent measures of risk, the same word refers to an integrated performance measure comprising of mean and variability measure simultaneously. Therefore, coherent measures of risk are extensions of the mean-risk analysis. It is known that coherent measures of risk are consistent with the first- and second-order stochastic dominance relations (see Shapiro, Dentcheva and Ruszczyński (2009)).

2.2.5 Coherent Measures of Risk

Artzner *et al.* (1999) initiated the general theory of coherent measures of risk, by specifying a number of axioms that these measures should satisfy. Here I provide a brief information sufficient for my purposes. I follow the abstract approach of Ruszczyński and Shapiro (2005, 2006a).

Let $(\Omega, \mathscr{F})$ be a certain measurable space. In my case, Ω is the probability space on which D is defined. An uncertain outcome is represented by a measurable function $X : \Omega \to \mathbb{R}$. Then, I assume that larger value of X is preferred to smaller value of X. I specify the vector space $\mathscr{X}$ of possible functions; in my case it is sufficient to consider $\mathscr{X} = \mathscr{L}_\infty(\Omega, \mathscr{F}, P)$. Indeed, for a fixed decision variable x, the function $\omega \to X(\omega)$ is bounded. For any X and $Y \in \mathscr{X}$, I denote $X \geq Y$ if and only if $X \geq Y$ with probability one.

A *coherent measure of risk* is a functional $\rho : \mathscr{X} \to \mathbb{R}$ satisfying the following axioms:

Convexity: $\rho(\alpha X + (1-\alpha)Y) \leq \alpha\rho(X) + (1-\alpha)\rho(Y)$, for all $X, Y \in \mathscr{X}$ and all $\alpha \in [0, 1]$;

Monotonicity: If $X, Y \in \mathscr{X}$ and $X \geq Y$, then $\rho(X) \leq \rho(Y)$;

Translation Equivariance: If $a \in \mathbb{R}$ and $X \in \mathscr{X}$, then $\rho(X + a) = \rho(X) - a$;

Positive Homogeneity: If $t \geq 0$ and $X \in \mathscr{X}$, then $\rho(tX) = t\rho(X)$.

All these axioms are defined in the minimization context. Thus, lower measured risk is always better than higher measured risk under coherent measures of risk.

Coherent measures of risk are special cases of mean-risk analysis. It implies that, for certain variability measure $r[\cdot]$ and a given range of $\lambda > 0$, the optimization model (2.10) satisfies all the four axioms – Convexity, Monotonicity, Translation Equivariance and Positive Homogeneity. Thus, if a mean-risk model satisfies the four axioms, then, with the abstract approach of Ruszczyński and Shapiro (2005, 2006a), the corresponding optimization model (2.10) can be reformulated as follows:

$$\min_{x \geq 0} \rho(X) = -\mathbb{E}[X] + \lambda r[X]. \tag{2.11}$$

A coherent measure of risk $\rho(\cdot)$ is called *law-invariant*, if the value of $\rho(X)$ depends only on the distribution of X, that is $\rho(X_1) = \rho(X_2)$ if X_1 and X_2 have identical distributions. Acerbi (2004) summarizes the meaning of this property that *a law-invariant coherent measure of risk gives the same risk for empirically exchangeable random outcomes*. Law-invariance looks so obvious that it is no wonder even if most risk practitioners take it for granted. However, it also implies that, for a coherent measure of risk ρ which is not law-invariant, $\rho(X_1)$ and $\rho(X_2)$ may be different even if X_1 and X_2 have same probability distribution. This apparent paradox can be resolved by reminding the formal definition of random variables. Actually, one needs to determine simultaneously ”probability law” and ”field of events” endowed with a σ-algebra structure to define a random variable. Thus, the two random variables with same probability distributions can be different and may have different values of ρ. An example of the coherent measure of risk which is not law-invariant is the so-called *worst conditional* *expectation* WCE_α defined in Artzner *et al.* (1999).

$$WCE_\alpha = -\inf\{\mathbb{E}[X|A] : A \in \mathscr{A}, \mathbb{P}[A] > \alpha\}$$

The infimum of conditional expectations $\mathbb{E}[X|A]$ is taken on all the events A with probability larger than α in the σ-algebra $\mathscr{A}$. However, under certain conditions on nonatomic probability space, this risk measure becomes law-invariant and coincides with *expected shortfall* described in (2.14) and (2.15). For more technical details, see Acerbi and Tasche (2002), Delbaen (2002) and Kusuoka (2003).

An important sequence of these four axioms is that ρ is associated with a convex set $\mathscr{P}$ of probability measures, such that the following dual representation holds:

$$\rho[X] = \sup_{P \in \mathscr{P}} \{\mathbb{E}^P[X]\}, \ \forall X \in \mathscr{X}.$$

where $\mathbb{E}^P[X]$ is the expectation of X under the probability measure P. Therefore, a coherent measure of risk is equivalent to calculating the maximum expectation under different distributions, and thus justify approaches used in practice. Artzner *et al.* (1999) also provided an alternative definition of coherent measures of risk by specifying *acceptance sets*. For a thorough discussion of the mathematical properties of coherent measures of risk, see Artzner *et al.* (1999), Föllmer and Schied (2004) and Ruszczyński and Shapiro (2005).

In my dissertation, I focus on the meaning of these four axioms in newsvendor models. These axioms imply the following properties:

- Convexity axiom means that the global risk of a portfolio should be equal or less than the sum of its partial risks. Thus, this axiom provides the diversification effects and this is especially valid in newsvendor models. Each product is very likely to have some nonzero value in newsvendor models because very small amounts will be sold almost always for each product.

- Monotonicity axiom is consistent to statewise dominance.

- Translation Equivariance axiom means that adding a constant cost is equivalent to increasing the vendor's performance measure by the same amount. On the contrary, adding a constant gain is equivalent to decreasing the vendor's performance measure by the same amount. More specifically, by excluding the impact of constant gains or losses, fixed parts can be separated equivalently from the vendor's random performance measure at every possible state of nature. Thus, this axiom allows one to draw a comparison between the only random parts of different random performance measures and thus rank risk properly.

- Positive Homogeneity axiom guarantees that the optimal solution is invariant to rescaling of units such as currency (e.g., from dollars to yuans) or considering the total profit or the average profit per product. More importantly, this axiom guarantees no diversification effects when demands are completely correlated. To see this, one notes that the subadditivity property, $\rho(X+Y) \leq \rho(X) + \rho(Y)$, is essential to risk measures that exhibit portfolio/diversification effect because it means that the risk measure of the sum is better than the sum of individual risk measures. Subadditivity implies $\rho(nX) \leq n\rho(X)$. However $\rho(nX) < n\rho(X)$ indicates that a diversification effect exists even when the random demands are completely correlated. To avoid this counter-intuitive effect, one is left with $\rho(nX) = n\rho(X)$ which is the Positive Homogeneity axiom.

Because coherent measures of risk are any risk functionals ρ which satisfy the four axioms, their functional forms are not determined uniquely. The popular examples of functionals $r[\cdot]$ are the *semideviation* of order $p \geq 1$:

$$\sigma_p[X] = \mathbb{E}[(\mathbb{E}[X] - X)_+^p]^{\frac{1}{p}}. \tag{2.12}$$

and *weighted mean deviation from quantile*:

$$r_\beta[X] = \min_{\eta \in \mathbb{R}} \mathbb{E}[\max((1-\beta)(\eta - X), \beta(X - \eta))], \quad \beta \in (0,1). \tag{2.13}$$

The optimal η in the problem above is the β-quantile of X. Optimization models with functionals (2.12) and (2.13) were considered in Ogryczak and Ruszczyński (1999, 2001 and 2002). In the maximization context, from the practical point of view, it is most reasonable to consider $\beta \in (0, 1/2]$, because then $r_\beta[X]$ penalizes the left tail of the distribution of X much higher than the right tail.

The functional $\rho[\cdot]$ defined in (2.11), with $r[\cdot] = \sigma_p[\cdot]$ and $p \geq 1$, is a coherent measure of risk, provided that $\lambda \in [0, 1]$. When $r[\cdot] = r_\beta[\cdot]$, the functional (2.11) is a coherent measure of risk, if $\lambda \in [0, 1/\beta]$. All these results can be found in Ruszczyński and Shapiro (2006a).

The mean-deviation from quantile $r_\beta[\cdot]$ is connected to the *Average Value at Risk* (AVaR), also known as *expected shortfall* or *Conditional Value at Risk* in Rockafellar and Uryasev (2000), as follows,

$$\mathrm{AVaR}_\beta(X) = -\max_{\eta\in\mathbb{R}} \left\{ \eta - \frac{1}{\beta}\mathbb{E}[(\eta - X)_+] \right\} = -\mathbb{E}[X] + \frac{1}{\beta} r_\beta[X]. \tag{2.14}$$

$$= \frac{1}{\beta}\int_0^\beta \mathrm{VaR}_p(X)\, dp. \tag{2.15}$$

Thus, AVaR is the special case of the mean-deviation from quantile when $\lambda = 1/\beta$. All these relations can be found in Ogryczak and Ruszczyński (2002), Ruszczyński and Vanderbei (2003) and Föllmer and Schied (2004) (with obvious adjustments for the sign change of X). The relation (2.14) allows me to interpret $\mathrm{AVaR}_\beta(X)$ as a mean–risk model. In the literature $\mathrm{AVaR}_\beta(X)$ is sometimes defined as the negative of (2.14) or in the right tail version. All these definitions are equivalent after appropriate sign adjustments.

One of the fundamental results in the theory of law-invariant measures is the theorem of Kusuoka (2003): *for every lower semicontinuous law-invariant coherent measure of risk* $\rho[\cdot]$ *on* $\mathscr{L}_\infty(\Omega, \mathscr{F}, P)$*, with an atomless probability space* $(\Omega, \mathscr{F}, P)$*, there exists a convex set* $\mathscr{M}$ *of probability measures on* $(0, 1]$ *such that*

$$\rho[X] = \sup_{\mu\in\mathscr{M}} \int_0^1 \mathrm{AVaR}_\beta[X]\, \mu(d\beta). \tag{2.16}$$

Using identity (2.14) I can rewrite $\rho[X]$ as follows:

$$\rho[X] = -\mathbb{E}[X] + \sup_{\mu \in \mathscr{M}} \int_0^1 \frac{1}{\beta} r_\beta[X]\, \mu(d\beta). \tag{2.17}$$

This means that every problem (2.11) with a coherent law-invariant measure of risk is a mean–risk model, with the risk functional

$$\varkappa_{\mathscr{M}}[X] = \sup_{\mu \in \mathscr{M}} \int_0^1 \frac{1}{\beta} r_\beta[X]\, \mu(d\beta). \tag{2.18}$$

and every nonatomic coherent measure of risk can be represented as a convex combination of other coherent measures of risk. In my dissertation, Kusuoka representation has a crucial role to derive analytical results under general coherent measures of risk.

The Kusuoka theorem also allows one to constructively define law-invariant coherent measures of risk, by specifying the set of measures $\mathscr{M}$. For example, setting $\mathscr{M}$ to the set of all probability measures on $[\beta_{\min}, \beta_{\max}]$, with $0 < \beta_{\min} \le \beta_{\max} < 1$, one obtains the risk functional as the worst scaled average deviation from quantile:

$$\varkappa_{\mathscr{M}}[Z] = \max_{\beta_{\min} \le \beta \le \beta_{\max}} \frac{1}{\beta} r_\beta[Z]. \tag{2.19}$$

2.2.6 Relationships between Different Risk Measures

The relation between stochastic dominance and utility functions is represented as follows:

$$X \succeq_{(1)} Y \Leftrightarrow \mathbb{E}[U(X)] \ge \mathbb{E}[U(Y)], \quad \text{for every nondecreasing } U[\cdot].$$

$$X \succeq_{(2)} Y \Leftrightarrow \mathbb{E}[U(X)] \ge \mathbb{E}[U(Y)], \quad \text{for every nondecreasing and concave } U[\cdot].$$

Then, the relation between stochastic dominance and coherent measures of risk is known as follows:

$$X \succeq_{(2)} Y \Rightarrow \rho(X) \le \rho(Y). \tag{2.20}$$

where $\rho(\cdot)$ is a law-invariant coherent measures of risk with non-atomic spaces $(\Omega, \mathscr{F}, \mathbb{P})$. It also implies that

$$X \succeq_{(1)} Y \Rightarrow \rho(X) \le \rho(Y). \tag{2.21}$$

because the second-order stochastic dominance is a weaker rule than the first-order stochastic dominance.

Lastly, some fundamental relations are represented among the concepts of Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR) and the stochastic dominance relations. Recall to §2.2.3. Then, by the definition of the first-order stochastic dominance relation, one obtains

$$X \succeq_{(1)} Y \Leftrightarrow \mathrm{VaR}_\alpha(X) \geq \mathrm{VaR}_\alpha(Y), \forall \alpha \in (0,1]. \tag{2.22}$$

The Conditional Value-at-Risk (CVaR), also called *Mean Shortfall or Tail VaR, at level* α, is defined in a simple way as follows (see Rockafellar and Uryasev (2000)):

$$\mathrm{CVaR}_\alpha(X) = \mathbb{E}[X|X \leq \mathrm{VaR}_\alpha(X)]. \tag{2.23}$$

The formula (2.23) is precise when $\mathrm{VaR}_\alpha(X)$ is not an atom of the distribution X.

Also, it is well known (see Ogryczak and Ruszczyński (2001) and Dentcheva and Ruszczyński (2006)) that the second-order stochastic dominance relation is equivalent to a continuum of CVaR constraints:

$$X \succeq_{(2)} Y \Leftrightarrow \mathrm{CVaR}_\alpha(X) \geq \mathrm{CVaR}_\alpha(Y), \quad \forall \alpha \in (0,1]. \tag{2.24}$$

2.3 Risk-Averse Inventory Models

In recent years, risk-averse inventory models have received increasing attention in the supply chain management literature. From my extensive literature review, I classify and summarize the literature of risk-averse inventory models at Table 2.4 by inventory models and risk measures used. Because there is no research so far directly applying stochastic dominance to this field, I drop it from the table.

In this 5-by-5 matrix, the rows represent risk measures used and the columns mean model types. Table 2.4 shows that most works to date focus on single-product inventory models for model types and utility function approach or mean-variance (or

	Single-product & Single-period	Sing-product & Multi-period	Multi-product & Single-period	Multi-echelon or Multi-agent
Utility Function	Lau (1980), Eeckhoudt *et al.* (1995), Agrawal & Seshadri (2000a), Gaur & Seshadri (2005)	Bouakiz & Sobel (1992), Chen *et al.* (2007)	van Mieghem (2007)	Agrawal & Seshadri (2000b), van Mieghem (2003), Gan *et al.* (2004)
Mean-Deviation from Quantile or Mean-Semideviation	Gotoh & Takano (2007)	None	Agrali & Soylu (2006), Tomlin & Wang (2005)	None
General Coherent Measures of Risk	Ahmed *et al.* (2007), Choi & Ruszczyński (2008)	Ahmed *et al.* (2007)	Choi, Ruszczyński and Zhao (2009)	None
Mean-Variance (or Mean-Standard Deviation)	Anvari (1987), Chung (1990), Chen & Federgruen (2000), Gaur & Seshadri (2005), Martinez-de-Albeniz & Simchi-Levi (2006)	None	van Mieghem (2007)	Lau & Lau (1999), Tsay (2002), van Mieghem (2003), Gan *et al.* (2004)
Chance Constraints (or Value-at-Risk)	Lau (1980), Özler *et al.* (2009)	None	Özler *et al.* (2009)	Gan *et al.* (2005)

Table 2.4: Summary of Literature on Risk-Averse Inventory Models.

mean-standard deviation) models for risk measures. Recently some studies use mean-deviation from quantile (including Conditional Value-at-Risk) or mean-semideviation, which are special cases of general coherent measures of risk. Then, on top of these studies, Ahmed *et al.* (2007), Choi and Ruszczyński (2008) and Choi, Ruszczyński and Zhao (2009) develop their inventory models further under general coherent measures of risk.

For newsvendor models, most research focuses on finding the optimal solution under a risk-averse performance measure, and studying the impact of the degree of risk aversion (and other model parameters) on the optimal solution. A typical finding is that as the degree of risk aversion increases, the optimal order quantity tends to decrease.

For single-product but multi-period dynamic inventory models under risk aversion, the literature focuses on characterizing the structure of the optimal ordering or pricing policies and quantifying the impact of the degree of risk aversion on the optimal polices. For the risk measures used, Chen, Sim, Simchi-Levi and Sun (2007) provide an excellent review and a summary of results for this literature.

For multi-product risk-averse newsvendor models, Tomlin and Wang (2005) study how characteristics of products (e.g., profit margin, demand correlation), resource reliability and firm's risk attitude affect the preference of resource flexibility and supply diversification. Under a downside risk measure and Conditional Value at Risk (CVaR), they show that for a risk-averse firm with unreliable resources, a supply chain can prefer dedicated resources than a flexible resource even if the cost of the latter is smaller than the former. While it seems counter-intuitive, it is possible because if the flexible resource fails then all products are negatively affected. However, in the dedicated resource case, for all products to be negatively affected, all resources have to fail.

Van Mieghem (2007) studies three newsvendor networks with many products and many resources under mean-variance and utility function approaches. These networks feature resource diversification, flexibility (e.g., ex post inventory capacity allocation) and/or demand pooling. The paper addresses the question of how the aforementioned operational strategies reduce total risk and create value. It shows that a risk-averse newsvendor may invest more resources in certain networks than a risk-neutral newsvendor (i.e., operational hedging) because such resources may reduce the profit variance and mitigate risk in the network. Among the three networks, the dedicated one is mostly related to our model. In this network, there are two products with correlated demand. The author characterizes the impact of demand correlation on the optimal order quantities in two extreme cases of complete positive or negative correlation. A numerical study is conducted to cover cases other than the extreme ones.

Özler *et al.* (2009) consider a multi-product newsvendor with a Value-at-Risk constraint. For a single-product system, they obtain the closed-form optimal ordering quantity which is the same result of Gan *et al.* (2004). Their biggest contribution to the literature is that for a two-product system, they obtain the mathematical formulation of mixed integer programming where the objective function is nonlinear and the constraints are mixed linear and nonlinear functions. Then, they conducted their numerical analysis to confirm their analytical results under multi-variate exponential demands.

Finally, Agrali and Soylu (2006) conduct a numerical investigation on a two-product newsvendor model under the risk measure of CVaR. Assuming a discretized multivariate normal demand distribution, the authors studied the sensitivity of the optimal solution with respect to the mean and variance of demand, demand correlation, and various cost parameters. Interestingly, the report shows that as the demand correlation increases, the optimal order quantities tend to decrease.

For multi-echelon or multi-agent models, so far all papers consider single-product and single-period models. Lau and Lau (1999) study a manufacturer's pricing strategy and return policy under the mean-variance risk measure. Agrawal and Seshadri (2000b) introduce a risk-neutral intermediaries to offer mutually beneficial contracts to risk-averse retailers. Tsay (2002) studies how a manufacturer can use return policies to share risk under the mean-standard deviation measure. Gan, Sethi and Yan (2004) study Pareto-optimality for suppliers and retailers under various risk-averse measures. Gan, Sethi and Yan (2005) design coordination schemes of buyback and risk-sharing contracts in a supply chain under a Value-at-Risk constraint. Van Mieghem (2003) provides an excellent review on the literature that incorporates risk aversion in capacity investment models and reduces risk via operational hedging.

2.4 Validity of Risk Measures in Newsvendor Models

Although the typical approaches to risk-averse inventory models are closely related and consistent to some extent, they are different from each other. There are advantages and disadvantages of each particular risk-averse inventory model. In this section, I compare various risk measures by two criteria in newsvendor models. The first criterion is consistency to stochastic dominance and the second one is consistency to the four axioms in coherent measures of risk. My axiomatic approach provides a clear standard to evaluate risk measures in risk-averse newsvendor models. This axiomatic approach can be also applicable to other types of risk-averse inventory models after appropriate adjustments.

	Consistency to Stochastic Dominance	Consistency to the Four Axioms for Coherent Risk Measures
Utility Function	First-order for nondecreasing utility function Second-order for nondecreasing and concave utility function	Convexity & Monotonicity
Mean-Variance	Not consistent	Translation Equivariance
Mean-Standard Deviation	Not consistent	Translation Equivariance & Positive Homogeneity
Chance Constraints (Value-at-Risk)	Relaxed version of First-order Could violate Second-order	Monotonicity & Translation Equivariance & Positive Homogeneity
General Coherent Measures of Risk	First and Second order	All of the four axioms
Convex Measures of Risk	Same as Coherent Measures of Risk	Convexity & Monotonicity & Translation Equivariance
Insurance Risk Measures	Same as Coherent Measures of Risk	Monotonicity & Positive Homogeneity & Translation Equivariance
Natural Risk Statistic	Same as Coherent Measures of Risk	Same as Insurance Risk Measures
Tradeable Measures of Risk	Same as Coherent Measures of Risk	Same as Insurance Risk Measures

Table 2.5: Comparison between Various Risk Measures.

Utility functions satisfy Convexity and Monotonicity when they are nondecreasing and concave. Nevertheless, they do not satisfy Translation Equivariance and Positive Homogeneity. Although these two axioms are not always desirable in supply chain applications (for instance, when initial endowment plays a significant role in risk preferences), they can capture risk preferences better in newsvendor models. This is true because in the newsvendor model, I am mainly concerned about the risk of random demand and the associated overage/underage costs. Thus, I use exponential utility function approach as a reference model to general coherent measures of risk. The work of multi-product newsvendors with exponential utility function approach is conducted in Chapter 5.

Mean-variance and mean-standard deviation model have been very well-known since Markowitz (1952). The mean-variance model satisfies the Translation Equivariance axiom only. Mean-standard deviation model satisfies Translation Equivariance and Positive Homogeneity, but not Convexity and Monotonicity. More importantly, these models are not consistent with any of stochastic dominance relations because they treat equally over-performance and under-performance.

Chance constraints and Value-at-Risk have been actively used in finance historically. In financial terms, they are easy to understand and intuitive. However, they generally violate Convexity, which implies that Value-at-Risk may penalize diversification instead of encouraging it. This situation may be justified in finance such as insurance company but very different from that in supply chain management.

General coherent measures of risk are consistent to the first- and second-order stochastic dominance relations and satisfy all the four axioms. Thus, they can arguably be the best risk measure for the single- and multi-product newsvendors by my axiomatic approach. That is the reason the work on single- and multi-product newsvendors under coherent measures of risk is at the heart of my dissertation in Chapter 3 and 4. However, these four axioms might be too strong in several financial markets such as insurance companies. More specifically, a very popular VaR (Value-at-Risk) is not a coherent measure of risk because it does not satisfy the Convexity axiom in general. Due to the popularity of VaR in financial markets, the Convexity has been a long controversial axiom in finance literature.

By this reason, several modifications and extensions of coherent measures of risk have been studied actively in finance literature to hold different subsets of the four axioms. They are *insurance risk measures, convex measures of risk, natural risk statistic* and *tradeable measures of risk*. Thus, in general, these four risk measures do not satisfy at least one of the four axioms in coherent measures of risk nor can avoid the potential problems from violating them in newsvendor models.

Föllmer and Schied (2002) consider convex measures of risk, in which the Positive

Homogeneity axiom is relaxed. Again, in our context, this may lead to a diversification effect when demands are completely correlated; it may also lead to counter-intuitive effects of changing my attitude to risk when the outcomes are re-scaled, by changing the currency in which profits are calculated, or by considering the average profit per product.

The other three risk measures do not satisfy the Convexity axiom in general. They are based on the reality of financial markets where non-coherent risk measures, such as VaR (Value-at-Risk), are still widely used. Insurance risk measures, initiated by Wang, Young and Panjer (1997), satisfy conditional state independence, monotonicity, comonotone additivity and continuity and this risk measure is law-invariant. Wang *et al.* (1997) set the comonotone additivity axiom based on their argument that comonotone portfolios cannot hedge each other. Heyde, Kou and Peng (2006) propose a natural risk statistic, which is law-invariant, and in which the convexity axiom is required only for comonotone random variables. Ahmed, Filipovic and Svidland (2008) show that such a risk measure can be represented as a composition of a coherent measure of risk and a certain law-preserving transformation, and thus our insights into models with coherent measures of risk are relevant for natural risk statistics. Pospišil, Večer and Xu (2008) propose tradeable measures of risk. They argue that the proper risk measures should be constructed by historically realized returns. However, comparing to the coherent measures of risk, these risk measures appears to be much more difficult to handle, due to non-convexity and/or nondifferentiability of the resulting model. We shall see that even in the case of coherent measures of risk, the technical difficulties are substantial.

Chapter 3
Single-Product Newsvendor Models

The work in this chapter was based on ”A risk-averse newsvendor with law-invariant coherent measures of risk” by S. Choi and A. Ruszczyński (2008) published at *Operations Research Letters*, **36**, 77-82 and partially complemented by another working paper, ”A multi-product risk-averse newsvendor with law-invariant coherent measures of risk” by S. Choi, A. Ruszczyński and Y. Zhao (2009).

In this chapter, I study a risk-averse newsvendor problem with coherent measures of risk. I first formulate this problem as a mean–risk model and consider risk-averse newsvendor solutions with mean-deviation from quantile and mean-semideviation. Then I study the impact of risk aversion on optimal ordering quantity with both risk measures and obtain a closed-form optimal solution under mean-deviation from quantile. Finally I generalize my analysis to every law-invariant coherent measure of risk and obtain a closed-form optimal solution.

By using the standard definition of the newsvendor problem and employing modern theory of law-invariant coherent measures of risk, I arrive that the more risk-averse the newsvendor is, the smaller his order is. This is in harmony with my intuition and with the results of Eeckhoudt *et al.* (1995), where a utility model was considered, and with Gotoh and Takano (2007), for a Conditional Value-at-Risk model.

3.1 Problem Formulation

Following from the typical format of the newsvendor model, I introduce the following parameters: unit resale price r, unit ordering cost c, and unit salvage value s. Backordering is not allowed in this model. I assume that $r > c > s \geq 0$ in order to avoid trivial cases. Let x denote ordering amounts by the newsvendor and let D be random demand. The net profit can be calculated as follows:

$$\Pi(x, D) = -cx + r\min(D, x) + s\max(0, x - D).$$

Simple manipulation yields an equivalent formula:

$$\begin{aligned}\Pi(x, D) &= -cx + s\min(D, x) + s\max(0, x - D) + (r - s)\min(D, x)\\ &= -cx + sx + (r - s)\min(D, x). = -(c - s)x + (r - s)\min(D, x).\end{aligned}$$

It is thus sufficient to consider the newsvendor problem with adjusted resale price $\bar{r} = r - s$ and adjusted ordering cost $\bar{c} = c - s > 0$, and with no salvage value. Here, $\bar{r} > \bar{c} > 0$. The objective function can be written simply as

$$\Pi(x, D) = -\bar{c}x + \bar{r}\min(x, D). \tag{3.1}$$

Thus, $\Pi(x; D)$ is piecewise linear and concave in x with a refraction point $x =$ D as follows:

$$\Pi(x, D) = \begin{cases} (\bar{r} - \bar{c})x, & \text{if } x \leq D. \\ -\bar{c}x + \bar{r}D, & \text{if } x > D. \end{cases}$$

As the function $\Pi(\cdot, D)$ is piecewise linear and concave in x for all given $D \in \mathbb{R}$, the expected newsvendor's profit, $\mathbb{E}[\Pi(x; D)]$ is piecewise and concave as well.

Similarly, the function $\Pi(\cdot; x)$ is piecewise linear and concave in D for all given $x \in \mathbb{R}$, the expected newsvendor's profit, $\mathbb{E}[\Pi(D; x)]$ is piecewise and concave as well.

The *risk-neutral* newsvendor's problem

$$\max_{x \geq 0} \; \mathbb{E}[\Pi(x, D)]. \tag{3.2}$$

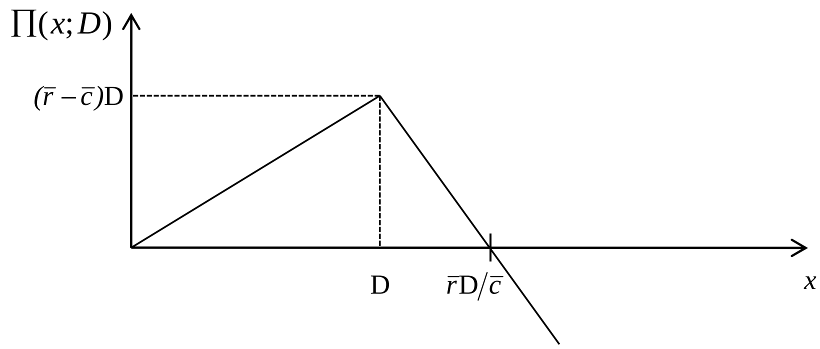

Figure 3.1: Profit, a Function of Ordering Quantity with Given Demand D

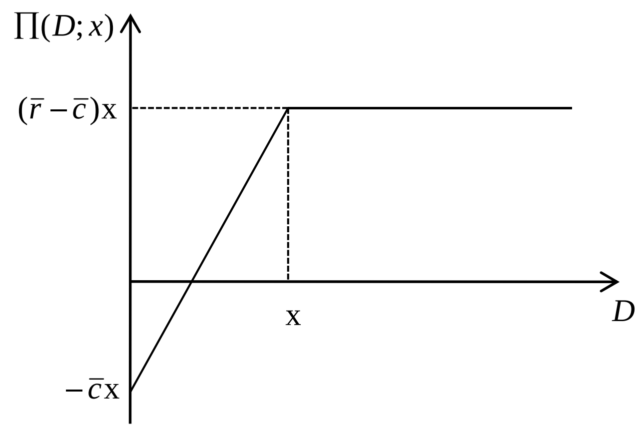

Figure 3.2: Profit, a Function of Random Demand with Given Ordering Quantity x

has the well-known solution

$$\hat{x}^{\text{RN}} \in [q_\alpha^-(D), q_\alpha^+(D)), \quad \text{with } \alpha = (\bar{r}-\bar{c})/\bar{r}, \tag{3.3}$$

where

$$q_\alpha^-(D) = \inf\{\eta : \mathbb{P}[D \le \eta] \ge \alpha\} : \text{ the left } \alpha\text{-quantile of } D, \tag{3.4}$$

$$q_\alpha^+(D) = \sup\{\eta : \mathbb{P}[D < \eta] \le \alpha\} : \text{ the right } \alpha\text{-quantile of } D. \tag{3.5}$$

When I assume a nonatomic probability space of the product demand, it implies that the demand distribution is continuous and the risk-neutral solution given at the equation (3.3) is equivalent as follows:

$$\hat{x}^{\text{RN}} = \bar{F}_D(\alpha), \text{ with } \alpha = (c-s)/(r-s). \tag{3.6}$$

where $\bar{F}_D(\cdot) = 1 - F_D(\cdot)$ and $F_D(\cdot)$ is a cumulative distribution function of demand D.

In this chapter, I aim to replace the new risk-averse objective function with coherent measures of risk. Then, I recall the objective function shown at the equation (2.11) and this is same as follows:

$$\min_{x \geq 0} \rho(X) = -\mathbb{E}[X] + \lambda r[X]. \tag{3.7}$$

In the next two sections, I will consider the two popular and special cases of coherent measures of risk, mean-deviation from quantile and mean-semideviation. Lastly, I will consider general coherent measures of risk.

3.2 Mean-Deviation from Quantile Model

In this section, I study the impact of the degree of risk aversion on the optimal ordering quantity and obtain the closed-form optimal solution for a mean-deviation from quantile model.

3.2.1 Impact of Degree of Risk Aversion

Lemma 1. *For every $\beta \in (0, 1)$ the function $x \mapsto r_\beta[\Pi(x)]$ is nondecreasing on $\mathbb{R}_+$.*

Proof. I use here the idea based on simple inequalities for quantiles. This proof idea does not require an assumption of continuous demand distribution. The result can be proved (for a continuous distribution of D) by differentiating $r_\beta[\Pi(x, D)]$ with respect to x and changing the order of integration (see Ahmed *et al.* (2007) and Gotoh and Takano (2007) for similar derivations).

Observe that $r_\beta[\Pi(x)] = r_\beta[\Pi(x) + \varphi(x)]$ for any deterministic function $\varphi : \mathbb{R} \to \mathbb{R}$. Also, the function $r_\beta[\cdot]$ is positively homogeneous. In particular, using $\varphi(x) = cx$, I see from (3.1) that $r_\beta[\Pi(x)] = \overline{r} \cdot r_\beta[\min(x, D)]$. It remains to prove the assertion for the function $x \mapsto r_\beta[\min(x, D)]$.

Consider $0 \leq x_1 \leq x_2$ and define the random variables $Z_1 = \min(x_1, D)$ and $Z_2 = \min(x_2, D)$. Let $q_\beta(D)$ be a fixed β-quantile of D.

Suppose $x_1 \leq q_\beta(D) \leq x_2$. Then $q_\beta(Z_1) = x_1$ and $q_\beta(Z_2) = q_\beta(D)$. By (2.13) with $\eta = q_\beta(D)$,

$$r[Z_2] = \mathbb{E}[\max(\beta(Z_2 - q_\beta(D)), (1-\beta)(q_\beta(D) - Z_2))]. \tag{3.8}$$

If $D < x_1$ then $D < q_\beta(D) \leq x_2$ and $Z_2 = D$. The expression under the expected value in (3.8) reads

$$\max(\beta(Z_2 - q_\beta(D)), (1-\beta)(q_\beta(D) - Z_2)) = (1-\beta)(q_\beta(D) - D) \geq (1-\beta)(x_1 - D).$$

If $D \geq x_1$ the "max" expression under the expected value in (3.8) is nonnegative. Therefore

$$r[Z_2] \geq (1-\beta)\mathbb{E}[(x_1 - D)_+] = r[Z_1].$$

as required. The other two cases ($q_\beta(D) < x_1$ and $q_\beta(D) > x_2$) can be analyzed in a similar way. □

Proposition 1. *Assume that* $0 \leq \lambda_1 \leq \lambda_2$. *Then for every solution* $\hat{x}^{\mathrm{RA}_1}$ *of the problem*

$$\min_{x \geq 0} -\mathbb{E}[\Pi(x, D)] + \lambda_1 r_\beta[\Pi(x, D)]. \tag{3.9}$$

there exists a solution $\hat{x}^{\mathrm{RA}_2}$ *of the problem*

$$\min_{x \geq 0} -\mathbb{E}[\Pi(x, D)] + \lambda_2 r_\beta[\Pi(x, D)]. \tag{3.10}$$

such that $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1}$. *Conversely, for every solution* $\hat{x}^{\mathrm{RA}_2}$ *of problem* (3.10) *there exists a solution* $\hat{x}^{\mathrm{RA}_1}$ *of problem* (3.9) *such that* $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1}$.

Proof. Suppose that $\hat{x}^{\mathrm{RA}_1}$ is an optimal solution of (3.9). Observe that the objective function of (3.10) differs from the objective function of (3.9) by $(\lambda_2 - \lambda_1) r_\beta[\Pi(x, D)]$ which, by Lemma 1, is a nondecreasing function of x. Thus, for every $x \geq \hat{x}^{\mathrm{RA}_1}$ I have the inequality

$$\begin{aligned}
&- \mathbb{E}[\Pi(x, D)] + \lambda_2 r_\beta[\Pi(x, D)] \\
=& - \mathbb{E}[\Pi(x, D)] + \lambda_1 r_\beta[\Pi(x, D)] + (\lambda_2 - \lambda_1) r_\beta[\Pi(x, D)]. \\
\geq& - \mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}_1}, D)] + \lambda_1 r_\beta[\Pi(\hat{x}^{\mathrm{RA}_1}, D)] + (\lambda_2 - \lambda_1) r_\beta[\Pi(\hat{x}^{\mathrm{RA}_1}, D)]. \\
=& - \mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}_1}, D)] + \lambda_2 r_\beta[\Pi(\hat{x}^{\mathrm{RA}_1}, D)].
\end{aligned}$$

The reverse implication is similar. □

3.2.2 Closed-Form Optimal Solution

Now I assume a continuous demand distribution, D. Then, I study mean-deviation from quantile model by differentiating $r_\beta[\Pi]$. In this subsection, I add a notation $Z_x^1 = \bar{r}\min(x, D)$. Then, let $f_D(\cdot)$ and $F_D(\cdot)$ be the probability density function (pdf) and the cumulative distribution function (cdf) of D, respectively. Denote $\bar{F}_D(\xi) = 1 - F_D(\xi)$, $\forall \xi \in \mathbb{R}$. Then,

$$\Pi(x, D) = -\bar{c}x + Z_x^1.$$

Thus,

$$\begin{aligned}
\rho[\Pi(x, D)] &= \bar{c}x + (-\mathbb{E}[Z_x^1] + \lambda r_\beta(Z_x^1)). \\
&= \bar{c}x + \Big(\mathbb{E}[Z_x^1](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}} \Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x^1)_+]\Big\}\Big).
\end{aligned} \tag{3.11}$$

Here $\hat{\eta}$ be the maximizer in (3.11), among $\eta \in \mathbb{R}$, at a fixed x. Then $\hat{\eta}$ is the β-quantile of Z_x^1. To take the derivative of $\rho[\Pi(x, D)]$ with respect to x, I consider two cases.

Case (i): $\hat{\eta} < \bar{r}x$.

Assuming that the quantile $\hat{\eta}$ is unique and differentiating the equation (3.11), I observe that

$$\frac{d\rho[\Pi(x, D)]}{dx} = \bar{c} + \bar{r}(\lambda\beta - 1)P[D > x] - \bar{r}\lambda\mathbb{P}[\{Z_x^1 < \hat{\eta}\} \cap \{D > x\}]. \tag{3.12}$$

Observe that in Case (i)

$$P[\{Z_x^1 < \hat{\eta}\} \cap \{D > x\}] = P[Z_x^1 < \hat{\eta} | D > x] P[D > x] = 0.$$

Therefore,

$$\frac{d\rho[\bar{\Pi}(x,D)]}{dx} = \bar{c} + \bar{r}(\lambda\beta - 1)P[D > x].$$

This yields the exact solution of the single product problem

$$\hat{x}^{\mathrm{RA}} = \bar{F}_D^{-1}\left(\frac{\bar{c}}{\bar{r}(1-\lambda\beta)}\right) \le \bar{F}_D^{-1}\left(\frac{\bar{c}}{\bar{r}}\right) = \hat{x}^{\mathrm{RN}}.$$

This special case solution is the same as the result of Gotoh and Takano (2007).

Case (ii): $\hat{\eta} = \bar{r}x$.

I have

$$\rho[\Pi(x,D)] = \bar{c}x + \Big(\mathbb{E}[Z_x^1](\lambda\beta - 1) - \lambda\beta\Big\{\bar{r}x - \frac{1}{\beta}\mathbb{E}[\bar{r}x - Z_x^1]\Big\}\Big).$$

Taking derivative with respect to x yields,

$$\begin{aligned} \frac{d\rho[\Pi(x,D)]}{dx} &= \bar{c} + \bar{r}(\lambda\beta - 1)P[D > x] - \bar{r}\lambda\beta\Big\{1 - \frac{1}{\beta}(1 - P[D > x])\Big\}. \\ &= \bar{c} + \bar{r}\lambda(1-\beta) + \bar{r}P[D > x](-1 + \lambda(\beta - 1)). \end{aligned} \tag{3.13}$$

Equating the right hand side to 0, I obtain an exact solution as follow:

$$\hat{x}^{\mathrm{RA}} = \bar{F}_D^{-1}\left(\frac{\bar{c} + \bar{r}\lambda(1-\beta)}{\bar{r}(1 + \lambda(1-\beta))}\right). \tag{3.14}$$

Clearly, if $\lambda = 0$, $\hat{x}^{\mathrm{RA}} = \hat{x}^{\mathrm{RN}}$ in both of case *(i)* and *(ii)*. As λ increases, $\hat{x}^{\mathrm{RA}}$ is decreasing, which confirms the Proposition 1. For any $0 \le \lambda \le 1/\beta$, $\hat{x}^{\mathrm{RA}}$ is well-defined. To determine whether Case (i) or Case (ii) applies, one can compute $\hat{x}^{\mathrm{RA}}$ for both cases, and then compute $\hat{\eta}$ to check the case conditions.

3.3 Mean-Semideviation Model

The property that the solution set of the problem (3.7) is a nonincreasing function of the risk aversion parameter λ can now be extended to the case of the semideviation risk functional. One way to do it is a direct proof, as Lemma 1 and Proposition 1 above.

I focus on the mean-semideviation with degree one and then choose another approach that uses the general identity proved by Ogryczak and Ruszczyński (2002):

$$\sigma_1[\Pi] = \max_{0<\beta<1} r_\beta[\Pi]. \tag{3.15}$$

Lemma 2. *For every $\beta \in (0, 1)$ the function $x \mapsto \sigma_1[\Pi(x)]$ is nondecreasing on $\mathbb{R}_+$.*

Proof. Owing to Lemma 1, for every $\beta \in (0, 1)$ the function $x \mapsto r_\beta[Z(x)]$ is nondecreasing. Consequently, their maximum is nondecreasing as well, and my result follows from the identity (3.15). □

It is now clear that Proposition 1 also holds true for the mean–risk model with the risk functional $r[\cdot] = \sigma_1[\cdot]$. The proof is identical. In addition, I confirm Lemma 2 alternatively by taking derivatives (for a continuous demand distribution) or subgradient (not necessarily differentiable at all points) in the following subsections. One needs to note that mean-semideviation is a special case of general law-invariant coherent measures of risk. Thus, similar analysis made in section 3.2 can be done for a mean-semideviation model. However, in a mean-semideviation model, it is very complicated to establish a convex set $\mathscr{M}$ of probability measures from the equation (2.16) through (2.19) and thus it is very difficult to obtain a closed-form optimal solution for a mean-semideviation model.

3.3.1 Impact of Degree of Risk Aversion with Uniform Demand Distribution

First, as the special case of continuous demand distribution, I assume that the demand has a uniform probability distribution where the minimum (or maximum) value of the demand is $d_{\min}$ (or $d_{\max}$) with $0 \le d_{\min} \le d_{\max}$. Then, the profit function shown from the equation (3.1) can be represented equivalently as follows:

$$\Pi(x,D)=\begin{cases}-\bar{c}x+\bar{r}D, & \text{if } d_{\min}\le D<x.\\ (\bar{r}-\bar{c})x, & \text{if } x\le D\le d_{\max}.\end{cases}$$

Then,

$$\begin{aligned}\mathbb{E}[\Pi(x,D)] &= (\bar{r}-\bar{c})x\left(\frac{d_{\max}-x}{d_{\max}-d_{\min}}\right)+\int_{d_{\min}}^{x}\left(\frac{\bar{r}D-\bar{c}x}{d_{\max}-d_{\min}}\right)dD.\\ &= (\bar{r}-\bar{c})x\left(\frac{d_{\max}-x}{d_{\max}-d_{\min}}\right)-\frac{\bar{c}x(x-d_{\min})}{d_{\max}-d_{\min}}+\frac{\bar{r}(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}.\end{aligned}$$

Thus,

$$\mathbb{E}[\Pi(x,D)]-\Pi(x,D)=\begin{cases}\bar{r}x\left(\frac{d_{\max}-x}{d_{\max}-d_{\min}}\right)+\frac{\bar{r}(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}-\bar{r}D, & \text{if } d_{\min}\le D<x.\\ -\bar{r}x\left(\frac{x-d_{\min}}{d_{\max}-d_{\min}}\right)+\frac{\bar{r}(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}, & \text{if } x\le D\le d_{\max}.\end{cases}$$

Here, when $x\le D\le d_{\max}$,

$$\begin{aligned}\mathbb{E}[\Pi(x,D)]-\Pi(x,D) &= -\bar{r}x\left(\frac{x-d_{\min}}{d_{\max}-d_{\min}}\right)+\frac{\bar{r}(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}.\\ &= \frac{\bar{r}}{2(d_{\max}-d_{\min})}\Big[-2x(x-d_{\min})+(x-d_{\min})(x+d_{\min})\Big].\\ &= -\frac{\bar{r}(x-d_{\min})^2}{2(d_{\max}-d_{\min})}<0.\end{aligned}$$

Thus, $(\mathbb{E}[\Pi(x,D)]-\Pi(x,D))_+$ is positive only if $d_{\min}\le D<\frac{x(d_{\max}-x)}{(d_{\max}-d_{\min})}+\frac{(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}$. Otherwise, it is zero. When I denote $\mathbb{E}[(\mathbb{E}[\Pi(x,D)]-\Pi(x,D))_+]=\sigma_1(x)$ and $a(x)=\frac{x(d_{\max}-x)}{(d_{\max}-d_{\min})}+\frac{(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}$,

$$\begin{aligned}\frac{d\sigma_1(x)}{dx} &= \frac{d}{dx}\int_{d_{\min}}^{a(x)}\left(\bar{r}x\left(\frac{d_{\max}-x}{d_{\max}-d_{\min}}\right)+\frac{\bar{r}(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}-\bar{r}D\right)\left(\frac{1}{d_{\max}-d_{\min}}\right)dD.\\ &= \frac{\bar{r}}{(d_{\max}-d_{\min})^2}\int_{d_{\min}}^{a(x)}(d_{\max}-x)dD, \quad \text{by the Leibnitz Integral Rule.}\\ &= \frac{\bar{r}(d_{\max}-x)}{(d_{\max}-d_{\min})^2}\left(\frac{x(d_{\max}-x)}{d_{\max}-d_{\min}}+\frac{(x-d_{\min})(x+d_{\min})}{2(d_{\max}-d_{\min})}-d_{\min}\right).\\ &= \frac{\bar{r}(d_{\max}-x)}{2(d_{\max}-d_{\min})^3}\Big(2x(d_{\max}-x)+(x-d_{\min})(x+d_{\min})-2d_{\min}(d_{\max}-d_{\min})\Big).\end{aligned}$$

Thus, $(d_{\max}-x)\ge 0$ and the remaining is to prove:

$$2x(d_{\max}-x)+(x-d_{\min})(x+d_{\min})\ge 2d_{\min}(d_{\max}-d_{\min}), \quad \text{where } d_{\min}\le x\le d_{\max}. \tag{3.16}$$

Now, the left-hand side of the equation (3.16) is equal to

$$-x^2 + 2xd_{\max} - d_{\min}^2 = -(x - d_{\max})^2 + d_{\max}^2 - d_{\min}^2. \tag{3.17}$$

Thus, the left-hand side of the equation (3.17) is an increasing function with $d_{\min} \leq x \leq d_{\max}$ and the left-hand side of the equation (3.17) is equal to the right-hand side of the equation (3.17) when $x = d_{\min}$. Thus, lemma 2 is confirmed again when product demand has a uniform distribution function.

3.3.2 Impact of Degree of Risk Aversion with (Arbitrarily) Continuous Demand Distribution

Next I relax the assumption of uniform demand distribution and then the product demand has an arbitrarily continuous distribution function. Again I define the objective function, $\rho(\Pi(x, D))$ as follows:

$$\begin{aligned}\min_{x\geq 0} \rho(\Pi(x, D)) &= -\mathbb{E}[\Pi(x, D)] + \lambda\sigma_1(\Pi(x, D)).\\ &= -\mathbb{E}[\Pi(x, D)] + \lambda\mathbb{E}[\mathbb{E}[\Pi(x, D)] - \Pi(x, D)]_+.\end{aligned}$$

Then,

$$\frac{\partial\rho(\Pi(x, D))}{\partial x} = \mathbb{E}\Big[\frac{\partial\rho(\Pi)}{\partial\Pi} \cdot \frac{d\Pi(x, D)}{dx}\Big].$$

where

$$\frac{\partial\rho(\Pi)}{\partial\Pi} = \frac{\partial(-\mathbb{E}(\Pi) + \lambda\sigma_1(\Pi))}{\partial\Pi} = -1 + \lambda G(\omega).$$

and

$$G(\omega) = \frac{\partial\mathbb{E}[(\mathbb{E}(\Pi) - \Pi)_+]}{\partial\Pi} = -\mathbb{1}_{\mathbb{E}(\Pi)>\Pi} + \mathbb{P}(\mathbb{E}(\Pi) > \Pi).$$

then

$$\Pi_x' = \frac{d\Pi(x, D)}{dx} = \begin{cases} -\bar{c}, & \text{if } D < x. \\ (\bar{r} - \bar{c}), & \text{if } D \geq x. \end{cases}$$

Therefore,

$$\frac{\partial\rho(x, D)}{\partial x} = \mathbb{E}[\Pi_x'(-1 + \lambda G(\omega))] = -\mathbb{E}[\Pi_x'] + \lambda\mathbb{E}[\Pi_x' \cdot G(\omega)].$$

In order to prove that $\hat{x}^{\mathrm{RA}} \leq \hat{x}^{\mathrm{RN}}$, it is enough to prove that $\frac{\partial \rho}{\partial x}(\hat{x}^{\mathrm{RN}}) \geq 0$ and also equivalent to prove that $\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot G(\omega)]$ because $\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}})] = 0$. Thus,

$$\begin{aligned}
\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot G(\omega)] &= -\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot \mathbb{1}_{\mathbb{E}(\Pi)>\Pi}] + \mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot \mathbb{P}(\mathbb{E}(\Pi) > \Pi)]. \\
&= -\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot \mathbb{1}_{\mathbb{E}(\Pi)>\Pi}] + \mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}})] \cdot \mathbb{P}(\mathbb{E}(\Pi) > \Pi). \\
&= -\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) \cdot \mathbb{1}_{\mathbb{E}(\Pi)>\Pi}]. \\
&= -\mathbb{P}(\mathbb{E}(\Pi) > \Pi) \cdot \mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) | \mathbb{E}(\Pi) > \Pi].
\end{aligned}$$

Then, $\Pi_x^{'}(\hat{x}^{\mathrm{RN}}) < 0$ if $\mathbb{E}(\Pi) > \Pi$ because $\mathbb{E}(\Pi) > \Pi$ implies that $D < x$. Therefore, $\frac{\partial \rho}{\partial x}(\hat{x}^{\mathrm{RN}}) > 0$ and $\hat{x}^{\mathrm{RA}} \leq \hat{x}^{\mathrm{RN}}$.

Then, the next step is to prove that $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1}$ when $\hat{x}^{\mathrm{RA}_1}$ and $\hat{x}^{\mathrm{RA}_2}$ are the optimal solutions for λ_1 and λ_2, respectively, with $0 \leq \lambda_1 \leq \lambda_2$. Then, I also denote the objective function similarly as ρ_1 and ρ_2, respectively. Similarly to the previous proof, it is enough to show that $\frac{\partial \rho_2}{\partial x}(\hat{x}^{\mathrm{RA}_1}) \geq 0$. Then,

$$\begin{aligned}
\frac{\partial \rho_2}{\partial x}(\hat{x}^{\mathrm{RA}_1}) &= -\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RA}_1})] + \lambda_2 \mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}). \\
&= -\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RA}_1})] + \lambda_1 \mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}) + (\lambda_2 - \lambda_1)\mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}). \\
&= (\lambda_2 - \lambda_1)\mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}).
\end{aligned}$$

Then, the remaining is to prove that $\mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}) \geq 0$. Thus,

$$\mathbb{E}[\Pi_x^{'} \cdot G(\omega)](\hat{x}^{\mathrm{RA}_1}) = \mathbb{E}[\Pi_x^{'} \cdot (-\mathbb{1}_{\mathbb{E}(\Pi)>\Pi} + \mathbb{P}(\mathbb{E}(\Pi) > \Pi))](\hat{x}^{\mathrm{RA}_1}). \tag{3.18}$$

$$= -\mathbb{P}(\mathbb{E}(\Pi) > \Pi) \cdot \mathbb{E}[\Pi_x^{'} | \mathbb{E}(\Pi) > \Pi](\hat{x}^{\mathrm{RA}_1}) + \mathbb{P}(\mathbb{E}(\Pi) > \Pi)\mathbb{E}[\Pi_x^{'}](\hat{x}^{\mathrm{RA}_1}). \tag{3.19}$$

Then, the first term of (3.19) is positive because $\Pi_x^{'}(\hat{x}^{\mathrm{RA}_1}) < 0$ when $\mathbb{E}(\Pi) > \Pi$. Also, the second term of (3.19) is also positive because $\hat{x}^{\mathrm{RA}_1} \leq \hat{x}^{\mathrm{RN}}$. Therefore, $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1} \leq \hat{x}^{\mathrm{RN}}$.

3.3.3 Impact of Degree of Risk Aversion with (Arbitrarily) General Demand Distribution

Finally I assume more general case - the risk-adjusted performance measure has only convexity and may not be differentiable in all domain of Π. Again,

$$\min_{x\geq 0} \rho(\Pi(x, D)) = -\mathbb{E}[\Pi(x, D)] + \lambda\sigma_1(\Pi(x, D)).$$

Here, $\rho(\Pi) = \inf_{\mu\in A} \int \Pi(\omega)\mu(d\omega)$, where μ is a measure on $(\mu, \mathscr{F})$. If $\Pi \in L = L_p$ which implies that the existence of p^{th} moments with $p = 1$ is equivalent to $\mathbb{E}[\Pi]$ exists and then $\mu(d\omega) = g(\omega) \cdot \mathbb{P}(d\omega)$. Thus,

$$\rho(\Pi) = \inf_{g\in A} \int \Pi(\omega)g(\omega)\mathbb{P}(d\omega) = \inf_{g\in A} \mathbb{E}[\Pi \cdot g].$$

where A: a convex set and $\forall g \in A$, $\int g \cdot d\mathbb{P} = 1$ and $g \geq 0$. Specifically, $g \equiv 1$ corresponds to the expected-value problem.

Now, we apply a directional derivative, $\rho'(\Pi_x; d)$ in a vector d. Then,

$$\begin{aligned}\rho_x'(\Pi_x; d) &= \lim_{\tau\downarrow 0} \frac{f(\Pi_x + \tau \cdot d) - f(\Pi_x)}{\tau}. \\ &= \inf_{g\in A(x)} \int \Pi_x^{'}(x; d) \cdot g(\omega) \cdot d\mathbb{P}(\omega) = \inf_{g\in A(x)} \mathbb{E}[\Pi_x^{'}(x; d) \cdot g].\end{aligned}$$

with the optimal densities $A(x) \subseteq A$ such that $A(x) = \{g \in A : \rho(\Pi_x) = \mathbb{E}[\Pi_x \cdot g]\}$ and $A(x)$ depends on Π.

Each gradient of $\int[\mathbb{E}(\Pi) - \Pi]_+ d\mathbb{P}$ is calculated as follows: First, I choose any function $\nu(\omega)$ such that

$$\nu(\omega) = \begin{cases} 1, & \text{if } \mathbb{E}[\Pi] - \Pi > 0. \\ 0, & \text{if } \mathbb{E}[\Pi] - \Pi < 0. \\ \{0, 1\}, & \text{if } \mathbb{E}[\Pi] - \Pi = 0. \end{cases}$$

Second, I calculate $h(\omega) = \mathbb{E}[\nu(\omega)] - \nu(\omega)$. Then, $g(\omega) = -1 + \lambda h(\omega)$. Now, to prove that $\hat{x}^{\mathrm{RA}} \leq \hat{x}^{\mathrm{RN}}$, it is enough to show that

$$\inf_{g\in A(x)} (\mathbb{E}[\Pi_x^{'}(x; d) \cdot (-1 + \lambda h(\omega))]) \geq 0, \quad \text{with } x = \hat{x}^{\mathrm{RN}}. \tag{3.20}$$

Also, it is equivalent to prove that

$$\inf_{g \in A(x)} (\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}; d) \cdot h(\omega)]) \geq 0. \tag{3.21}$$

because $\inf_{g \in A(x)}(\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RN}}; d)]) = 0$. When $x = \hat{x}^{\mathrm{RN}}$,

$$\begin{aligned}\mathbb{E}[\Pi_x^{'}(x; d) \cdot h(\omega)] &= \mathbb{E}[\Pi_x^{'}(x; d) \cdot (\mathbb{E}[\nu(\omega)] - \nu(\omega))]. \\ &= \mathbb{E}[\Pi_x^{'}(x; d) \cdot (-\nu(\omega))] = -\mathbb{E}[\Pi_x^{'}(x; d) | \mathbb{E}(\Pi) > \Pi].\end{aligned}$$

Thus, $\mathbb{E}[\Pi_x^{'}(x; d) \cdot h(\omega)] > 0$ because $\Pi_x^{'}(x; d) = -\bar{c} < 0$ when $\mathbb{E}[\Pi] > \Pi$.

The next step is to prove that $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1}$ similar to the previous section. Thus, it is enough to show that

$$\inf_{g \in A(x)} (\mathbb{E}[\Pi_x^{'}(x; d) \cdot (-1 + \lambda_2 h(\omega))]) \geq 0, \quad \text{with } x = \hat{x}^{\mathrm{RA}_1}. \tag{3.22}$$

Also similar to the last subsection, it is also enough to prove that equivalently:

$$\inf_{g \in A(x)} (\mathbb{E}[\Pi_x^{'}(\hat{x}^{\mathrm{RA}_1}; d) \cdot h(\omega)]) \geq 0. \tag{3.23}$$

Then, when $x = \hat{x}^{\mathrm{RA}_1}$,

$$\mathbb{E}[\Pi_x^{'}(x; d) \cdot h(\omega)] = -\mathbb{E}[\Pi_x^{'}(x; d) | \mathbb{E}(\Pi) > \Pi] > 0.$$

Therefore, $\hat{x}^{\mathrm{RA}_2} \leq \hat{x}^{\mathrm{RA}_1} \leq \hat{x}^{\mathrm{RN}}$.

3.4 Extension to General Law-Invariant Coherent Measures of Risk

Lemma 3. *The function $x \mapsto \varkappa_{\mathscr{M}}[Z(x)]$ is nondecreasing on $\mathbb{R}_+$.*

Proof. By Lemma 1, each function $x \mapsto r_\beta[Z(x)]$ is nondecreasing, for every $\beta \in (0, 1)$. Then the integral over β with respect to any nonnegative measure μ is nondecreasing as well. Taking the supremum in (2.18) does not change this property. □

Owing to this result, Proposition 1 also holds true for the mean–risk model with the risk functional $r[\cdot] = \varkappa_{\mathscr{M}}[\cdot]$. The proof is identical.

For closed-form optimal solutions under general law-invariant nonatomic coherent measures of risk, I apply the equation (3.11) to the equation (2.18) and obtain the following equation such as:

$$\rho[\bar{\Pi}(x,D)] = \bar{c}x + \sup_{\mu\in\mathscr{M}} \int_0^1 \left(\mathbb{E}[Z_x^1](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}} \left\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x^1)_+]\right\}\right)\mu(d\beta). \tag{3.24}$$

Then,

$$\frac{d\rho[\bar{\Pi}(x,D)]}{dx} = \bar{c} + \sup_{\mu\in\mathscr{M}} \int_0^1 \left(\bar{r}(\lambda\beta - 1)P[D > x] - \bar{r}\lambda P[\{Z_x^1 < \hat{\eta}\} \cap \{D > x\}]\right)\mu(d\beta).$$

Similarly to mean-deviation from quantile case,

$$P[\{Z_x^1 < \hat{\eta}\} \cap \{D > x\}] = P[Z_x^1 < \hat{\eta} | D > x]P[D > x] = 0.$$

Therefore,

$$\begin{aligned}\frac{d\rho[\bar{\Pi}(x,D)]}{dx} &= \bar{c} + \sup_{\mu\in\mathscr{M}} \int_0^1 \left(\bar{r}(\lambda\beta - 1)P[D > x]\right)\mu(d\beta).\\ &= \bar{c} + \bar{r}\left(-1 + \lambda \sup_{\mu\in\mathscr{M}} \int_0^1 \beta\hat{\mu}(d\beta)\right)P[D > x].\end{aligned}$$

Therefore, the closed-form solutions for general coherent measures of risk are as follows:

$$\hat{x}^{\mathrm{RA}} = \bar{F}_D^{-1}\left(\frac{\bar{c}}{\bar{r}(1 - \lambda\bar{\beta})}\right) \le \bar{F}_D^{-1}\left(\frac{\bar{c}}{\bar{r}}\right) = \hat{x}^{\mathrm{RN}}, \quad \text{where } \bar{\beta} = \sup_{\mu\in\mathscr{M}} \int_0^1 \beta\hat{\mu}^{\mathrm{RN}}(d\beta).$$

Chapter 4

Multi-Product Newsvendor Model - Coherent Measures of Risk

The work in this chapter was constructed by the working paper, ”A multi-product risk-averse newsvendor with law-invariant coherent measures of risk” by S. Choi, A. Ruszczyński and Y. Zhao (2009). This work was also presented by me at the special interest group of iFORM (Interfaces of Finance, Operations and Risk Management) at 2009 INFORMS (INstitute For Operations Research and Management Sciences) MSOM (Manufacturing and Service Operations Management) Annual Conference in Boston at June 28.

This model presents a considerable challenge, both analytically and computationally, because the objective function cannot be decomposed by product and I have to look at the totality of all products as a portfolio rather than one-by-one. In particular, one has to characterize the impact of risk aversion and demand dependence on the optimal solution, identify efficient ways to find the optimal solution, and connect this model to the financial portfolio theory. While Tomlin and Wang (2005) study a two-product system under CVaR, their focus is on the design of material flow topology and thus is very different from mine.

I should also point out that in most practical cases where this model is relevant (either manufacturing or retailing), firms may have a large number of heterogenous products. Due to the complex nature of risk optimization models, they become practically intractable for problems of these dimensions. Thus, it is theoretically interesting and practically useful to obtain fast approximation for large size problems and study the asymptotic behavior of the system as the number of products tends to infinity.

This work contributes to literature in the following ways: First, this chapter considers a multi-product risk-averse newsvendor using law-invariant coherent measures of risk. As I argued in §2.4, coherent measures of risk can be more attractive than the expected utility theory in multi-product newsvendor problems due to their additional axioms of Translation Equivariance and Positive Homogeneity.

Then I first establish several fundamental properties for the model at §4.3, e.g., the convexity of the model, the symmetry of the solution, the impact of risk aversion and the impact of the shift in mean demand.

I then consider large but finite number of independent heterogenous products, for which I develop closed-form approximations at §4.4.2 which are exact in single-product case. The approximations are as simple to compute as the risk-neutral solutions. I also show that under certain regularity conditions, risk-neutral solutions are asymptotically optimal under risk aversion as the number of products tends to be infinity. This asymptotic result has an important economic implication: companies with many products or product families with low demand dependence need to look only at risk-neutral solutions, even if they are risk-averse.

The impact of dependent demand under risk aversion poses a substantial analytical challenge. By utilizing the concept of associated random variables, I am able to prove at §4.5 that in a two-identical product system, positively dependent demand leads to a lower optimal order quantity than independent demand under risk aversion, while negatively dependent demand leads to a higher optimal order quantity under risk aversion. Then I study the impact of risk aversion with dependent demands. More interestingly, the optimal order quantity can increase in the degree of risk aversion when the demands are strongly negatively correlated while in most cases the order quantity decreases in the degree of risk aversion. In §4.6, I analyze three special examples with perfectly negative correlations in which each marginal demand distribution is uniform. Then I obtain a closed-form optimal solution in each case, which is equal to, higher than or lower than risk-neutral solutions, respectively. These analytical results can be extended to general uniform distributions and other symmetric random variables.

4.1 Multi-Product Inventory Models

The multi-product newsvendor problem is a classical model in the inventory management literature. In this model, a newsvendor has multiple products to be procured and sold in a single selling season. The newsvendor only knows the demand for each product as a form of probability distribution at the time of making orders and the demand realization is determined after some time during the season. When the newsvendor orders less than the actual demand realization for any product, the excessive demand is lost. On the other hand, when the newsvendor orders more than the actual demand realization, the excessive inventory is sold at a loss. The objective of the newsvendor is to determine his ordering quantity for each product to optimize a certain profit or cost function.

The literature of the multi-product newsvendor model has mainly used risk-neutral performance measures as an objective function. For example, the newsvendor optimizes the expected profit or cost. Under these objective functions, the model is decomposable and the newsvendor can consider each product separately as multiple single-product newsvendor models, unless resource constraints are imposed or demand substitution is allowed. Under risk-averse objective functions, however, the model is generally not decomposable. One needs to consider all products simultaneously, as a portfolio.

Below, I first review the literature of risk-neutral multi-product inventory models by ways products interact. Hadley and Whitin (1962) consider a multi-product newsvendor model with storage capacity or budget constraints, and provide the solution methods based on Lagrangian multiplier. Porteus (1990) presents a thorough review of various newsvendor models. Veinott (1965) considers the dynamic version of the multi-product inventory models in a multi-period setting, with general assumptions in demand process, cost parameters and lead times. Conditions under which myopic policy is optimal are identified. Ignall and Veinott (1969) and Heyman and Sobel (1984) extend the work by identifying new conditions for the myopic policy in models with risk-neutral assumption, see Evans (1967), Federgruen and Zipkin (1984), DeCroix and Arreola-Risa (1998) and Aviv and Federgruen (2001) for exact analysis and approximations.

Starting from 1990s', multi-product newsvendor models gain attention again by allowing demand substitution, where unsatisfied demand of one product can be satisfied by on-hand inventory of another product. Bassok, Anupindi and Akella (1999) studies the manufacturer-driven downward substitution based on a stochastic programming approach. Van Ryzin and Mahajan (1999) consider a multi-product newsvendor with customer-driven substitution. The demand is modeled by the multinomial logit model (MNL). The paper focuses on the static substitution and develops structural results for product assortment and replenishment planning. Smith and Agrawal (2000) considers a similar problem but using a different demand model: the exogenous demand model, where a customer does not find her first choice may either pick her second choice or leave. The authors developed upper and lower bounds for the performance measures. Mahajan and van Ryzin (2001) studies dynamic substitution where customer makes choice dynamically based on current inventory levels of all products. Sample path based simulation algorithm is proposed to determine the optimal ordering quantity. Van Ryzin and Mahajan (1999) provides a review on multi-item inventory systems with substitution.

4.2 Problem Formulation

Given products $j = 1, \ldots, n$, let $x = (x_1, x_2, \ldots, x_n)$ be the vector of ordering quantities and let $D = (D_1, \ldots, D_n)$ be the demand vector. I also define $r = (r_1, \ldots, r_n)$ to be the price vector, $c = (c_1, \ldots, c_n)$ to be cost vector, and $s = (s_1, \ldots, s_n)$ to be the vector of salvage values. Finally, let $f_{D_j}(\cdot)$ and $F_{D_j}(\cdot)$ be the marginal probability density function (pdf), if it exists, and the marginal cumulative distribution function (cdf) of D_j, respectively. Denote $\bar{F}_{D_j}(\xi) = 1 - F_{D_j}(\xi)$.

Setting $\bar{c}_j = c_j - s_j$ and $\bar{r}_j = r_j - s_j$, I can write the profit function as follows:

$$\Pi(x, D) = \sum_{j=1}^{n} \Pi_j(x_j, D_j). \tag{4.1}$$

where

$$\begin{aligned} \Pi_j(x_j, D_j) &= -\bar{c}_j x_j + \bar{r}_j \min\{x_j, D_j\}. \\ &= (r_j - c_j)x_j - (r_j - s_j)(x_j - D_j)_+, \quad j = 1, \ldots, n. \end{aligned} \tag{4.2}$$

Then, $(r_j - c_j)$ and $(c_j - s_j)$ are underage and overage costs for product j, respectively. I assume that the demand vector D is random and nonnegative. Therefore, for every $x \geq 0$ the profit $\Pi(x, D)$ is a real bounded random variable.

The risk-neutral multi-product newsvendor optimization problem is to maximize the expected profit:

$$\max_{x \geq 0} \mathbb{E}[\Pi(x, D)]. \tag{4.3}$$

This problem can be decomposed into independent problems, one for each product. Thus, under risk-neutrality, a multi-product newsvendor problem is equivalent to multiple single-product newsvendor problems. Then the optimal solution is given similarly to the equation (3.6) as follows.

$$\hat{x}_j^{\mathrm{RN}} = \bar{F}_{D_j}(\alpha), \text{ with } \alpha = (c_j - s_j)/(r_j - s_j), \quad j = 1, \ldots, n. \tag{4.4}$$

Under a coherent measure of risk, the optimization problem of the risk-averse newsvendor is defined as follows:

$$\min_{x \geq 0} \rho[\Pi(x, D)]. \tag{4.5}$$

where $\rho[\cdot]$ is a law-invariant coherent measure of risk, and $\Pi(x, D)$ represents the profit of the newsvendor, as defined in (4.1). It is worth stressing that problem (4.5) cannot be decomposed into independent subproblems, one for each product. Thus, it is necessary to consider the portfolio of products as a whole.

4.3 Analytical Results for Independent Demands

In this section, I provide several analytical results for the multi-product newsvendor model under coherent measures of risk. I assume independent demands. These results lay the theoretical foundation for the paper.

Proposition 2. *If $\rho[\cdot]$ is a coherent measure of risk, then $\rho[\Pi(x, D)]$ is a convex function of x.*

Proof. I first note that $\Pi(x, D) = \sum_{j=1}^{n} \Pi_j(x_j, D_j)$ is concave in x. That is, for any $0 \leq \alpha \leq 1$ and all x, y,

$$\Pi(\alpha x + (1 - \alpha)y, D) \geq \alpha\Pi(x, D) + (1 - \alpha)\Pi(y, D), \qquad \text{for all } D.$$

Using the monotonicity axiom, I obtain

$$\begin{aligned}\rho[\Pi(\alpha x + (1 - \alpha)y, D)] &\leq \rho[\alpha\Pi(x, D) + (1 - \alpha)\Pi(y, D)]. \\ &\leq \alpha\rho[\Pi(x, D)] + (1 - \alpha)\rho[\Pi(y, D)].\end{aligned}$$

The second inequality follows by the axiom of convexity. □

Observe that I did not use the axiom of positive homogeneity, and thus Proposition 2 extends to more general convex measures of risk. I next prove an intuitively clear argument that identical products should be ordered in equal quantities under coherent measures of risk.

Proposition 3. *Assume that all products are identical, i.e., prices, ordering costs and salvage values are the same across all products. Furthermore, let the joint probability distribution of the demand be symmetric, that is, invariant with respect to permutations of the demand vector. Then, for every law invariant coherent measure of risk $\rho[\cdot]$, one of optimal solutions of problem* (4.5) *is a vector with equal coordinates, $\hat{x}_1^{\mathrm{RA}} = \hat{x}_2^{\mathrm{RA}} = \cdots = \hat{x}_n^{\mathrm{RA}}$.*

Proof. An optimal solution exists, because without loss of generality I can assume that x is bounded by some large constant, and $\rho[\Pi(x, D)]$ is continuous with respect to x (see Ruszczyński and Shapiro (2006a)).

Let me consider an arbitrary order vector $x = (x_1, \ldots, x_n)$ and let P be an $n \times n$ permutation matrix. Then, the distribution of profit associated with Px is the same as that associated with x. There are $n!$ different permutations of x and let us denote them $x^1, \ldots, x^{n!}$. Consider the point

$$y = \frac{1}{n!} \sum_{i=1}^{n!} x^i.$$

It has all coordinates equal to the average of the coordinates x_j. As the joint probability distribution of $D_1, D_2, \ldots, D_n$ is symmetric, the distribution of $\Pi(x^i, D)$ is the same for each i. By Proposition 2 and by law-invariance of $\rho[\cdot]$, I obtain

$$\rho[\Pi(y, D)] \le \frac{1}{n!} \sum_{i=1}^{n!} \rho[\Pi(x^i, D)] = \rho[\Pi(x, D)].$$

This means that for every plan x, the corresponding plan y with equal orders is at least as good. As an optimal plan exists, there is an optimal plan with equal orders. $\square$

Note that Proposition 3 only requires symmetric joint demand distribution, but not independent demands. To study the impact of the degree of risk aversion, let me first focus on a specific variability functional – the weighted mean-deviation from quantile, given by (2.13). The corresponding measure of risk has the form,

$$\rho[V] = -\mathbb{E}[V] + \lambda r_\beta[V]. \tag{4.6}$$

Due to (2.14), I can write

$$\rho[V] = -(1 - \lambda\beta)\mathbb{E}[V] + \lambda\beta \mathrm{AVaR}_\beta[V]. \tag{4.7}$$

Consider the problem

$$\min_{x \ge 0} \Big\{ - \mathbb{E}[\Pi(x, D)] + \lambda r_\beta[\Pi(x, D)] \Big\}. \tag{4.8}$$

Proposition 4. *Assume that all products are identical and demands for all products are iid and have a continuous distribution. Let $\hat{x}^{\mathrm{RA}_1}$ be the solution of problem* (4.8) *for $\lambda = \lambda_1 > 0$, having equal coordinates. If $\lambda_2 \geq \lambda_1$ then there exists a solution $\hat{x}^{\mathrm{RA}_2}$ of problem* (4.8) *for $\lambda = \lambda_2$, having equal coordinates and such that $\hat{x}_j^{\mathrm{RA}_2} \leq \hat{x}_j^{\mathrm{RA}_1}$, $j = 1, \ldots, n$.*

Proof. In view of Proposition 3, I can assume that the coordinates of $\hat{x}^{\mathrm{RA}_i}$ are equal, $i = 1, 2$. My argument extends Proposition 1 at the single-product case in Chapter 3 to the multi-product case.

Since all coordinates of the solutions are assumed equal, with a slight abuse of notation, I consider a fixed decision variable x and simplify (4.1)–(4.2) to

$$\varPi(x, D) = nx(r-c) - (r-s)\sum_{j=1}^{n}(x - D_j)_+ = -n(c-s)x + (r-s)\sum_{j=1}^{n}\min(x, D_j). \quad (4.9)$$

For every nonrandom a we have $r_\beta[V + a] = r_\beta[V]$ and thus

$$\begin{aligned} -\mathbb{E}[\varPi(x,D)] + \lambda_2 r_\beta[\varPi(x,D)] &= -\mathbb{E}[\varPi(x,D)] + \lambda_1 r_\beta[\varPi(x,D)] + (\lambda_2 - \lambda_1) r_\beta[\varPi(x,D)]. \\ &= -\mathbb{E}[\varPi(x,D)] + \lambda_1 r_\beta[\varPi(x,D)] + (\lambda_2 - \lambda_1)(r-s) r_\beta\Big[\sum_{j=1}^{n}\min(x, D_j)\Big]. \end{aligned}$$

As $\hat{x}^{\mathrm{RA}_1}$ minimizes the sum of the first two terms, it remains to show that the function

$$x \mapsto r_\beta\Big[\sum_{j=1}^{n}\min(x, D_j)\Big].$$

is nondecreasing on $\mathbb{R}_+$. Consider the random variable $Z_x = \sum_{j=1}^{n}\min(x, D_j)$. From formula (2.14) I obtain:

$$\frac{1}{\beta} r_\beta[Z_x] = \mathbb{E}[Z_x] - \max_{\eta\in\mathbb{R}}\Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)_+]\Big\}.$$

Then I differentiate both terms of the right-hand side with respect to x. I have:

$$\frac{d\mathbb{E}[Z_x]}{dx} = \sum_{j=1}^{n} P[D_j > x] = nP[D_j > x].$$

To differentiate the second term, I define $\hat{\eta}$ to be the maximizer (among $\eta \in \mathbb{R}$) at a given x, equal to the β-quantile of Z_x. Clearly, $\hat{\eta}$ depends on x, but I suppress this dependence here for the ease of exposition. I consider two cases.

Case (i): $\hat{\eta} < nx$.

If $\hat{\eta}$ is unique, I can use the differential properties of the optimal value:

$$\frac{d}{dx}\Big[\max_{\eta\in\mathbb{R}}\Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)_+]\Big\}\Big] = -\frac{1}{\beta}\frac{d}{dx}\Big\{\mathbb{E}[(\hat{\eta} - Z_x)_+]\Big\}.$$

Note that differentiation here is only on Z_x (see Theorem 4.13 of Bonnans and Shapiro (2000)). By differentiation I obtain

$$\frac{d}{dx}\Big\{\mathbb{E}[(\hat{\eta} - Z_x)_+]\Big\} = -\mathbb{E}[\mathbb{1}_{\{Z_x<\hat{\eta}\}}\sum_{j=1}^{n}\mathbb{1}_{\{D_j>x\}}] = -\sum_{j=1}^{n}P[\{Z_x < \hat{\eta}\} \cap \{D_j > x\}].$$

The events $\{Z_x < \hat{\eta}\}$ and $\{D_j > x\}$ are dependent, but for independent D_j we have

$$\begin{aligned} P[\{Z_x < \hat{\eta}\} \cap \{D_j > x\}] &= P[Z_x < \hat{\eta}|D_j > x]P[D_j > x]. \\ &\leq P[Z_x < \hat{\eta}]P[D_j > x] = \beta P[D_j > x]. \end{aligned}$$

The inequality holds true because

$$\begin{aligned} P[Z_x < \hat{\eta}|D_j > x] &= P\Big[\sum_{i\neq j}\min(x, D_i) < \hat{\eta} - x\Big]. \\ &\leq P\Big[\sum_{i\neq j}\min(x, D_i) < \hat{\eta} - \min(x, D_j)\Big] = P[Z_x < \hat{\eta}]. \end{aligned}$$

Thus

$$\frac{d}{dx}\Big[\max_{\eta\in\mathbb{R}}\Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)_+]\Big\}\Big] \leq \sum_{j=1}^{n}P[D_j > x] = nP[D_j > x].$$

I conclude that

$$\frac{d}{dx}r_\beta[Z_x] \geq 0.$$

If $\hat{\eta}$ is not unique, one can consider the left and the right derivatives of the optimal value, by substituting the largest and the smallest β-quantile for $\hat{\eta}$ in the calculations above. I observe that the event $\{Z_x < \hat{\eta}\}$ does not change, and conclude that the right derivative is non-negative.

Case (ii): $\hat{\eta} = nx$.

As Z_x has an atom at nx, for sufficiently small x I can just substitute $\hat{\eta} = nx$ in formula (2.14):

$$\frac{1}{\beta} r_\beta[Z_x] = \mathbb{E}[Z_x] - \Big\{nx - \frac{1}{\beta}\mathbb{E}[(nx - Z_x)]\Big\}.$$

Taking derivative with respect to x, I conclude that

$$\frac{d}{dx} r_\beta[Z_x] = \beta\{nP[D_j > x] - (n - \frac{1}{\beta}(n - nP[D_j > x]))\}.$$
$$= n(1 - P[D_j > x])(1 - \beta) \geq 0.$$

as required. In the general case, I consider the left derivative here, because if $\hat{\eta}(x) = nx$ then $\hat{\eta}(y) = ny$ for all $y < x$, and I arrive at the same conclusion. □

I can extend this Proposition of the monotonicity property to all law-invariant coherent measures of risk. Observe that my assumption about continuous distribution of the demand implies that the probability space is nonatomic.

Consider the problem

$$\min_{x \geq 0} \Big\{ - \mathbb{E}[\Pi(x, D)] + \lambda \varkappa_{\mathscr{M}}[\Pi(x, D)]\Big\}. \tag{4.10}$$

where $\varkappa_{\mathscr{M}}[V]$ is given by (2.18).

Proposition 5. *Assume that all products are identical and demands for all products are iid. Let $\hat{x}^{\mathrm{RA}_1}$ be the solution of problem* (4.10) *for $\lambda = \lambda_1 > 0$, having equal coordinates. If $\lambda_2 \geq \lambda_1$ then there exists a solution $\hat{x}^{\mathrm{RA}_2}$ of problem* (4.10) *for $\lambda = \lambda_2$, having equal coordinates and such that $\hat{x}_j^{\mathrm{RA}_2} \leq \hat{x}_j^{\mathrm{RA}_1}$, $j = 1, \dots, n$.*

Proof. As in the proof of Proposition 4, each function $x \mapsto r_\beta[\Pi(x, D)]$ is nondecreasing, for every $\beta \in (0, 1)$. Then the integral over β with respect to any nonnegative measure μ is nondecreasing as well. Taking the supremum in (2.18) does not change this property. Therefore, Proposition 5 holds true also for the mean–risk model with the risk functional $r[\cdot] = \varkappa_{\mathscr{M}}[\cdot]$. □

I need to point out that the parameter λ represents how sensitive a newsvendor is to the risk defined by another parameter β. Thus, β is another important factor to determine risk in my model. However, the analysis of β is much more demanding than that of β.

Finally, I discuss the impact of the shifts in mean demands on the optimal order quantities under general coherent measures of risk. For this purpose, I consider identical products and demands with identical and independent probability distributions except that $\mu_j = \mathbb{E}[D_j]$, $j = 1, \dots, n$, may be different. Without loss of generality, I assume that $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$. Consider the demand vector $\tilde{D}_j = D_j - \mu_j + \mu_1$. Because it has *iid* components, due to Proposition 3 there exists an optimal order vector $\tilde{x}$ with equal coordinates: $\tilde{x}_1 = \tilde{x}_2 = \cdots = \tilde{x}_n$, for the risk-averse multi-product newsvendor with $\tilde{D}$ as the demand vector. I can interpret the demand D as a sum of the random demand $\tilde{D}$ and a deterministic demand vector h with coordinates $h_j = \mu_j - \mu_1$. If $\tilde{x}_j > 0$ then

$$\hat{x}_j = \tilde{x}_j + \mu_j - \mu_1, \quad j = 1, \dots, n$$

are the optimal order quantities for the risk-averse problem with D as the demand vector.

This can be easily shown analytically. From formula (4.2) I obtain

$$\Pi_j(x_j, \tilde{D}_j) = \Pi_j(x_j + \mu_j - \mu_1, D_j) - (\bar{r}_j - \bar{c}_j)(\mu_j - \mu_1).$$

Using the Translation Equivariance axiom of a coherent measure of risk $\rho[\cdot]$, I obtain

$$\rho[\Pi(x, \tilde{D})] = \rho[\Pi(x + h, D)] + \langle \bar{r} - \bar{c}, h \rangle. \tag{4.11}$$

where $\langle \cdot, \cdot \rangle$ is the inner product. Therefore, if $\tilde{x}$ minimizes the left hand side (over $x \geq 0$) and is positive, then it also minimizes the right hand side over $x + h \geq 0$. Thus $\hat{x} = \tilde{x} + h$ is the solution of the problem

$$\min_{x \geq 0} \rho[\Pi(x, D)].$$

This is true for *every* coherent measure of risk $\rho[\cdot]$.

For law-invariant measures of risk I can also analyze the impact of the parameter λ on the solution with the shifts in mean demands . The considerations in this case follow directly from Proposition 5, with the use of identity (4.11).

4.4 Asymptotic Analysis and Closed-Form Approximations

In this section, I study the asymptotic behavior of the risk-averse newsvendor model when the number of products tends to infinity, and develop closed-form approximations to its optimal solution in the case of a large but finite number of products. I assume heterogenous products with independent demands.

4.4.1 Asymptotic Optimality of Risk-Neutral Solutions

I start from the derivation of error bounds for the risk-neutral solution. Consider a sequence of products $j = 1, 2, \ldots$, with corresponding prices r_j, costs c_j, and salvage values s_j. I assume that $s_j < c_j < r_j$, and that all these quantities are uniformly bounded for $j = 1, 2, \ldots$. As before, I set $\bar{c}_j = c_j - s_j$ and $\bar{r}_j = r_j - s_j$. The demands $D_1, D_2, \ldots$ are assumed to be independent.

Consider the risk-neutral optimal order quantities

$$\hat{x}_j^{\mathrm{RN}} = \bar{F}_j^{-1}\Big(\frac{\bar{c}_j}{\bar{r}_j}\Big), \quad j = 1, 2, \ldots. \tag{4.12}$$

I assume that the following conditions are satisfied:

(i) There exist $x^{\min} > 0$ and $x^{\max}$ such that

$$x^{\min} \leq \hat{x}_j^{\mathrm{RN}} \leq x^{\max}, \quad j = 1, 2, \ldots.$$

(ii) There exists $\sigma_{\min} > 0$ such that

$$\mathbb{V}\mathrm{ar}[\,\min(\hat{x}_j^{\mathrm{RN}}, D_j)] \geq \sigma_{\min}^2, \quad j = 1, 2, \ldots.$$

My intention is to evaluate the quality of the risk-neutral solution $\hat{x}^{\mathrm{RN}}$ in the risk-averse problem such that

$$\min_{x_1,\dots,x_n} \rho\Big[\frac{1}{n}\sum_{j=1}^{n} \Pi_j(x_j, D_j)\Big]. \tag{4.13}$$

Observe that in (4.13) I consider the average profit per product, rather than the total profit, as in (4.5). The reason is that I intend to analyze properties of the optimal value of this problem as $n \to \infty$ and I want the limit of the objective value of (4.13) to exist. Owing to the Positive Homogeneity axiom, problems (4.5) and (4.13) are equivalent.

I denote by $\hat{\rho}_n$ the optimal value of problem (4.13). I also introduce the following notation,

$$\mu_j^{\mathrm{RN}} = \mathbb{E}[\min(\hat{x}_j^{\mathrm{RN}}, D_j)], \qquad \bar{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} \bar{r}_j \mu_j^{\mathrm{RN}},$$

$$(\sigma_j^{\mathrm{RN}})^2 = \mathbb{V}\mathrm{ar}[\min(\hat{x}_j^{\mathrm{RN}}, D_j)], \qquad \bar{s}_n^2 = \frac{1}{n^2}\sum_{j=1}^{n} \bar{r}_j^2 (\sigma_j^{\mathrm{RN}})^2.$$

Finally, I denote by $\mathscr{N}$ the standard normal variable.

Proposition 6. *Assume that $\rho[\cdot]$ is a law-invariant coherent measure of risk and the space $(\Omega, \mathscr{F}, P)$ is nonatomic. Then*

$$\limsup_{n\to\infty} \frac{1}{\bar{s}_n}\left(\rho\Big[\frac{1}{n}\sum_{j=1}^{n} \Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] - \hat{\rho}_n\right) \le \rho[\mathscr{N}]. \tag{4.14}$$

Proof. Denote $Z^n = \frac{1}{n}\sum_{j=1}^{n} \bar{r}_j \min(\hat{x}_j^{\mathrm{RN}}, D_j)$. I have $\mathbb{E}[Z^n] = \bar{\mu}_n$, $\mathbb{V}\mathrm{ar}[Z^n] = \bar{s}_n^2$, and

$$\frac{1}{n}\sum_{j=1}^{n} \Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left[\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\right] + (Z^n - \bar{\mu}_n).$$

Owing to conditions (i) and (ii), the sequence $\{\bar{r}_j \min(\hat{x}_j^{\mathrm{RN}}, D_j)\}$, $j = 1, 2 \dots$, satisfies the Lindeberg condition (see, e.g, Feller (1971), p. 262). I can therefore apply the Central Limit Theorem for non-identical independent random variables, to conclude that

$$\frac{Z^n - \bar{\mu}_n}{\bar{s}_n} \xrightarrow{\mathscr{D}} \mathscr{N}. \tag{4.15}$$

Here the symbol $\xrightarrow{\mathscr{D}}$ denotes convergence in distribution. By the Translation Equivariance axiom

$$\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] = -\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left[\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\right] + \rho[Z^n - \bar{\mu}_n].$$

At any other value of x, in particular, at a solution of problem (4.13), I have

$$\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(x_j, D_j)\Big] \geq -\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left[\Pi_j(x_j, D_j)\right] \geq -\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left[\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\right].$$

because of (2.17) and because $\hat{x}_j^{\mathrm{RN}}$ maximizes $\mathbb{E}[\Pi_j(x_j, D_j)]$. Combining the last two relations I conclude that

$$\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] - \hat{\rho}_n \leq \rho[Z^n - \bar{\mu}_n].$$

Dividing both sides by $\bar{s}_n$ and using the Positive Homogeneity axiom I obtain

$$\frac{1}{\bar{s}_n}\left(\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] - \hat{\rho}_n\right) \leq \rho\left[\frac{Z^n - \bar{\mu}_n}{\bar{s}_n}\right]. \tag{4.16}$$

Let $\Phi_n(\cdot)$ be the cdf of $(Z^n - \bar{\mu}_n)/\bar{s}_n$. By (4.15), $\Phi_n \to \Phi$ pointwise, where $\Phi(\cdot)$ is the cdf of the standard normal distribution. As the risk measure $\rho[\cdot]$ is law-invariant and the space is nonatomic, I have $\rho[(Z^n - \bar{\mu}_n)/\bar{s}_n] = \rho[\Phi_n^{-1}(\mathscr{U})]$, where $\mathscr{U}$ is a uniform random variable on $[0, 1]$. By the continuity of $\rho[\cdot]$ in the space of integrable random variables, the right-hand side of inequality (4.16) tends to $\rho[\mathscr{N}]$ as $n \to \infty$. Passing to the limit in (4.16), I obtain (4.14). $\square$

Conditions (i) and (ii) imply that $\bar{s}_n = O(1/\sqrt{n})$, and thus it follows from (4.14) that

$$\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] \leq \min_{x_1,\dots,x_n}\rho\Big[\frac{1}{n}\sum_{j=1}^{n}\Pi_j(x_j, D_j)\Big] + O\Big(\frac{1}{\sqrt{n}}\Big).$$

Asymptotically, the difference between the optimal value of (4.13) and the value obtained by using the risk-neutral solution disappears at the rate of $1/\sqrt{n}$. For a firm dealing with very many products having independent demands, the risk-neutral solution is a reasonable sub-optimal alternative to the risk-averse solution.

4.4.2 Adjustments in Mean–Deviation from Quantile Models

When the number of products is moderately large, I develop close-form approximations to the optimal risk-averse solution. My idea is to use the risk-neutral solution as the starting point, and to calculate an appropriate correction to account for risk-aversion.

I first consider deviation from quantile as a measure of variability defined in (2.13). Recall that the corresponding mean–risk model (4.6) is equivalent to the minimization of a combination of the mean and the Conditional Value-at-Risk, as in (4.7). I then consider a general coherent measure of risk in §4.4.3. I finally discuss several iterative methods that are based on the approximations in §4.4.4.

Similarly to the previous subsection, I use the notation $Z_x^n = \frac{1}{n}\sum_{j=1}^n \bar{r}_j \min(x_j, D_j)$ (with x as a subscript to stress the dependence of Z^n on x). Using (4.1)–(4.2), I can calculate the average profit per product as follows:

$$\bar{\Pi}(x, D) = \frac{1}{n}\sum_{j=1}^n \Pi_j(x_j, D_j) = -\frac{1}{n}\sum_{j=1}^n \bar{c}_j x_j + Z_x^n.$$

Thus,

$$\begin{aligned}\rho[\bar{\Pi}(x, D)] &= \frac{1}{n}\sum_{j=1}^n \bar{c}_j x_j + \big(-\mathbb{E}[Z_x^n] + \lambda r_\beta(Z_x^n)\big). \\ &= \frac{1}{n}\sum_{j=1}^n \bar{c}_j x_j + \Big(\mathbb{E}[Z_x^n](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}}\Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x^n)_+]\Big\}\Big).\end{aligned} \tag{4.17}$$

Similarly to the proof of Proposition 4, let $\hat{\eta}$ be the maximizer in (4.17), among $\eta \in \mathbb{R}$, at a fixed x. $\hat{\eta}$ is the β-quantile of Z_x^n. To take the partial derivative of $\rho[\bar{\Pi}(x, D)]$ with respect to x_j, I consider two cases similarly in single-product models at Chapter 3.

Case (i): $\hat{\eta} < \frac{1}{n}\sum_{j=1}^n \bar{r}_j x_j$.

Assuming that the quantile $\hat{\eta}$ is unique and differentiating the equation (4.17), I observe again that

$$\frac{\partial \rho[\bar{\Pi}(x, D)]}{\partial x_j} = \frac{\bar{c}_j}{n} + \frac{\bar{r}_j(\lambda\beta - 1)}{n} P[D_j > x_j] - \frac{\bar{r}_j \lambda}{n} P\big[\{Z_x^n < \hat{\eta}\} \cap \{D_j > x_j\}\big]. \tag{4.18}$$

Here I used Theorem 4.13 of Bonnans and Shapiro (2000) to avoid differentiating on $\hat{\eta}$.

Then let me analyze the last term on the right-hand side for $j = 1, 2, \ldots, n$:

$$\begin{aligned} P[\{Z_x^n < \hat{\eta}\} \cap \{D_j > x_j\}] &= P[Z_x^n < \hat{\eta} | D_j > x_j] P[D_j > x_j]. \\ &= P\Big[\frac{1}{n}\sum_{k\neq j}^{n} \bar{r}_k \min(x_k, D_k) < \hat{\eta} - \frac{\bar{r}_j x_j}{n}\Big] \cdot P[D_j > x_j]. \end{aligned} \tag{4.19}$$

Suppose $x_j \geq x_{\min}$, $j = 1, 2, \ldots,$. Owing to conditions (i) and (ii), exactly as in §4.4.1, for large n the random variable Z_x^n is approximately normally distributed with the mean $\bar{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} \bar{r}_j \mu_j$ and the variance $\bar{s}_n^2 = \frac{1}{n^2}\sum_{j=1}^{n} \bar{r}_j^2 \sigma_j^2$, where $\mu_j = \mathbb{E}[\min\{x_j, D_j\}]$ and $\sigma_j^2 = \mathbb{V}\mathrm{ar}(\min\{x_j, D_j\})$. Under normal approximation, the β-quantile of Z_x^n can be approximated as follows: $\hat{\eta} \simeq \bar{\mu}_n + z_\beta \bar{s}_n$, where z_β is the β-quantile of the standard normal variable. Similarly, $\frac{1}{n-1}\sum_{k\neq j}^{n} \bar{r}_k \min(x_k, D_k)$ is approximately normal with mean $\frac{1}{n-1}\sum_{k\neq j}^{n} \bar{r}_k \mu_k$ and variance $\frac{1}{(n-1)^2}\sum_{k\neq j}^{n} \bar{r}_k^2 \sigma_k^2$. Using these approximations and denoting by $\mathscr{N}$ the standard normal random variable I obtain:

$$\begin{aligned} P\Big[\frac{1}{n}\sum_{k\neq j}^{n} \bar{r}_k \min(x_k, D_k) < \hat{\eta} - \frac{\bar{r}_j x_j}{n}\Big] &\simeq P\Bigg[\mathscr{N} < \frac{-\bar{r}_j(x_j - \mu_j) + z_\beta \sqrt{\sum_{k=1}^{n} \bar{r}_k^2 \sigma_k^2}}{\sqrt{\sum_{k\neq j} \bar{r}_k^2 \sigma_k^2}}\Bigg]. \\ &= P\Bigg[\mathscr{N} < \frac{-\bar{r}_j(x_j - \mu_j)}{\sqrt{n-1}\,\gamma_{nj}} + z_\beta \sqrt{1 + \frac{\bar{r}_j^2 \sigma_j^2}{(n-1)\gamma_{nj}^2}}\Bigg]. \end{aligned} \tag{4.20}$$

where $\gamma_{nj} = \sqrt{\frac{1}{n-1}\sum_{k\neq j} \bar{r}_k^2 \sigma_k^2}$. As $\bar{r}_k^2 \sigma_k^2$ is uniformly bounded from above and below across all products, I conclude that γ_{nj} is bounded from above and below for all j and n.

This estimate can be put into (4.19) and thus (4.18) can be approximated as follows:

$$\begin{aligned} &\frac{\partial \rho[\bar{\Pi}(x, D)]}{\partial x_j} \\ &\simeq \frac{\bar{c}_j}{n} + \frac{\bar{r}_j P[D_j > x_j]}{n}\left(\lambda\beta - 1 - \lambda P\Bigg[\mathscr{N} < \frac{-\bar{r}_j(x_j - \mu_j)}{\sqrt{n-1}\,\gamma_{nj}} + z_\beta \sqrt{1 + \frac{\bar{r}_j^2 \sigma_j^2}{(n-1)\gamma_{nj}^2}}\Bigg]\right). \end{aligned} \tag{4.21}$$

My next step is to approximate the probability on the right-hand side of (4.21). To this end, I derive its limit and calculate a correction to this limit for a finite n. When $n \to \infty$ I have

$$P\left[\mathcal{N} < \frac{-\bar{r}_j(x_j - \mu_j)}{\sqrt{n-1}\gamma_{nj}} + z_\beta \sqrt{1 + \frac{\bar{r}_j^2 \sigma_j^2}{(n-1)\gamma_{nj}^2}}\right] \to \beta. \tag{4.22}$$

and thus

$$\frac{\partial \rho[\bar{\Pi}(x, D)]}{\partial x_j} \to \frac{1}{n}(\bar{c}_j - \bar{r}_j P[D_j > x_j]).$$

This means that the conditions of the risk-averse solution

$$\frac{\partial \rho[\bar{\Pi}(x, D)]}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n. \tag{4.23}$$

approaches that of the risk-neutral solution (4.12). Thus the risk-neutral solution will be used as the base value, to which corrections will be calculated.

I can estimate the difference between the probability in (4.22) and β for a large but finite n, by assuming that x is close to $\hat{x}^{\mathrm{RN}}$. Thus μ_j is close to $\mu_j^{\mathrm{RN}} = \mathbb{E}[\min\{\hat{x}_j^{\mathrm{RN}}, D_j\}]$ and σ_j is close to $\sigma_j^{\mathrm{RN}} = \sqrt{\mathbb{V}\mathrm{ar}(\min\{\hat{x}_j^{\mathrm{RN}}, D_j\})}$. Considering only the leading term with respect to $1/\sqrt{n-1}$, I obtain

$$P\left[\mathcal{N} < \frac{-\bar{r}_j(x_j - \mu_j)}{\sqrt{n-1}\gamma_{nj}} + z_\beta \sqrt{1 + \frac{\bar{r}_j^2 \sigma_j^2}{(n-1)\gamma_{nj}^2}}\right] \simeq P\left[\mathcal{N} < \frac{-\bar{r}_j(\hat{x}_j^{\mathrm{RN}} - \mu_j^{\mathrm{RN}})}{\sqrt{n-1}\gamma_{nj}^{\mathrm{RN}}} + z_\beta\right].$$

where $\gamma_{nj}^{\mathrm{RN}} = \sqrt{\frac{1}{n-1}\sum_{k \neq j} \bar{r}_k^2 (\sigma_k^{\mathrm{RN}})^2}$. The last probability can be estimated by the linear approximation derived at z_β. Observing that $P[\mathcal{N} < z_\beta] = \beta$ and that its derivative at $z = z_\beta$ is the standard normal density at z_β, I get

$$P\left[\mathcal{N} < \frac{-\bar{r}_j(\hat{x}_j^{\mathrm{RN}} - \mu_j^{\mathrm{RN}})}{\sqrt{n-1}\gamma_{nj}^{\mathrm{RN}}} + z_\beta\right] \simeq \beta - \delta_{nj}^{\mathrm{RN}},$$

with

$$\delta_{nj}^{\mathrm{RN}} = \frac{e^{-z_\beta^2/2}}{\sqrt{2\pi}} \frac{\bar{r}_j(\hat{x}_j^{\mathrm{RN}} - \mu_j^{\mathrm{RN}})}{\sqrt{n-1}\gamma_{nj}^{\mathrm{RN}}}, \quad j = 1, \ldots, n. \tag{4.24}$$

These estimates can be substituted to the formula (4.21) for the derivative and yield

$$\frac{\partial \rho[\bar{\Pi}(x, D)]}{\partial x_j} \simeq \frac{\bar{c}_j}{n} + \frac{\bar{r}_j}{n}(-1 + \lambda \delta_{nj}^{\mathrm{RN}}) P[D_j > x_j]. \tag{4.25}$$

Using the above approximations of the derivatives in equations (4.23), I obtain the first-order approximation of the risk-averse solution:

$$\boxed{\hat{x}_j^{\mathrm{RA}} \simeq \bar{F}_{D_j}^{-1}\left[\frac{\bar{c}_j}{\bar{r}_j(1-\lambda\delta_{nj}^{\mathrm{RN}})}\right], \quad j = 1, 2, \dots, n.} \tag{4.26}$$

Clearly, this approximation of $\hat{x}_j^{\mathrm{RA}}$ is increasing in n, decreasing in λ and tends to the risk-neutral solution as $n \to \infty$.

Case (ii): $\hat{\eta} = \frac{1}{n}\sum_{j=1}^{n} \bar{r}_j x_j$.

I have

$$\rho[\bar{\Pi}(x, D)] = \frac{1}{n}\sum_{j=1}^{n} \bar{c}_j x_j + \Big(\mathbb{E}[Z_x^n](\lambda\beta - 1) - \lambda\beta\Big\{\frac{1}{n}\sum_{j=1}^{n} \bar{r}_j x_j - \frac{1}{\beta}\mathbb{E}\big[\frac{1}{n}\sum_{j=1}^{n} \bar{r}_j x_j - Z_x^n\big]\Big\}\Big).$$

Taking derivative with respect to x_j yields,

$$\begin{aligned}
\frac{\partial\rho[\bar{\Pi}(x, D)]}{\partial x_j} &= \frac{\bar{c}_j}{n} + \frac{\bar{r}_j}{n}(\lambda\beta - 1)P[D_j > x_j] - \frac{\bar{r}_j\lambda\beta}{n}\Big\{1 - \frac{1}{\beta}(1 - P[D_j > x_j])\Big\}. \\
&= \frac{1}{n}[\bar{c}_j + \bar{r}_j\lambda(1-\beta) + \bar{r}_j P[D_j > x_j](\lambda(\beta - 1) - 1)].
\end{aligned} \tag{4.27}$$

Equating the right-hand side to zero, I get

$$\hat{x}_j^{\mathrm{RA}} = \bar{F}_{D_j}^{-1}\left(\frac{\bar{c}_j + \bar{r}_j\lambda(1-\beta)}{\bar{r}_j(1 + \lambda(1-\beta))}\right). \tag{4.28}$$

Note that the solution in Case (ii) is a closed-form optimal solution, not an approximation. In addition, it does not depend on the number of products, n, nor any demand dependence structure. Clearly, if $\lambda = 0$, $\hat{x}_j^{\mathrm{RA}} = \hat{x}_j^{\mathrm{RN}}$. As λ increases, $\hat{x}_j^{\mathrm{RA}}$ is decreasing. For any $0 \le \lambda \le 1/\beta$, $\hat{x}_j^{\mathrm{RA}}$ is well-defined.

It should be emphasized that Case (i) is more important, because for large n the distribution of Z_x^n is close to normal and for a small β, the β-quantile of Z_x^n tends to be smaller than $\frac{1}{n}\sum_{j=1}^{n} \bar{r}_j x_j$, for the values of x of interest.

Consider the special case of identical products. With a slight abuse of notation, let $c_j = c$, $r_j = r$ and $s_j = s$ for all $j = 1, 2, \dots, n$. In Case (i), the first-order approximation of the risk-averse solution yields:

$$\frac{d\rho[\bar{\Pi}(x, D)]}{dx} \simeq \bar{c} + \bar{r}P[D_1 > x](\delta_n^{\mathrm{RN}}\lambda - 1).$$

with $\delta_n^{\mathrm{RN}} = \frac{e^{-z_\beta^2/2}}{\sqrt{2\pi}} \frac{\hat{x}^{\mathrm{RN}} - \mu_x^{\mathrm{RN}}}{\sqrt{n-1}\sigma_x^{\mathrm{RN}}}$ where $\hat{x}^{\mathrm{RN}}$, μ_x^{RN} and σ_x^{RN} are the counterparts of $\hat{x}_j^{\mathrm{RN}}$, μ_j^{RN} and σ_j^{RN} in the case of identical products, respectively. Equating the right-hand side to zero, I obtain

$$\boxed{\hat{x}_j^{\mathrm{RA}} \simeq \bar{F}_{D_j}^{-1}\left(\frac{\bar{c}}{\bar{r}(1-\lambda\delta_n^{\mathrm{RN}})}\right), \quad j = 1, \ldots, n.} \tag{4.29}$$

The equation (4.29) is similar to the equation (4.26) except that the terms $\bar{c}_j$, $\bar{r}_j$ and $\delta_{nj}^{\mathrm{RN}}$ are now identical. In Case (ii), the equation (4.28) reduces to

$$\hat{x}^{\mathrm{RA}} = \bar{F}_{D_1}^{-1}\left(\frac{\bar{c} + \bar{r}\lambda(1-\beta)}{\bar{r}(1+\lambda(1-\beta))}\right). \tag{4.30}$$

4.4.3 General Law-Invariant Coherent Measures of Risk

So far my analysis has focused on a special risk measure, weighted mean-deviation from quantile, given in the equation (2.13). I now generalize the results to any law-invariant coherent measure of risk $\rho[\cdot]$.

Consider problem (4.10) where $\varkappa_{\mathscr{M}}[V]$ is given by (2.18). By Kusuoka theorem, for nonatomic spaces, every law-invariant coherent measure of risk has such representation. The equation (4.17) can be replaced by

$$\rho[\bar{\Pi}(x, D)] = \frac{1}{n}\sum_{j=1}^{n} \bar{c}_j x_j + \sup_{\mu \in \mathscr{M}} \int_0^1 \left(\mathbb{E}[Z_x^n](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}} \left\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x^n)_+]\right\}\right)\mu(d\beta).$$

Suppose the maximum over $\mathscr{M}$ is attained at a unique measure $\hat{\mu}$ (this is certainly true for *spectral* measures of risk, where the set $\mathscr{M}$ has just one element). Similarly to (4.25),

$$\frac{\partial\rho[\bar{\Pi}(x, D)]}{\partial x_j} \simeq \frac{\bar{c}_j}{n} + \frac{\bar{r}_j}{n}\left(-1 + \lambda\int_0^1 \delta_{nj}(\beta)\,\hat{\mu}(d\beta)\right) P[D_j > x_j]. \tag{4.31}$$

I denote here the quantity given in formula (4.24) by $\delta_{nj}^{\mathrm{RN}}(\beta)$, to stress its dependence on β. Let me approximate $\hat{\mu}$ by the measure $\hat{\mu}^{\mathrm{RN}}$, obtained for the risk-neutral solution $\hat{x}^{\mathrm{RN}}$.

Equating the approximate derivatives (4.31) to zero, I obtain an approximate solution:

$$\boxed{\hat{x}_j^{\mathrm{RA}} \simeq \bar{F}_{D_j}^{-1}\left(\frac{\bar{c}_j}{\bar{r}_j\Big(1 - \lambda \displaystyle\int_0^1 \delta_{nj}^{\mathrm{RN}}(\beta)\,\hat{\mu}^{\mathrm{RN}}(d\beta)\Big)}\right), \quad j = 1, 2, \ldots, n.} \tag{4.32}$$

Again, $\delta_{nj}^{\mathrm{RN}}(\beta) \downarrow 0$ as $n \to \infty$, and thus $\hat{x}_j^{\mathrm{RA}}$ increases in n and approaches the risk-neutral solution $\hat{x}_j^{\mathrm{RN}}$. This is consistent with Proposition 6.

4.4.4 Iterative Methods

So far, I have discussed the approximations based on expansions about the risk-neutral solution $\hat{x}^{\mathrm{RN}}$. But exactly the same argument can be used to develop an iterative method, in which the best approximation known so far is substituted for the risk-neutral solution. I explain the simplest idea for the approximation developed in §4.4.2; an identical extension can be done for Kusuoka measures discussed in §4.4.3.

The idea of the *iterative method* is to generate a sequence of approximations $\hat{x}^{(\nu)}$, $\nu = 0, 1, 2, \ldots$. We set $\hat{x}^{(0)} = \hat{x}^{\mathrm{RN}}$. Then I calculate $\hat{x}^{(1)}$ by applying formula (4.26). In the iteration $\nu = 1, 2, \ldots,$ I use $\hat{x}^{(\nu)}$ instead of $\hat{x}^{\mathrm{RN}}$ in my approximation, calculating:

$$\begin{aligned}
\mu_j^{(\nu)} &= \mathbb{E}[\min\{\hat{x}_j^{(\nu)}, D_j\}], \quad j = 1, \ldots, n, \\
\sigma_j^{(\nu)} &= \sqrt{\mathbb{V}\mathrm{ar}(\min\{\hat{x}_j^{(\nu)}, D_j\})}, \quad j = 1, \ldots, n, \\
\gamma_{nj}^{(\nu)} &= \sqrt{\frac{1}{n-1} \sum_{k \neq j} \bar{r}_k^2 \big(\sigma_k^{(\nu)}\big)^2}, \quad j = 1, \ldots, n, \\
\delta_{nj}^{(\nu)} &= \frac{e^{-z_\beta^2/2}}{\sqrt{2\pi}} \frac{\bar{r}_j(\hat{x}_j^{(\nu)} - \mu_j^{(\nu)})}{\sqrt{n-1}\,\gamma_{nj}^{(\nu)}}, \quad j = 1, \ldots, n.
\end{aligned}$$

Finally, formula (4.26) is applied to generate the next approximate solution $\hat{x}^{(\nu+1)}$, and the iteration continues.

The iterative method is efficient, if the initial approximation $\hat{x}^{(0)}$ is sufficiently close to the risk-averse solution. This is true, if the risk aversion coefficient $\kappa = \lambda\beta$ is close to zero, or when the number of products is very large. I must point out that my approximation given in the equation (4.26) may result in infeasible solution as the term

$\dfrac{\bar{c}_j}{\bar{r}_j(1-\delta_{nj}^{(v)}\lambda)}$ can be negative or greater than 1 (due to approximation). When this occurs less likely, one can say that the approximation is more stable. Generally, the approximation is more stable for larger number of products and smaller κ. For large κ and a moderate number of products, the risk-neutral solution may not be a good starting point for the iterative method. An alternative method is then the *continuation method*. In this approach, I apply the iterative method for a small value of κ, starting from the risk-neutral solution. Then I increase κ a little, and I apply the iterative method again, but starting from the best solution found for the previous value of κ. In this way, I gradually increase κ, until I recover optimal solutions for all values of the risk aversion coefficient which are of interest for me (usually, between 0 and 1).

4.5 Impact of Dependent Demands

In this section, I provide analytical insights on the impact of dependent demand in the multi-product risk-averse newsvendor model. Due to the significant analytical challenge, I focus on a two-product system and the mean–deviation from quantile measure of variability defined in (2.13).

Under the risk-neutral measure, dependence of product demands has no impact on the optimal order quantities. However, under risk-averse measures, it can greatly affect the optimal order decisions for the newsvendor. Intuitively, positively (negatively) dependent demand generates larger (smaller) variability and thus poses a larger (smaller) risk than independent demand. Thus, one tends to decrease (increase) the order quantity in case of positively (negatively) dependent demand relative to the case of independent demand.

To characterize the impact of demand dependence on the optimal order quantity under the coherent risk measure, I utilize the concept of "associated" random variables. Consider random variables $D_1, D_2, \ldots, D_n$, denote vector $D = (D_1, D_2, \ldots, D_n)$. The following definition is due to Esary, Proschan and Walkup (1967); see Tong (1980) for a review.

Definition 1. *The set of random variables* $\{D_1, D_2, \dots, D_n\}$ *is associated, or the random variables* $D_1, D_2, \dots, D_n$ *are associated, if*

$$Cov[f(D), g(D)] \geq 0, \tag{4.33}$$

or, equivalently,

$$\mathbb{E}[f(D)g(D)] \geq \mathbb{E}[f(D)]\mathbb{E}[g(D)]. \tag{4.34}$$

for all non-decreasing real functions f, g *for which* $\mathbb{E}[f(D)], \mathbb{E}[g(D)]$ *and* $\mathbb{E}[f(D)g(D)]$ *exist.*

Lemma 4. *Associated random variables have the following properties:*

(i) *Any subset of associated random variables is associated.*

(ii) *If two sets of associated random variables are independent of each other, their union is a set of associated random variables.*

(iii) *Non-decreasing (or non-increasing) functions of associated random variables are associated.*

(iv) *Let* $D_1, D_2, \dots, D_n$ *be associated random variables, then*

$$P\{D_1 \leq y_1, D_2 \leq y_2, \dots, D_n \leq y_n\} \geq \Pi_{k=1}^{n} P\{D_k \leq y_k\},$$

and

$$P\{D_1 \geq y_1, D_2 \geq y_2, \dots, D_n \geq y_n\} \geq \Pi_{k=1}^{n} P\{D_k \geq y_k\},$$

for all $(y_1, y_2, \dots, y_n) \in R^n$.

One refers to Tong (1980) for proofs. Intuitively, Part (iv) of Lemma 4 means that associated random variables are dependent in such a way that they tend to "hang on" together.

Association is closely related to correlation. By Tong (1980, pg. 99), a set of multi-variate normal random variables is associated if their correlation matrix has the structure l (Tong 1980, pg. 13) in which the correlation coefficient $\rho_{ij} = \gamma_i \gamma_j$ for all

$i \neq j$ and $0 \leq \gamma_i < 1$ for all i. This means that I can represent the demands as having one common factor:

$$D_i = \gamma_i D_0 + \Delta_i, \quad i = 1, \ldots, n,$$

where the factor D_0 and Δ_i, $i = 1, \ldots, n$, are independent. A special case is the bivariate normal random variable with a positive correlation coefficient.

Consider a two-identical product system and its solution with equal coordinates. Let $Z_x = \min\{x, D_1\} + \min\{x, D_2\}$. Clearly, $\Pi(x, D) = -2\bar{c}x + \bar{r}Z_x$ and

$$\rho(\Pi(x, D)) = 2\bar{c}x + \bar{r}\rho(Z_x). \tag{4.35}$$

$$\rho(Z_x) = E(Z_x)(\lambda\beta - 1) - \lambda\beta \max_{\eta \in \mathbb{R}} \left\{ \eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)_+] \right\}. \tag{4.36}$$

Let $\hat{\eta}$ be the maximizer. If $\hat{\eta}$ is not an atom of the distribution of Z_x, similarly to Case (i) of Proposition 4, I obtain

$$\frac{d\rho(Z_x)}{dx} = \frac{dE[Z_x]}{dx}(\lambda\beta - 1) + \lambda\frac{d\mathbb{E}[(\hat{\eta} - Z_x)_+]}{dx},$$

where $\hat{\eta}$ is the β quantile of Z_x and $\hat{\eta} < 2x$. Because the first term depends only on the marginal distributions of the demands, I focus on the second term, which is affected by the dependence of D_1 and D_2. I have

$$\frac{d\mathbb{E}[(\hat{\eta} - Z_x)_+]}{dx} = -\sum_{j=1}^{2} P[\{Z_x < \hat{\eta}\} \cap \{D_j > x\}] = -2P[\min\{x, D_2\} < \hat{\eta} - x, D_1 > x]. \tag{4.37}$$

Consider three cases of (D_1, D_2). Without changing the marginal distribution of D_1 and D_2, we let (D_1, D_2) be associated random variables in case 1 where $\hat{\eta}_P$ denotes the β quantile of the corresponding Z_x; In case 2, (D_1, D_2) are independent with $\hat{\eta}_I$ as the β quantile of the Z_x; In case 3, $(D_1, -D_2)$ are associated random variables with $\hat{\eta}_N$ as the β quantile of the Z_x. I also let x_P^*, x_I^* and x_N^* be the optimal order quantity of problem (4.8) for case 1, 2 and 3 respectively.

Proposition 7. *If $\hat{\eta}_P \leq \hat{\eta}_I \leq \hat{\eta}_N < 2x$, then*

$$x_P^* \leq x_I^* \leq x_N^*. \tag{4.38}$$

That is, positively (negatively) dependent (D_1, D_2) *results in smaller (larger) optimal order quantity than independent* (D_1, D_2)*.*

Proof. I first consider associated (D_1, D_2). I have

$$P[\min\{x, D_2\} < \hat{\eta}_P - x, D_1 > x] = P[D_2 < \hat{\eta}_P - x, D_1 > x].$$
$$= P[D_1 > x] - P[D_2 \geq \hat{\eta}_P - x, D_1 > x] \leq P[D_1 > x] - P[D_2 \geq \hat{\eta}_P - x]P[D_1 > x].$$
$$= P[D_2 < \hat{\eta}_P - x]P[D_1 > x] \leq P[D_2 < \hat{\eta}_I - x]P[D_1 > x].$$

The first inequality follows by Lemma 4 part (iv). The second inequality follows by $\hat{\eta}_P \leq \hat{\eta}_I$. Note that the last term corresponds to independent (D_1, D_2). Thus, by the equation (4.37), associated (D_1, D_2) have the derivatives $d\rho(Z_x)/dx$ at least as large as independent (D_1, D_2), which implies that $x_P^* \leq x_I^*$.

I then consider associated $(D_1, -D_2)$. I obtain

$$P[D_2 < \hat{\eta}_N - x, D_1 > x] = P[-D_2 > -\hat{\eta}_N + x, D_1 > x] \geq P[-D_2 > -\hat{\eta}_N + x]P[D_1 > x].$$
$$= P[D_2 < \hat{\eta}_N - x]P[D_1 > x] \geq P[D_2 < \hat{\eta}_I - x]P[D_1 > x].$$

The first inequality follows by Lemma 4 part (iv). The second inequality follows by $\hat{\eta}_I \leq \hat{\eta}_N$. Note that the last term corresponds to independent (D_1, D_2). Thus, by the equation (4.37), associated $(D_1, -D_2)$ have the derivatives $d\rho(Z_x)/dx$ no larger than independent (D_1, D_2), which implies that $x_I^* \leq x_N^*$. □

The condition $\hat{\eta}_P \leq \hat{\eta}_I \leq \hat{\eta}_N$ holds when $Y_1 = \min\{x, D_1\}$ and $Y_2 = \min\{x, D_2\}$ follow bivariate normal distribution and $\beta \leq 0.5$. One can approximate the joint distribution of Y_1 and Y_2 very closely by bivariate normal when (D_1, D_2) follow bivariate normal and x is set to cover most of the demand, which is very likely in practice when the underage cost $r - c$ is much greater than the overage cost $c - s$.

4.6 Special Cases with Dependent Demands

For independent demands and identical products case under coherent measures of risk, risk aversion decreases optimal solutions monotonously and any risk-averse solution is at most equal to the risk-neutral solution. However, for dependent demands case, similar analysis is analytically challenging even if products are identical. In order to see if a risk-averse solution is equal to, greater than or smaller than, I focus on two-identical product systems with dependent demands in this section, which is followed by the numerical examples at §6.3.2.

For simplicity, I assume that the marginal distribution of each demand follows a uniform distribution from zero to ten with perfectly negative correlation. That is, the correlation between the demands of product 1 and 2 is -1 and the relation between the demands of product 1 and 2 is $D_1 + D_2 = 10$. Then I discuss three examples of two-identical product systems as follows:

$$\text{Case A: } r_1 = r_2 = 2, c_1 = c_2 = 1, s_1 = s_2 = 0.$$

$$\text{Case B: } r_1 = r_2 = 2, c_1 = c_2 = 0.8, s_1 = s_2 = 0.$$

$$\text{Case C: } r_1 = r_2 = 2, c_1 = c_2 = 1.2, s_1 = s_2 = 0.$$

Again the case (i) solution is more important where the distribution of Z_x does not have an atom at the maximizer $\hat{\eta}$. I study the case (ii) solution alternatively only if the analysis of the case (i) solution is not enough to describe each example. Then,

$$\frac{d\rho(\Pi(x, D))}{dx} = 2\bar{c} - 2\bar{r}\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 2\bar{r}\lambda\mathbb{P}[\min\{x, D_2\} < \hat{\eta} - x, D_1 > x]. \quad (4.39)$$

1. Case A: $r_1 = r_2 = 2, c_1 = c_2 = 1, s_1 = s_2 = 0$.

   In case A, the risk-neutral solution is 5 which is the median of each marginal distribution. This is the special case when $\bar{r} = 2\bar{c}$. By applying the given parameters at Case A into the equation (4.39), I obtain

$$\frac{d\rho(\Pi(x, D))}{dx} = 2 - 4(1 - \lambda\beta)\mathbb{P}(D_1 > x) - 4\lambda\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x].$$

Thus, the key question is to compute $\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x]$ and, specifically, $\hat{\eta}$ which is the β-quantile of $Z_x = \min(x, D_1) + \min(x, D_2)$. Then I divide the range of x into the two subcases when $x \leq 5$ and $x > 5$.

- Case A-1: if $x \leq 5$

  Because $Z_x = \min(x, D_1) + \min(x, 10 - D_1)$, Z_x can be represented as a function of D_1, a random outcome.

$$Z_x = \begin{cases} D_1 + x, & \text{if } 0 \leq D_1 \leq x. \\ 2x, & \text{if } x \leq D_1 \leq 10 - x. \\ 10 + x - D_1, & \text{if } 10 - x \leq D_1 \leq 10. \end{cases}$$

  Then, I obtain

$$\hat{\eta} = \begin{cases} x + 5\beta, & \text{if } 0 < \beta \leq 2x/10. \\ 2x, & \text{if } \beta > 2x/10. \end{cases}$$

  When $\hat{\eta} = 2x$, $\hat{\eta}$ has an atom at this point and it is equivalent to $0 \leq x < 5\beta$. Thus, I focus on the case when $\hat{\eta} = x + 5\beta$ which is equivalent to $5\beta \leq x \leq 5$. When $5\beta \leq x \leq 5$,

$$\begin{aligned} \frac{d\rho(\Pi(x, D))}{dx} &= 2 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 4\lambda \cdot 1/2\beta \\ &= 2(1 - \lambda\beta)(1 - 2\mathbb{P}(D_1 > x)) \leq 0. \end{aligned}$$

  where the equality only holds when $x = 5$ which is equal to the risk-neutral solution and the median of the demand distribution.

- Case A-2: if $x > 5$

  Z_x can be also represented as a function of D_1, a random outcome.

$$Z_x = \begin{cases} D_1 + x, & \text{if } 0 \leq D_1 \leq 10 - x. \\ 10, & \text{if } 10 - x \leq D_1 \leq x. \\ 10 + x - D_1, & \text{if } x \leq D_1 \leq 10. \end{cases}$$

  Similar to case A-1, I obtain

$$\hat{\eta} = \begin{cases} x + 5\beta, & \text{if } 0 < \beta \leq 2(1 - x/10). \\ 10, & \text{if } \beta > 2(1 - x/10). \end{cases}$$

When $\hat{\eta} = 10$, $\hat{\eta}$ has an atom at this point and it is equivalent to $10 - 5\beta < x \leq 10$. Thus, I also focus on the case when $\hat{\eta} = x + 5\beta$ which is equivalent to $5 < x \leq 10 - 5\beta$. When $5 < x \leq 10 - 5\beta$,

$$\begin{aligned}\frac{d\rho(\Pi(x, D))}{dx} &= 2 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 4\lambda \cdot 1/2\beta \\ &= 2(1 - \lambda\beta)(1 - 2\mathbb{P}(D_1 > x)) > 0.\end{aligned}$$

In summary, $\frac{d\rho(\Pi(x,D))}{dx}$ is only equal to zero when $x = 5$ and negative value with $x < 5$ and positive value with $x > 5$. Therefore, $\hat{x}^{\mathrm{RA}} = 5$ is the unique risk-averse solution and it is exactly the same as the risk-neutral solution.

Next, I consider two more cases, $c_1 = c_2 = 0.8$ and $c_1 = c_2 = 1.2$ in Case B and C. All other conditions are the same as above. This means that when $c_1 = c_2 = 0.8$, $\bar{r} > 2\bar{c}$ and the risk-neutral solution is 6 which is higher than the median of the demand distribution. On the other hand, in case C, the risk-neutral solution is 4 and it is lower than the median of the demand distribution. Similar to Case A, I divide the range of x into two subcases when $x \leq 5$ and $x > 5$ ($x \geq 5$ and $x < 5$) in Case B (Case C).

2. Case B: $r_1 = r_2 = 2, c_1 = c_2 = 0.8, s_1 = s_2 = 0$.

- Case B-1: if $x \leq 5$

  Z_x can be represented as a function of D_1 as follows:

$$Z_x = \begin{cases} D_1 + x, & \text{if } 0 \leq D_1 \leq x. \\ 2x, & \text{if } x \leq D_1 \leq 10 - x. \\ 10 + x - D_1, & \text{if } 10 - x \leq D_1 \leq 10. \end{cases}$$

  This is the same as in the case A-1. Thus, when $\beta \leq 2x/10 \Leftrightarrow 5\beta \leq x \leq 5$, I obtain

$$\begin{aligned}&\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x] \\ =&\mathbb{P}[\min\{x, 10 - D_1\} < x + 5\beta - x, D_1 > x] = 1/2\beta.\end{aligned}$$

Because $\mathbb{P}(D_1 > x) \geq 0.5$, then

$$\begin{aligned}\frac{d\rho(\Pi(x, D))}{dx} &= 1.6 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 4\lambda \cdot 1/2\beta \\ &\leq 1.6 - 4 \cdot 0.5(1 - \lambda\beta) - 2\lambda\beta = -0.4 < 0.\end{aligned}$$

- Case B-2: if $x > 5$

  Z_x can be represented as a function of D_1 as follows:

$$Z_x = \begin{cases} D_1 + x, & \text{if } 0 \leq D_1 \leq 10 - x. \\ 10, & \text{if } 10 - x \leq D_1 \leq x. \\ 10 + x - D_1, & \text{if } x \leq D_1 \leq 10. \end{cases}$$

  This is also the same as in the case A-2. Thus, when $\hat{\eta} = x + 5\beta \Leftrightarrow \beta \leq 2(1 - x/10)$, then $5 < x \leq 10 - 5\beta$ and I obtain

$$\begin{aligned}&\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x] \\ =&\mathbb{P}[\min\{x, 10 - D_1\} < x + 5\beta - x, D_1 > x] = 1/2\beta.\end{aligned}$$

  Thus,

$$\begin{aligned}\frac{d\rho(\Pi(x, D))}{dx} &= 1.6 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 2\lambda\beta. \\ &= 1.6 - 4(1 - \lambda\beta)(1 - x/10) - 2\lambda\beta. \\ &= 0.4x(1 - \lambda\beta) + 2\lambda\beta - 2.4.\end{aligned}$$

  Equating the right-hand side, I obtain

$$\hat{x}^{\text{RA}} = \left(\frac{6 - 5\lambda\beta}{1 - \lambda\beta}\right) = 6 + \left(\frac{\lambda\beta}{1 - \lambda\beta}\right) > 6. \tag{4.40}$$

  Thus, this solution at (4.40) is clearly higher than the risk-neutral solution, but there is no guarantee that this solution belongs to $(5, 10 - 5\beta]$. When the solution at (4.40) is higher than $10 - 5\beta$, I consider the case where $\hat{\eta}$ has an atom at 10. Then I obtain

$$\begin{aligned}&\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x] \\ =&\mathbb{P}[\min\{x, 10 - D_1\} < 10 - x, D_1 > x] = \mathbb{P}(D_1 > 10 - x).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{d\rho(\Pi(x, D))}{dx} &= 1.6 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 4\lambda \cdot \mathbb{P}(D_1 > 10 - x)\\ &= 1.6 - 4(1 - x/10)(1 - \lambda\beta) - 4\lambda x/10\\ &= 0.4x(1 - \lambda\beta - \lambda) + 1.6 - 4(1 - \lambda\beta).\end{aligned}$$

Equating the right-hand side, I also obtain

$$\hat{x}^{\mathrm{RA}} = \left(\frac{6 - 10\lambda\beta}{1 - \lambda\beta - \lambda}\right). \tag{4.41}$$

It should be emphasized that the solution at (4.41) is also higher than the risk-neutral solution because this solution is always higher than $10-5\beta$ and $10 - 5\beta > 6$ with $\beta \in (0, 0.5]$.

Therefore, considering the solutions at (4.40) and (4.41), I obtain

$$\hat{x}^{\mathrm{RA}} = \begin{cases} \frac{6-5\lambda\beta}{1-\lambda\beta}, & \text{if } \frac{6-5\lambda\beta}{1-\lambda\beta} \leq 10 - 5\beta. \\ \frac{6-10\lambda\beta}{1-\lambda\beta-\lambda}, & \text{if } \frac{6-5\lambda\beta}{1-\lambda\beta} > 10 - 5\beta \ \& \ \frac{6-10\lambda\beta}{1-\lambda\beta-\lambda} < 10. \\ 10, & \text{if } \frac{6-5\lambda\beta}{1-\lambda\beta} > 10 - 5\beta \ \& \ \frac{6-10\lambda\beta}{1-\lambda\beta-\lambda} \geq 10. \end{cases}$$

Thus, in case B, the risk-averse solution is higher than the risk-neutral solution and can be located at $(6, 10]$ depending on the values of λ and β. Then the newsvendor profits under risk-neutrality and risk-aversion can be represented as follows:

$$\begin{aligned}&\Pi^{\mathrm{RN}}(x = \hat{x}^{\mathrm{RN}} = 6|D_1)\\ =&(2 - 0.8) \cdot 6 + (2 - 0.8) \cdot 6 - 2(6 - D_1)_+ - 2(6 - D_2)_+\\ =&1.2 \cdot 6 + 1.2 \cdot 6 - 2(6 - D_1)_+ - 2(6 - (10 - D_1))_+\\ =&14.4 - 2(6 - D_1)_+ - 2(-4 + D_1)_+.\\ &\Pi^{\mathrm{RA}}(x = \hat{x}^{\mathrm{RA}} = 6 + v|D_1)\\ =&(2 - 0.8) \cdot (6 + v) + (2 - 0.8) \cdot (6 + v) - 2(6 + v - D_1)_+ - 2(6 + v - D_2)_+\\ =&1.2 \cdot (6 + v) + 1.2 \cdot (6 + v) - 2(6 + v - D_1)_+ - 2(6 + v - (10 - D_1))_+\\ =&14.4 + 2.4v - 2(6 + v - D_1)_+ - 2(-4 + v + D_1)_+.\end{aligned}$$

where $0 \leq D_1 \leq 10$ and $0 < v < 4$ whatever solution I need to choose. Then,

$$\Pi^{\mathrm{RN}}(D_1) - \Pi^{\mathrm{RA}}(D_1) = \begin{cases} -0.4v, & \text{if } 0 \leq D_1 \leq 4 - v. \\ 1.6v + 2D_1 - 8, & \text{if } 4 - v \leq D_1 \leq 4. \\ 1.6v, & \text{if } 4 \leq D_1 \leq 6. \\ 12 + 1.6v - 2D_1, & \text{if } 6 \leq D_1 \leq 6 + v. \\ -0.4v, & \text{if } 6 + v \leq D_1 \leq 10. \end{cases}$$

Thus,

$$\begin{aligned} \mathbb{E}\left[\Pi^{\mathrm{RN}} - \Pi^{\mathrm{RA}}\right] = & \int_0^{4-v} -0.4v \cdot \frac{1}{10} dD_1 + \int_{4-v}^{4} (1.6v + 2D_1 - 8) \cdot \frac{1}{10} dD_1 \\ & + \int_4^6 1.6v \cdot \frac{1}{10} dD_1 + \int_6^{6+v} (12 + 1.6v - 2D_1) \cdot \frac{1}{10} dD_1 \\ & + \int_{6+v}^{10} -0.4v \cdot \frac{1}{10} dD_1 \\ & = \frac{1}{5} v^2 > 0 \end{aligned}$$

In summary, the profit under the risk-neutral solution does not outperform that under risk-averse solution in any event, even though the *expected* profit under the risk-neutral solution outperforms that under risk-averse solution (obviously) regardless of the value of $0 < v < 4$.

3. Case C: $r_1 = r_2 = 2, c_1 = c_2 = 1.2, s_1 = s_2 = 0$.

- Case C-1: if $x \geq 5$

  Here Z_x is defined the same in case B. If $\beta \leq 2(1 - x/10) \Leftrightarrow 5 \leq x \leq 10 - 5\beta$, then I obtain

$$\begin{aligned} & \frac{d\rho(\Pi(x, D))}{dx} \\ = & 2.4 - 4\mathbb{P}(D_1 > x)(1 - \lambda\beta) - 4\lambda\mathbb{P}[\min\{x, 10 - D_1\} < \hat{\eta} - x, D_1 > x] \\ = & 2.4 - 4(1 - x/10)(1 - \lambda\beta) - 2\lambda\beta > 0. \end{aligned}$$

- Case C-2: if $x < 5$

  If $\beta \leq 2x/10 \Leftrightarrow 5\beta \leq x < 5$, then

$$\begin{aligned}
&\frac{d\rho(\Pi(x,D))}{dx} \\
=&2.4 - 4\mathbb{P}(D_1 > x)(1-\lambda\beta) - 4\lambda\mathbb{P}[\min\{x, 10-D_1\} < \hat{\eta} - x, D_1 > x] \\
=&2.4 - 4(1-x/10)(1-\lambda\beta) - 4\lambda \cdot 1/2\beta \\
=&0.4x(1-\lambda\beta) - 1.6 + 2\lambda\beta.
\end{aligned}$$

  Equating the right-hand side, I obtain

$$\hat{x}^{\mathrm{RA}} = \left(\frac{4-5\lambda\beta}{1-\lambda\beta}\right) = 4 - \left(\frac{\lambda\beta}{1-\lambda\beta}\right) < 4. \tag{4.42}$$

  Similar to case B, this solution at (4.42) is clearly lower than the risk-neutral solution, but there is no guarantee that this solution belongs to $[5\beta, 5)$. When the solution (4.42) is lower than 5β, I consider the case when $\hat{\eta}$ has an atom at $2x$. Then I obtain

$$\begin{aligned}
&\mathbb{P}[\min\{x, 10-D_1\} < \hat{\eta} - x, D_1 > x] \\
=&\mathbb{P}[\min\{x, 10-D_1\} < 2x - x, D_1 > x] = \mathbb{P}(D_1 > 10 - x).
\end{aligned}$$

  Thus,

$$\begin{aligned}
\frac{d\rho(\Pi(x,D))}{dx} &= 2.4 - 4(1-x/10)(1-\lambda\beta) - 4\lambda x/10 \\
&= 0.4x(1-\lambda\beta-\lambda) + 4\lambda\beta - 1.6.
\end{aligned}$$

  Equating the right-hand side, I obtain

$$\hat{x}^{\mathrm{RA}} = \left(\frac{4-10\lambda\beta}{1-\lambda\beta-\lambda}\right). \tag{4.43}$$

  Similar to Case B-2, the solution at (4.43) is also lower than the risk-neutral solution because this solution is always higher than 5β and $5\beta < 4$ with $\beta \in (0, 0.5]$.

Therefore, considering the solutions at (4.42) and (4.43), I obtain

$$\hat{x}^{\mathrm{RA}} = \begin{cases} \frac{4-5\lambda\beta}{1-\lambda\beta}, & \text{if } \frac{4-5\lambda\beta}{1-\lambda\beta} \geq 5\beta. \\ \frac{4-10\lambda\beta}{1-\lambda\beta-\lambda}, & \text{if } \frac{4-5\lambda\beta}{1-\lambda\beta} < 5\beta \ \& \ \frac{4-10\lambda\beta}{1-\lambda\beta-\lambda} > 0. \\ 0, & \text{if } \frac{4-5\lambda\beta}{1-\lambda\beta} < 5\beta \ \& \ \frac{4-10\lambda\beta}{1-\lambda\beta-\lambda} \leq 0. \end{cases}$$

Thus, similar to Case B, the risk-averse solution is lower than the risk-neutral solution in Case C and can be located at $[0, 4)$ depending on the values of λ and β. Then the newsvendor profits under risk-neutrality and risk-aversion can be represented as follows:

$$\begin{aligned} & \Pi^{\mathrm{RN}}(x = \hat{x}^{\mathrm{RN}} = 4|D_1) \\ =& (2-1.2)\cdot 4 + (2-1.2)\cdot 4 - 2(4-D_1)_+ - 2(4-D_2)_+ \\ =& 0.8\cdot 4 + 0.8\cdot 4 - 2(4-D_1)_+ - 2(4-(10-D_1))_+ \\ =& 6.4 - 2(4-D_1)_+ - 2(-6+D_1)_+ \\ & \Pi^{\mathrm{RA}}(x = \hat{x}^{\mathrm{RA}} = 4-v|D_1) \\ =& (2-1.2)\cdot(4-v) + (2-1.2)\cdot(4-v) - 2(4-v-D_1)_+ - 2(4-v-D_2)_+ \\ =& 0.8\cdot(4-v) + 0.8\cdot(4-v) - 2(4-v-D_1)_+ - 2(4-v-(10-D_1))_+ \\ =& 6.4 - 1.6v - 2(4-v-D_1)_+ - 2(-6-v+D_1)_+ \end{aligned}$$

where $0 \leq D_1 \leq 10$ and $0 < v < 4$ in the case C. Then,

$$\Pi^{\mathrm{RN}}(D_1) - \Pi^{\mathrm{RA}}(D_1) = \begin{cases} -0.4v, & \text{if } 0 \leq D_1 \leq 4-v. \\ 1.6v + 2D_1 - 8, & \text{if } 4-v \leq D_1 \leq 4. \\ 1.6v, & \text{if } 4 \leq D_1 \leq 6. \\ 12 + 1.6v - 2D_1, & \text{if } 6 \leq D_1 \leq 6+v. \\ -0.4v, & \text{if } 6+v \leq D_1 \leq 10. \end{cases}$$

Thus, $\Pi^{\mathrm{RN}}(D_1) - \Pi^{\mathrm{RA}}(D_1)$ is exactly the same as in case B for all possible $0 \leq D_1 \leq 10$ and $\mathbb{E}\left[\Pi^{\mathrm{RN}} - \Pi^{\mathrm{RA}}\right] = \frac{1}{5}v^2 > 0$. Therefore, the same conclusion can be drawn here as Case B.

Chapter 5

Multi-Product Newsvendor Model - Exponential Utility Function

The work in this chapter was constructed by a working paper, ”A multi-product risk-averse newsvendor with exponential utility function” by S. Choi and A. Ruszczyński (2009).

In this chapter, I focus on the exponential utility function of the profit per product to model risk aversion in the multi-product newsvendor problem. Exponential utility function is a particular form of a nondecreasing and concave utility function. It is also the unique function to satisfy constant absolute risk aversion (CARA) property. By those reasons, exponential utility function has been used frequently in finance and also in the supply chain management literature such as Bouakiz and Sobel (1992) and Chen et al. (2007). Lastly, I should also point out that in most industrial examples companies have a large number of products. The exponential utility of the *total* profit becomes very flat when the number of products increases. From that perspective, it is more appropriate to consider average profit per product instead of total profit.

This paper contributes to literature in the following ways. In §5.2.1 I establish two basic analytical results for the model when the product demands are independent: the convexity of the model and monotonicity of the impact of risk aversion on the solution. I then consider the model with respect to the ratio of the degree of risk aversion to the number of products for independent demands case. When this ratio is sufficiently small but not zero, a closed-form approximation is obtained which is as easy to compute as the risk-neutral solution. I also show the asymptotic behaviors of the solution when this ratio converges to zero or infinity.

When product demands are dependent, the model becomes to be more complicated than with independent demands. Similar to the model under coherent measures of risk, but without the assumptions of identical products in two-product systems, I prove the same proposition of the impact of demand correlation under risk aversion for a system with any number of products. Then I study the impact of risk aversion with dependent demands. For a comparison purpose between risk-neutral and risk-averse solutions, I re-analyze in §5.4 the exactly same two-identical product cases of perfectly negative correlations discussed at §4.6. Then I show that when product demands have a perfectly negative correlation, risk-averse solutions can be equal to, higher than or lower than risk-neutral solutions depending on the degree of risk aversion. These analytical results can be also extended to general uniform distribution and other symmetric marginal distribution cases.

5.1 Problem Formulation

Again, given products $j = 1, \ldots, n$, let $x = (x_1, \ldots, x_n)$ be the vector of ordering quantities and let $D = (D_1, \ldots, D_n)$ be the demand vector. I also define $r = (r_1, \ldots, r_n)$ to be the price vector, $c = (c_1, \ldots, c_n)$ to be the cost vector, and $s = (s_1, \ldots, s_n)$ to be the vector of salvage values. Finally, let $f_{D_j}(\cdot)$ and $F_{D_j}(\cdot)$ be the marginal probability density function (pdf), if it exists, and the marginal cumulative distribution function (cdf) of D_j, respectively. Denote $\bar{F}_{D_j}(\xi) = 1 - F_{D_j}(\xi)$.

Setting $\bar{c}_j = c_j - s_j$ and $\bar{r}_j = r_j - s_j$, I can write the profit function as follows:

$$\bar{\Pi}(x, D) = \frac{1}{n} \sum_{j=1}^{n} \Pi_j(x_j, D_j).$$

where

$$\begin{aligned} \Pi_j(x_j, D_j) &= -\bar{c}_j x_j + \bar{r}_j \min\{x_j, D_j\} \\ &= (r_j - c_j)x_j - (r_j - s_j)(x_j - D_j)_+, \quad j = 1, \ldots, n. \end{aligned}$$

I assume that the demand vector D is random and nonnegative. Thus, for every $x \geq 0$ the profit $\Pi(x, D)$ is a real bounded random variable. Here these definitions of parameters, decision variables and profit functions are exactly same as those under coherent measures of risk.

The risk-neutral multi-product newsvendor optimization problem is to maximize the expected average profit per product:

$$\max_{x \geq 0} \mathbb{E}[\bar{\Pi}(x, D)]. \tag{5.1}$$

Again the optimal solution of the problem (5.1) is the same as given at the equation (4.4).

The exponential utility function of a profit $z \in \mathbb{R}$ is defined as follows:

$$u(z) = -e^{-\lambda z}.$$

It is nondecreasing and concave. Here, λ is a positive degree of risk aversion. The expected utility of a random profit Z is defined as follows:

$$U(Z) = \mathbb{E}\Big[-e^{-\lambda Z}\Big].$$

Setting $Z = \bar{\Pi}(x, D)$, I obtain the expected utility in the newsvendor problem,

$$U(\bar{\Pi}(x, D)) = \mathbb{E}\Big[-e^{-\lambda \bar{\Pi}(x,D)}\Big].$$

Thus, the problem to maximize the expected utility can be represented equivalently as follows:

$$\min_{x \geq 0} \mathbb{E}\Big[e^{-\lambda \bar{\Pi}(x,D)}\Big]. \tag{5.2}$$

Here it should be emphasized that the optimization model (5.2) cannot be defined when λ is equal to zero. Instead, I will show the asymptotic behavior of the solution when λ converges to zero in the following section.

5.2 Analytical Results for Independent Demands

5.2.1 Basic Analytical Results

In this subsection, I provide two analytical results for the multi-product newsvendor model under the exponential utility function. These results lay the theoretical foundation for the paper.

Proposition 8. $\mathbb{E}\left[e^{-\lambda\bar{\Pi}(x,D)}\right]$ *is a convex function of* x*.*

Proof. I first note that $\bar{\Pi}(x,D) = \frac{1}{n}\sum_{j=1}^{n}\Pi_j(x_j,D_j)$ is concave of x. Then, $-\lambda\bar{\Pi}(x,D)$ is a convex function of x. Because the function e^t is increasing and convex of t, the composition is convex as well. □

Lemma 5. *Assume that all products have independent demands. Let* $\hat{x}^{\mathrm{RA}_1}$ *be the solution of problem* (5.2) *for* $\lambda = \lambda_1 > 0$*. If* $\lambda_2 \geq \lambda_1$ *then there exists a solution* $\hat{x}^{\mathrm{RA}_2}$ *of problem* (5.2) *for* $\lambda = \lambda_2$ *such that* $\hat{x}_j^{\mathrm{RA}_2} \leq \hat{x}_j^{\mathrm{RA}_1}$*,* $j = 1,\ldots,n$*.*

Proof. Because all products have independent demands, the problem (5.2) is separable into each product. The result follows by Eeckhoudt, Gollier and Schlesinger (1995). □

5.2.2 Asymptotic Analysis and Closed-Form Approximations

In this subsection, I first consider the asymptotic results with respect to the ratio of the number of products to the degree of risk aversion $\left(\frac{\lambda}{n}\right)$. Then I develop a closed-form approximation to the optimal solution for a sufficiently low $\left(\frac{\lambda}{n}\right)$.

Due to the independence of demands, the problem is decomposable to each product:

$$\begin{aligned}
\mathbb{E}\left[e^{-\lambda\bar{\Pi}(x,D)}\right] &= \mathbb{E}\left[e^{-\lambda\cdot\frac{1}{n}\sum_{j=1}^{n}\Pi_j(x_j,D_j)}\right] \\
&= \mathbb{E}\left[e^{-\frac{\lambda}{n}\Pi_1(x_1,D_1)}\right]\cdot\mathbb{E}\left[e^{-\frac{\lambda}{n}\Pi_2(x_2,D_2)}\right]\cdots\mathbb{E}\left[e^{-\frac{\lambda}{n}\Pi_n(x_n,D_n)}\right].
\end{aligned}$$

It implies that one can optimize each $\mathbb{E}\Big[e^{-\frac{\lambda}{n}\Pi_j(x_j,D_j)}\Big]$ separately. Consider product $j = 1, 2, \ldots, n$. Interchanging the differentiation and expectation operations, I obtain

$$\begin{aligned}
&\frac{\partial}{\partial x_j}\left(\mathbb{E}\Big[e^{-\frac{\lambda}{n}\Pi_j}\Big]\right) = \mathbb{E}\Big[\frac{\partial}{\partial x_j}\left(e^{-\frac{\lambda}{n}\Pi_j}\right)\Big] = \mathbb{E}\Big[-\frac{\lambda}{n}e^{-\frac{\lambda}{n}\Pi_j}\cdot\frac{\partial \Pi_j}{\partial x_j}\Big] \\
&\quad = \mathbb{E}\Big[-\frac{\lambda}{n}e^{-\frac{\lambda}{n}\Pi_j}\cdot\Big(\{r_j - c_j\} - \{r_j - s_j\}\cdot\mathbb{1}_{\{D_j<x_j\}}\Big)\Big] \\
&\quad = -\frac{\lambda}{n}e^{-\frac{\lambda}{n}(r_j-c_j)x_j}\cdot\mathbb{E}\Big[e^{\frac{\lambda}{n}\{r_j-s_j\}\{x_j-D_j\}_+}\cdot\Big(\{r_j - c_j\} - \{r_j - s_j\}\cdot\mathbb{1}_{\{D_j<x_j\}}\Big)\Big].
\end{aligned}$$

The derivative is zero if the following equation is satisfied:

$$\mathbb{E}\Big[e^{\frac{\lambda}{n}\{r_j-s_j\}\{x_j-D_j\}_+}\cdot\Big(\{r_j - c_j\} - \{r_j - s_j\}\cdot\mathbb{1}_{\{D_j<x_j\}}\Big)\Big] = 0. \tag{5.3}$$

Then I consider three cases.

Case 1: The degree of risk aversion, λ, converges to zero or the number of products, n, goes to infinity.

The solution x_j^* of (5.3) satisfies the equation,

$$\int_0^{x_j^*} e^{\frac{\lambda}{n}(r_j-s_j)(x_j-\xi)}\cdot(c_j - s_j)f_j(\xi)d\xi = \int_{x_j^*}^{\infty}(r_j - c_j)f_j(\xi)d\xi = (r_j - c_j)\bar{F}_j(x_j^*). \tag{5.4}$$

If $\lambda/n \to 0$, then $e^{\frac{\lambda}{n}(r_j-s_j)(x_j-\xi)} \to 1$. It follows that the optimal solution x_j^* converges to the risk-neutral solution at (4.4). The economic implication is the same as the asymptotic behavior under coherent measures of risk when the number of products goes to infinity at §4.4.1.

Case 2: The degree of risk aversion to the number of products, λ, goes to infinity.

In this case, $e^{\frac{\lambda}{n}(r_j-s_j)(x_j-\xi)}$ goes to infinity. Then, the resulting optimal solution x_j^* by the equation (5.4) converges to be zero.

Case 3: The ratio of the degree of risk aversion to the number of products $\left(\frac{\lambda}{n}\right)$ is sufficiently small, but not zero.

In this case, I can use the following approximation:

$$e^{\frac{\lambda}{n}(r_j-s_j)(x_j-D_j)} \simeq 1 + \frac{\lambda}{n}(r_j - s_j)(x_j - D_j). \tag{5.5}$$

Then, substituting (5.5) into (5.4), I obtain:

$$(c_j - s_j)F_j(x_j) + \frac{\lambda}{n}(r_j - s_j)(c_j - s_j)\mathbb{E}[x_j - D_j]_+ = (r_j - c_j)\bar{F}_j(x_j). \tag{5.6}$$

The risk-averse solution $\hat{x}_j^{\text{RA}}$ is very close to $\hat{x}_j^{\text{RN}}$ by Case 1. Therefore, I can use the following first-order approximation:

$$\mathbb{E}\Big[\hat{x}_j^{\text{RA}} - D_j\Big]_+ \simeq \mathbb{E}\Big[\hat{x}_j^{\text{RN}} - D_j\Big]_+. \tag{5.7}$$

Substituting (5.7) into (5.6), I get the following closed-form approximation of the risk-averse solution:

$$\hat{x}_j^{\text{RA}} \simeq \bar{F}_j^{-1}\left(\frac{c_j - s_j}{r_j - s_j} + \frac{\lambda}{n} \cdot (c_j - s_j) \cdot \mathbb{E}\Big[\hat{x}_j^{\text{RN}} - D_j\Big]_+\right), \quad j = 1, \ldots, n. \tag{5.8}$$

From the analytical results obtained in this subsection, it can be said that the impacts of the number of products and the degree of risk aversion are combined as $\frac{\lambda}{n}$. More specifically, the n-product model with a degree of risk aversion λ is equivalent to the single-product counterpart with the new degree of risk aversion $\lambda_1 = \frac{\lambda}{n}$.

5.2.3 Iterative Methods

I cannot expect the approximation obtained in the equation (5.8) to be very accurate unless $\frac{\lambda}{n}$ is sufficiently small. In order to make the error rates smaller, the idea of the iterative method is to generate a sequence of approximations $\hat{x}_j^{(\nu)}, \nu = 0, 1, 2, \ldots$ At first, $\hat{x}_j^{(0)} = \hat{x}_j^{\text{RN}}$. Then I calculate $\hat{x}_j^{(1)}$ by applying the equation (5.8). In the iteration $\nu = 1, 2, \ldots$, I use $\hat{x}_j^{(\nu-1)}$ instead of $\hat{x}_j^{\text{RN}}$ in our approximation, calculating

$$\hat{x}_j^{(\nu)} \simeq \bar{F}_j^{-1}\left(\frac{c_j - s_j}{r_j - s_j} + \frac{\lambda}{n} \cdot (c_j - s_j) \cdot \mathbb{E}\Big[\hat{x}_j^{(\nu-1)} - D_j\Big]_+\right), \quad j = 1, \ldots, n. \tag{5.9}$$

If $\frac{\lambda}{n}$ is sufficiently small, the iterative method is efficient because the initial approximation $\hat{x}^{(0)}$ is sufficiently close to the risk-averse solution. Otherwise, the risk-neutral solution may not be a good starting point for this method. Similar to the multi-product model under general coherent measures of risk, I apply for a straightforward iterative method starting from the risk-neutral solution and continuation method for exponential utility functions. Again I must point out that my approximation given at the equation (5.8) does not guarantee a feasible solution because the term $\frac{c_j - s_j}{r_j - s_j} + \frac{\lambda}{n} \cdot (c_j - s_j) \cdot \mathbb{E}\left[\hat{x}_j^{\mathrm{RN}} - D_j\right]_+$ might be negative or greater than 1 (due to approximation). Correspondingly to the case with coherent measures of risk, the approximation method is generally more stable for smaller values of $\frac{\lambda}{n}$. The same reasons and implications with coherent measures of risk can be also applied. In addition, another reason is that in the approximation with exponential utility function, I multiply $\mathbb{E}\left[\hat{x}_j^{\mathrm{RN}} - D_j\right]_+$ by $\frac{\lambda}{n}$, which makes the risk-neutral solution further from the risk-averse solution.

5.3 Impact of Dependent Demands

Under risk-averse performance measures, dependence of product demands can greatly affect the optimal solutions. The joint distribution function is not decomposable and it makes this problem analytically very challenging.

In this section, I provide analytical insights on the impact of dependent demand under exponential utility function and consider three cases of $(D_1, D_2, \ldots, D_n)$. Similar to the multi-product model under general coherent measures of risk, an intuitive and appealing property is that positively (or negatively) dependent demands generate larger (or smaller) variability and thus poses a larger (or smaller) risk than independent demands.

To characterize the impact of demand dependence on the optimal order quantity under the coherent risk measure, I also utilize the concept of "associated" random variables defined and explained at Definition 1 and Lemma (iv). Then I similarly categorize into the three cases as follows: In case 1, $(D_1, D_2, \ldots, D_n)$ are positively associated random variables and $\hat{x}_j^{P,\lambda}$ is the optimal solution for product $j = 1, 2, \ldots, n$ when the degree of risk aversion is λ; In case 2, $(D_1, D_2, \ldots, D_n)$ are independent and $\hat{x}_j^{I,\lambda}$ is the optimal solution for product $j = 1, 2, \ldots, n$ when the degree of risk aversion is λ; In case 3, $(D_1, D_2, \ldots, D_n)$ are negatively associated random variables and $\hat{x}_j^{N,\lambda}$ is the optimal solution for product $j = 1, 2, \ldots, n$ when the degree of risk aversion is λ;

Proposition 9. $\hat{x}_j^{P,\lambda} \leq \hat{x}_j^{I,\lambda} \leq \hat{x}_j^{N,\lambda}, \quad j = 1, 2, \ldots, n.$

Proof. Let's start considering positively associated random variables $(D_1, D_2, \ldots, D_n)$. Differentiating $\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]$ with respect to x_1, I obtain:

$$\begin{aligned}\frac{\partial}{\partial x_1}\left(\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]\right) &= \frac{\partial}{\partial x_1}\mathbb{E}\left[e^{-\frac{\lambda}{n}\Pi_1} \cdot e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right] \\ &= \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1} \cdot e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right].\end{aligned}$$

Then I have

$$e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)} = e^{-\frac{\lambda}{n}\sum_{j=2}^{n}((r_j-c_j)x_j-(r_j-s_j)(x_j-D_j)_+)}$$

and

$$\frac{\partial}{\partial x_1}\left(e^{-\frac{\lambda}{n}\Pi_1}\right) = \begin{cases}\frac{\lambda}{n}(c_1 - s_1)e^{-\frac{\lambda}{n}(r_1-c_1)x_1+\frac{\lambda}{n}(r_1-s_1)(x_1-D_1)}, & \text{if } D_1 < x_1, \\ -\frac{\lambda}{n}(r_1 - c_1)e^{-\frac{\lambda}{n}(r_1-c_1)x_1}, & \text{if } D_1 \geq x_1.\end{cases}$$

Thus, $e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}$ is a nonincreasing and positive function of $D_2, \ldots, D_n$. Also, $\frac{\partial}{\partial x_1}\left(e^{-\frac{\lambda}{n}\Pi_1}\right)$ is a nonincreasing and positive function of D_1 if $D_1 < x_1$. If $D_1 \geq x_1$, then $\frac{\partial}{\partial x_1}\left(e^{-\frac{\lambda}{n}\Pi_1}\right)$ does not depend on D_1. As $D_1, D_2, \ldots, D_n$ are positively associated, I obtain the inequality

$$\frac{\partial}{\partial x_1}\left(\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]\right) \geq \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1}\right] \cdot \mathbb{E}\left[e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right].$$

Thus,

$$\frac{\partial}{\partial x_1}\left(\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]\right)_{\left\{x_1=\hat{x}_j^{I,\lambda}\right\}} \geq \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1}\right]_{\left\{x_1=\hat{x}_j^{I,\lambda}\right\}} \cdot \mathbb{E}\left[e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right] = 0.$$

This implies that $\hat{x}_1^{P,\lambda} \leq \hat{x}_1^{I,\lambda}$. Similarly, $\hat{x}_j^{P,\lambda} \leq \hat{x}_j^{I,\lambda}, \quad j = 2, \ldots, n.$

Consider negatively associated random variables $(D_1, D_2, \cdots, D_n)$. Then, arguing as before, I obtain the inequality

$$\begin{aligned}\frac{\partial}{\partial x_1}\left(\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]\right) &= \frac{\partial}{\partial x_1}\mathbb{E}\left[e^{-\frac{\lambda}{n}\Pi_1}\cdot e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right]\\ &= \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1}\cdot e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right]\\ &\leq \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1}\right]\cdot\mathbb{E}\left[e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right].\end{aligned}$$

Thus, the inequality direction is reversed as follows:

$$\frac{\partial}{\partial x_1}\left(\mathbb{E}\left[e^{-\lambda\bar{\Pi}}\right]\right)_{\left\{x_1=\hat{x}_j^{I,\lambda}\right\}} \leq \mathbb{E}\left[\frac{\partial}{\partial x_1}e^{-\frac{\lambda}{n}\Pi_1}\right]_{\left\{x_1=\hat{x}_j^{I,\lambda}\right\}}\cdot\mathbb{E}\left[e^{-\frac{\lambda}{n}(\Pi_2+\cdots+\Pi_n)}\right]=0.$$

Consequently, $\hat{x}_1^{I,\lambda} \leq \hat{x}_1^{N,\lambda}$. Similarly, $\hat{x}_j^{I,\lambda} \leq \hat{x}_j^{N,\lambda}, \quad j = 2,\ldots,n.$ □

5.4 Special Cases with Dependent Demands

For independent product demands case under an exponential utility function, a risk-averse solution is smaller than a risk-neutral solutions by Lemma 5. As the degree of risk aversion increases, the risk-averse solution decreases monotonously and it converges to zero in the limit of the degree of risk aversion, infinity. It implies that, for independent demands case, any risk-averse solution under exponential utility function is located between zero and the corresponding risk-neutral solution. However, for dependent product demands case under an exponential utility function, similar analysis is analytically very challenging. Thus, I consider in this section the exactly same special examples of two-identical product systems analyzed at §4.6.

Let's recall Proposition 3. The same proposition can be also proved similarly under exponential utility functions as follows:

Proposition 10. Assume that all products are identical, i.e., prices, ordering costs and salvage values are the same across all products. Further, let the joint probability distribution of the demand be symmetric, that is, invariant with respect to permutations of the demand vector. Then, for the exponential utility function, one of optimal solutions of problem (5.2) is a vector with equal coordinates, $\hat{x}_1^{\mathrm{RA}} = \hat{x}_2^{\mathrm{RA}} = \cdots = \hat{x}_n^{\mathrm{RA}}$.

Proof. Let me consider an arbitrary order vector $x = (x_1, \ldots, x_n)$ and let P be an $n \times n$ permutation matrix. Then, the distribution of profit associated with Px is the same as that associated with x. There are $n!$ different permutations of x and let us denote them $x^1, \ldots, x^{n!}$. Consider the point

$$y = \frac{1}{n!} \sum_{i=1}^{n!} x^i.$$

It has all coordinates equal to the average of the coordinates x_j. As the joint probability distribution of $D_1, D_2, \ldots, D_n$ is symmetric, the distribution of $\bar{\Pi}(x^i)$ is the same for each i. By Proposition 8, I obtain

$$\mathbb{E}\Big[e^{-\lambda\bar{\Pi}(y,D)}\Big] \leq \frac{1}{n!} \sum_{i=1}^{n!} \mathbb{E}\Big[e^{-\lambda\bar{\Pi}(x^i,D)}\Big] = \mathbb{E}\Big[e^{-\lambda\bar{\Pi}(x,D)}\Big].$$

This means that for every plan x, the corresponding plan y with equal orders is at least as good. As an optimal plan exists, there is an optimal plan with equal orders. □

1. Case A: $r_1 = r_2 = 2, c_1 = c_2 = 1, s_1 = s_2 = 0$.

   In this case, the risk-neutral solutions for product 1 and 2 are both 5, which are the medians of each marginal demand distribution, respectively. Then,

   $$\mathbb{E}[e^{-\lambda\bar{\Pi}}] = \mathbb{E}\Big[e^{-\frac{\lambda}{2}(\Pi_1+\Pi_2)}\Big], \quad \text{where } \Pi_j = (\bar{r}_j - \bar{c}_j)x_j - \bar{r}_j(x_j - D_j)_+, \quad \text{with } \forall j = 1, 2.$$

   Thus,

   $$\begin{aligned} U[\Pi(x,D)] &= \mathbb{E}\Big[e^{-\lambda(\bar{r}-\bar{c})x + \frac{\lambda}{2}\bar{r}(x-D_1)_+ + \frac{\lambda}{2}\bar{r}(x-D_2)_+}\Big] = \mathbb{E}\Big[e^{-\lambda x + \lambda(x-D_1)_+ + \lambda(x-D_2)_+}\Big]. \\ &= \mathbb{E}\Big[e^{-\lambda x + \lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big]. \end{aligned}$$

   because the optimal solutions of product 1 and 2 are always same by Proposition 10. Then,

   $$\frac{dU[\Pi(x,D)]}{dx} = \lambda\mathbb{E}\Big[\left(-1 + \mathbb{1}_{\{D_1 \leq x\}} + \mathbb{1}_{\{D_1 > 10-x\}}\right) \cdot e^{-\lambda x + \lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big].$$

- Case A-1: if $x > 5$, then

$$\begin{aligned}
\frac{dU[\Pi(x,D)]}{dx} &= \lambda \mathbb{E}\Big[-e^{-\lambda x+\lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big] \\
&+ \lambda \mathbb{E}\Big[\mathbb{1}_{\{D_1 \le x\}} \cdot e^{-\lambda x+\lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big] \\
&+ \lambda \mathbb{E}\Big[\mathbb{1}_{\{D_1 > 10-x\}} \cdot e^{-\lambda x+\lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big]. \\
&= \lambda \int_0^{10-x} -e^{-\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} -e^{-\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{x}^{10} -e^{-\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_0^{10-x} e^{-\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} e^{-\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} e^{-\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{x}^{10} e^{-\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1. \\
&= \lambda \int_{10-x}^{x} e^{-\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 > 0.
\end{aligned}$$

- Case A-2: if $x \le 5$, then

$$\begin{aligned}
\frac{dU[\Pi(x,D)]}{dx} &= \lambda \int_0^{x} -e^{-\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{x}^{10-x} -e^{-\lambda x} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{10} -e^{-\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_0^{x} e^{-\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{10} e^{-\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&= \lambda \int_{x}^{10-x} -e^{-\lambda x} \cdot \frac{1}{10} dD_1 \le 0.
\end{aligned}$$

Therefore, $\hat{x}_1^{\mathrm{RA}} = \hat{x}_2^{\mathrm{RA}} = \hat{x}_1^{\mathrm{RN}} = \hat{x}_2^{\mathrm{RN}} = 5$.

2. Case B: $r_1 = r_2 = 2, c_1 = c_2 = 0.8, s_1 = s_2 = 0$ with all the same other conditions.

In this case, the risk-neutral solutions for product 1 and 2 are both 6, which are $\frac{2-0.8}{2} = 0.6$-quantile of each marginal demand distribution, respectively. Then,

- Case B-1: if $x \le 5$, then

$$U[\Pi(x, D)] = \mathbb{E}\Big[e^{-1.2\lambda x + \lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big].$$

Then,

$$\begin{aligned}
&\frac{dU[\Pi(x, D)]}{dx} \\
=&\lambda\mathbb{E}\Big[\,(-1.2 + \mathbb{1}_{\{D_1 \le x\}} + \mathbb{1}_{\{D_1 > 10-x\}}) \cdot e^{-1.2\lambda x + \lambda(x-D_1)_+ + \lambda(x-10+D_1)_+}\Big].\\
=&\lambda \int_0^x -1.2e^{-1.2\lambda x + \lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
+&\lambda \int_x^{10-x} -1.2e^{-1.2\lambda x} \cdot \frac{1}{10} dD_1\\
+&\lambda \int_{10-x}^{10} -1.2e^{-1.2\lambda x + \lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1\\
+&\lambda \int_0^x e^{-1.2\lambda x + \lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
+&\lambda \int_{10-x}^{10} e^{-1.2\lambda x + \lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1.\\
=&-0.2\lambda \int_0^x e^{-1.2\lambda x + \lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
&-0.2\lambda \int_{10-x}^{10} e^{-1.2\lambda x + \lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1\\
&-1.2\lambda \int_x^{10-x} e^{-1.2\lambda x} \cdot \frac{1}{10} dD_1 < 0.
\end{aligned}$$

- Case B-2: if $x > 5$, then

$$
\begin{aligned}
\frac{dU[\Pi(x,D)]}{dx} = & \lambda \int_0^{10-x} -1.2e^{-1.2\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_{10-x}^{x} -1.2e^{-1.2\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_{x}^{10} -1.2e^{-1.2\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_0^{10-x} e^{-1.2\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_{10-x}^{x} e^{-1.2\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_{10-x}^{x} e^{-1.2\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
& + \lambda \int_{x}^{10} e^{-1.2\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1. \\
= & - 0.2\lambda \int_0^{10-x} e^{-1.2\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
& - 0.2\lambda \int_{x}^{10} e^{-1.2\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
& + 0.8\lambda \int_{10-x}^{x} e^{-1.2\lambda x+\lambda(x-D_1)_+ +\lambda(x-10+D_1)_+} \cdot \frac{1}{10} dD_1. \\
& = \frac{0.2}{10} e^{0.8\lambda x-10\lambda} - \frac{0.2}{10} e^{-0.2\lambda x} - \frac{0.2}{10} e^{-0.2\lambda x} + \frac{0.2}{10} e^{0.8\lambda x-10\lambda} \\
& + \frac{0.8}{10} e^{0.8\lambda x-10\lambda} \cdot (2x-10). \\
& = 0.04e^{-0.2\lambda x}\left((4x-19)e^{\lambda x-10\lambda} - 1\right).
\end{aligned}
$$

Thus, I cannot solve the risk-averse solution in a closed-form. Instead, I will check if the risk-averse solution is higher or lower than the risk-neutral solution.

$$
\frac{dU[\Pi(x,D)]}{dx}\left(x = \hat{x}^{\mathrm{RN}} = 6\right) = 0.04e^{-1.2\lambda}\left(5e^{-4\lambda} - 1\right).
$$

Therefore,

$$
\hat{x}^{\mathrm{RA}} = \begin{cases} 5 < \hat{x}^{\mathrm{RA}} \le \hat{x}^{\mathrm{RN}} = 6, & \text{if } \lambda \le \frac{1}{4}\log 5. \\ 6 = \hat{x}^{\mathrm{RN}} < \hat{x}^{\mathrm{RA}} \le 10, & \text{if } \lambda > \frac{1}{4}\log 5. \end{cases}
$$

Thus, depending on the value of λ, the risk-averse solution may be equal to, higher than or lower than the risk-neutral solution.

3. Case C: $r_1 = r_2 = 2, c_1 = c_2 = 1.2, s_1 = s_2 = 0$ with all the same other conditions. In this case, the risk-neutral solutions for product 1 and 2 are both 4, which are $\frac{2-1.2}{2} = 0.4$-quantile of each marginal demand distribution, respectively. Then,

   - Case C-1: if $x \geq 5$, then

$$U[\Pi(x,D)] = \mathbb{E}\Big[e^{-0.8\lambda x+\lambda(x-D_1)_+ +\lambda(x-10+D_1)_+}\Big].$$

   Then,

$$\begin{aligned}
&\frac{dU[\Pi(x,D)]}{dx} \\
&= \lambda\mathbb{E}\Big[\left(-0.8 + \mathbb{1}_{\{D_1 \leq x\}} + \mathbb{1}_{\{D_1 > 10-x\}}\right) \cdot e^{-0.8\lambda x+\lambda(x-D_1)_+ +\lambda(x-10+D_1)_+}\Big]. \\
&= \lambda \int_0^{10-x} -0.8e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} -0.8e^{-0.8\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{x}^{10} -0.8e^{-0.8\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_0^{10-x} e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} e^{-0.8\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{10-x}^{x} e^{-0.8\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ \lambda \int_{x}^{10} e^{-0.8\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1. \\
&= 0.2\lambda \int_0^{10-x} e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1 \\
&+ 0.2\lambda \int_{x}^{10} e^{-0.8\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 \\
&+ 1.2\lambda \int_{10-x}^{x} e^{-0.8\lambda x+\lambda(x-D_1)+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1 > 0.
\end{aligned}$$

- Case C-2: if $x < 5$, then

$$\begin{aligned}
&\frac{dU[\Pi(x,D)]}{dx}\\
&= \lambda \int_0^x -0.8e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
&+ \lambda \int_{10-x}^x -0.8e^{-0.8\lambda x} \cdot \frac{1}{10} dD_1\\
&+ \lambda \int_{10-x}^{10} -0.8e^{-0.8\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1\\
&+ \lambda \int_0^x e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
&+ \lambda \int_{10-x}^{10} 0.2e^{-0.8\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1.\\
&= 0.2\lambda \int_0^x e^{-0.8\lambda x+\lambda(x-D_1)} \cdot \frac{1}{10} dD_1\\
&+ 0.2\lambda \int_{10-x}^{10} e^{-0.8\lambda x+\lambda(x-10+D_1)} \cdot \frac{1}{10} dD_1\\
&- 0.8\lambda \int_x^{10-x} e^{-0.8\lambda x} \cdot \frac{1}{10} dD_1.\\
&= -2 \cdot \frac{0.2}{10} e^{-0.8\lambda x} + 2 \cdot \frac{0.2}{10} e^{0.2\lambda x} - \frac{0.8}{10} e^{-0.8\lambda x} \cdot (10 - 2x).\\
&= 0.04e^{-0.8\lambda x} \left(e^{\lambda x} + (4x - 21)\right).
\end{aligned}$$

Thus, a risk-averse solution cannot be solved in a closed-form similarly done in Case B. Instead, I will check if the risk-averse solution is higher or lower than the risk-neutral solution.

$$\frac{dU[\Pi(x,D)]}{dx} \left(x = \hat{x}^{\mathrm{RN}} = 4\right) = 0.04e^{-3.2\lambda} \left(e^{4\lambda} - 5\right).$$

Therefore,

$$\hat{x}^{\mathrm{RA}} = \begin{cases} 4 = \hat{x}^{\mathrm{RN}} < \hat{x}^{\mathrm{RA}} < 5, & \text{if } \lambda < \frac{1}{4} \log 5. \\ 0 \le \hat{x}^{\mathrm{RA}} \le \hat{x}^{\mathrm{RN}} = 4, & \text{if } \lambda \ge \frac{1}{4} \log 5. \end{cases}$$

Similarly, depending on the value of λ, the risk-averse solution may be also equal to, higher than or lower than the risk-neutral solution.

Chapter 6
Computational Study

6.1 Sample-based Optimization

Sample-based optimization has been applied to solve large-scale stochastic programming problems using Monte Carlo sampling method and the theory and detailed procedure are summarized in Shapiro (2003 and 2008). However, the technique is developed mostly for expected value (risk-neutral) models and thus the usage has been limited mainly to expected value models.

Recently, Ruszczyński and Vanderbei (2003) propose solvable mean-risk models by applying Monte Carlo method. Thus, I follow their method in this chapter, which can be summarized as follows:

Step 1: Generating N uniform random numbers and transforming it properly to another random variable with a certain probability distribution;

Step 2: Calculating statistic such as expectation and measure of risk;

Step 3: Solving the particular optimization model;

In step 1, each sample is *iid* and randomly chosen scenario outside the optimization model and it has a probability of $1/N$. As the number of samples increases to infinity, each statistic from the samples converges to the true value with probability one by the Law of Large Numbers. Finally, uncertainty is removed in the optimization model and one can solve it deterministically. However, the resulting models are still large-scale and difficult to solve.

6.2 Numerical Analysis for Single-Product Models

The objective of this section is as follows: I consider the three risk measures at single-product models - mean-semideviation with the degree one, mean-deviation from quantile and mean-worst scaled average deviation from quantile. I solve the linear programming problems obtained by sample-based optimization and tabulate and illustrate the impact of the degree of risk aversion on optimal solutions. Then I confirm previous analytical results by numerical examples.

I randomly generate a large sample from the demand distribution: $D_k, k = 1, \ldots, N$. Then I solve the mean–risk model with the empirical distribution of D, by treating sampled values D_k as equally likely scenarios (with probabilities $p_k = 1/N$).

In the case of a semideviation $\sigma_1[\cdot]$, the mean-risk model (3.7) is equivalent to the following linear programming problem:

$$
\begin{aligned}
\max \quad & \mu - \lambda \textstyle\sum_{k=1}^{N} p_k r_k \\
\text{subject to} \quad & \mu = \textstyle\sum_{k=1}^{N} p_k \Pi_k, \\
& \Pi_k \le -\bar{c}x + \bar{r}D_k, \quad k = 1, \ldots, N, \\
& \Pi_k \le -\bar{c}x + \bar{r}x, \quad k = 1, \ldots, N, \\
& r_k \ge \mu - \Pi_k, \quad k = 1, \ldots, N, \\
& x \ge 0 \quad \text{and} \quad r_k \ge 0, \quad k = 1, \ldots, N.
\end{aligned}
$$

where Π_k and r_k are the profit generated and the risk calculated from k^{th} sample demand, $k = 1, \ldots, N$. The equivalent linear programming problem for the mean-deviation from a quantile is:

$$\begin{aligned} \max \quad & \textstyle\sum_{k=1}^{N} p_k \Pi_k - \lambda \sum_{k=1}^{N} p_k [(1-\beta)u_k + \beta v_k] \\ \text{subject to} \quad & \Pi_k = \eta + v_k - u_k, \quad k = 1, \ldots, N, \\ & \Pi_k \le -\bar{c}x + \bar{r}D_k, \quad k = 1, \ldots, N, \\ & \Pi_k \le -\bar{c}x + \bar{r}x, \quad k = 1, \ldots, N, \\ & x \ge 0, u_k \ge 0, v_k \ge 0, \quad k = 1, \ldots, N. \end{aligned}$$

The last model may also serve as a building block for the mean–risk model associated with a general law-invariant measure of risk, given by a convex set of probability measures $\mathscr{M}$ on $[0, 1]$ in (2.18). I approximate $\mathscr{M}$ by a convex set of probability measures Q supported on the points $\beta_i = i/n$, $i = 1, \ldots, N$. I obtain the semi-infinite linear programming problem:

$$\begin{aligned} \max \quad & \textstyle\sum_{k=1}^{N} p_k \Pi_k - \lambda \varkappa \\ \text{subject to} \quad & \varkappa \ge \textstyle\sum_{i=1}^{N} q_i \varkappa_i N/i, \quad \text{for all} \quad q \in Q, \\ & \varkappa_i \ge \textstyle\sum_{k=1}^{N} p_k [(1 - \frac{i}{N})u_{ik} + \frac{i}{N} v_{ik}], \quad i = 1, \ldots, N, \\ & \Pi_k = \eta_i + v_{ik} - u_{ik}, \quad i = 1, \ldots, N, \quad k = 1, \ldots, N, \\ & \Pi_k \le -\bar{c}x + \bar{r}D_k, \quad k = 1, \ldots, N, \\ & \Pi_k \le -\bar{c}x + \bar{r}x, \quad k = 1, \ldots, N, \\ & x \ge 0 \quad \text{and} \quad u_{ik} \ge 0, v_{ik} \ge 0, \quad i = 1, \ldots, N, \quad k = 1, \ldots, N. \end{aligned}$$

If the set Q is a polyhedron, it is sufficient to satisfy the semi-infinite constraint only for the vertices $\hat{q}^j$ of Q. For example, in the mean–risk model with the risk functional (2.19), I can calculate $i_{\min} = \lfloor \beta_{\min} N \rfloor$, $i_{\max} = \lceil \beta_{\max} N \rceil$ (the round-down and the roundup) and consider only atomic measures $\hat{q}^j$, $j = i_{\min}, \ldots, i_{\max}$, with $\hat{q}^j_i = 1$, if $i = j$, and $\hat{q}^j_i = 0$, if $i \neq j$. The semi-infinite constraint becomes:

$$\varkappa \ge \varkappa_i N/i, \quad i = i_{\min}, \ldots, i_{\max}.$$

	Semideviation			Deviation from Median			Worst Scaled Deviation		
λ	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$
0	63.22	157.70	74.41	63.22	157.70	76.46	63.04	157.70	346.63
0.2	58.66	157.15	68.40	58.93	157.22	76.13	53.68	154.56	298.38
0.4	55.55	155.78	64.08	56.46	156.21	73.21	38.62	134.24	232.89
0.6	52.09	153.27	59.07	53.14	154.14	68.96	6.10	28.84	29.91
0.8	48.38	149.45	53.63	48.24	149.29	62.05	3.37	16.33	10.29
1.0	44.25	143.95	47.50	42.09	140.56	52.39	2.73	13.29	6.82

Table 6.1: Solutions for Different Levels of Risk Aversion with Uniform Demand Distribution.

It should be stressed that the representation of all law-invariant coherent measures of risk by Kusuoka theorem holds true in nonatomic spaces, in general. In a space with atoms, the mean–risk model with the risk term given by (2.18) still defines a coherent measure of risk, but it may not be possible to construct any law-invariant coherent measure of risk this way. Also, the issues of convergence of discrete approximations call for precise analysis.

In order to illustrate the results of this note, I consider the problem with $r = 15, c = 10, s = 7$. I consider two distributions of the demand: uniform in $[0, 100]$ and lognormal. The parameters of the lognormal distribution are chosen such that the mean and variance of both distributions are same. The risk-neutral solution $\hat{x}^{\mathrm{RN}}$ equals 62.5 (51.37) for the uniform (lognormal) distribution. Therefore, there are six numerical examples by the combination of three risk measures and two demand distributions. In mean-deviation from quantile models, I use $\beta = 0.5$. For each of the six models, $\lambda = 0, 0.2, 0.4, 0.6, 0.8, 1$.

To find risk-averse solutions, I generate a sample of size $N = 1000$ and solve the resulting linear programming models by the way described in section 6.1. The results are summarized in Tables 6.1 and 6.2, and illustrated in Figure 6.1. My numerical solutions monotonously decrease as λ increases through the whole range of λ with different demand distributions and risk measures, which confirms Proposition 1 in single-product models.

	Semideviation			Deviation from Median			Worst Scaled Deviation		
λ	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$	$\hat{x}^{\mathrm{RA}}$	$\mathbb{E}[\Pi(\hat{x}^{\mathrm{RA}})]$	$r[\Pi(\hat{x}^{\mathrm{RA}})]$
0	50.86	163.39	42.14	50.86	163.39	40.12	50.86	163.39	267.25
0.2	48.15	163.00	38.48	47.70	162.86	36.33	45.01	161.44	181.09
0.4	46.04	162.12	35.52	45.01	161.44	31.59	36.28	149.75	143.21
0.6	44.03	160.64	32.58	41.67	158.21	25.07	18.54	91.02	31.26
0.8	41.41	157.91	28.69	39.79	155.76	21.58	16.50	81.63	16.71
1.0	39.49	155.33	25.81	38.17	153.24	18.80	15.84	78.50	13.33

Table 6.2: Solutions for Different Levels of Risk Aversion with Lognormal Demand Distribution.

Figure 6.1 shows the empirical cumulative distribution function for the net profit with different λ, at the optimal solutions of the problem with the lognormal distribution. The most distinguishing features are that there is a jump of each curve (corresponding to the mass probability at $\hat{x}^{\mathrm{RA}}$) and that the curves with larger λ have a thinner left tail, which reflects risk aversion.

6.3 Numerical Analysis for the Multi-Product Model under Coherent Measures of Risk

The objective of this section is two-fold. First, I study the accuracy and convergence rates of the approximations. Second, I provide insights (in addition to the analyzes in §§4.3-4.5) on the impact of demand dependence and risk aversion.

In all examples considered I apply sample-based optimization to solve the resulting stochastic programming problems. I generate a sample $D^1, D^2, \ldots, D^T$ of the demand vector, where

$$D^t = (d_{1t}, d_{2t}, \ldots, d_{nt}), \quad t = 1, \ldots, T.$$

Then I replace the original demand distribution by the empirical distribution based on the sample, that is, I assign to each of the sample points the probability $p_t = 1/T$. It is known that when $T \to \infty$, the optimal value of the sample problem approaches the optimal value of the original problem (see Shapiro (2008)). In all our examples I used

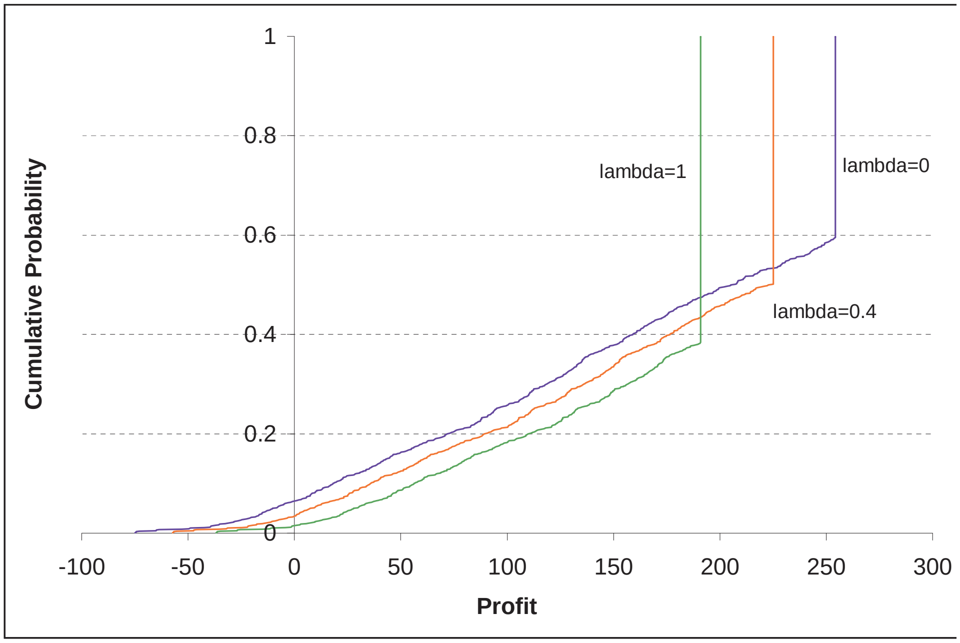

Figure 6.1: Cumulative Distribution Functions of Profit for Different Levels of Risk Aversion in the Problem with Lognormal Distribution. Mean-Deviation from Median is used as the Risk Functional.

$T = 10,000$.

For the empirical distribution, the corresponding optimization problem (4.8) has an equivalent linear programming formulation. For each $j = 1, \ldots, n$ and $t = 1, \ldots, T$ we introduce the variable w_{jt} to represent the salvaged number of product j in scenario t. The variable u_t represents the shortfall of the profit in scenario t to the quantile η. It is also convenient to introduce the parameter $\kappa = \lambda\beta$ to represent the relative risk aversion

$(0 \le \kappa \le 1)$. I obtain the formulation

$$\begin{aligned}
\max \quad & (1-\kappa)\sum_{j=1}^{n}\Big[(r_j-c_j)x_j-(r_j-s_j)\sum_{t=1}^{T}p_t w_{jt}\Big]+\kappa\Big(\eta-\frac{1}{\beta}\sum_{t=1}^{T}p_t u_t\Big) \qquad (6.1)\\
\text{subject to} \quad & \sum_{j=1}^{n}\Big[(r_j-c_j)x_j-(r_j-s_j)w_{jt}\Big]+u_t\ge\eta, \quad t=1,\ldots,T,\\
& x_j-d_{jt}\le w_{jt}, \quad j=1,\ldots,n; \quad t=1,\ldots,T,\\
& w_{jt}\ge 0, \quad j=1,\ldots,n; \quad t=1,\ldots,T,\\
& u_t\ge 0, \quad t=1,\ldots,T,\\
& x_j\ge 0, \quad j=1,\ldots,n.
\end{aligned}$$

Indeed, suppose the order quantities x_j are fixed. Then $w_{jt} = (x_j - d_{jt})_+$ and $u_t = (\eta - \Pi(x, D^t))_+$ are optimal, and I maximize with respect to η the last term in (6.2), that is,

$$\max_{\eta}\left\{\eta-\frac{1}{\beta}\mathbb{E}[(\eta-\Pi(x,D))_+]\right\}=-\mathrm{AVaR}_{\beta}[\Pi(x,D)].$$

In the last expression I used (2.14). Therefore, (6.2) equals $(1-\kappa)\mathbb{E}[\Pi(x,D)] - \kappa\mathrm{AVaR}_{\beta}[\Pi(x,D)]$.

6.3.1 Accuracy of Approximations

In this subsection, I assess the accuracy of the closed-form approximations of §4.4.2. I first consider identical products, then non-identical products.

For identical products, I assume that all products have identical cost structure and identical and independent demands. I set $r = 15, c = 10, s = 7$. I also set the demand distribution of each product to be lognormal with $\mu = 3$ and $\sigma = 0.4724$ (to achieve the desirable coefficient of variance (CV) of 0.5). Thus, the mean and standard deviation of each demand are $e^{\mu+\sigma^2/2} = 22.46$ and $e^{\mu+\sigma^2/2}\cdot\sqrt{(e^{\sigma^2}-1)} = 11.23$. Because the joint demand distribution is invariant with respect to the permutations of the demand vector, there exists an order vector with equal coordinates, which is optimal for the model.

I use $\beta = 0.5$, that is, I am concerned with the expected shortfall below the median as the risk measure. Then I choose the number of products, n, to be 1, 3, 10 and 30, and I study the impact of the number of products on the gap between the sample-based LP solutions and the approximate solutions (generated by the iterative method with $\nu = 3$, see §4.4.4). While the sample-based LP solutions can take hours to solve, especially for large n and T (by CPLEX 9.0 at Intel Pentium IV PC with CPU 2.0 GHz and 1 GB RAM), the approximate solution can be obtained instantly, which is 1 or 2 seconds when $n = 30$.

Number of Products	Identical Products	Heterogenous Products
3	232	337
10	3572	9006
30	32607	50889

Table 6.3: Comparison of CPU Running Time between Identical and Heterogenous Products Model with respect to Number of Products – Time unit: second.

In my numerical study, the optimal order quantities of different products are very close to each other but not necessarily identical due to the random sample error, which confirms Proposition 3. Thus, I obtain the numerical solution by taking the average of all products' numerical solutions. The corresponding results are illustrated in Figure 6.2, where on the horizontal axis I display the degree of relative risk aversion parameter $\kappa = \lambda\beta$. The term "exact", "numerical" and "approximation" represent the solution obtained by exact calculation, sample-based LP, and a closed-form approximation by the continuation method, respectively.

Figure 6.2 shows that my analytical solution is very close to the numerical solution when $n = 1$. This is obvious, because the solution is exact for the single-product case (in this problem the case $\hat{\eta} = \bar{r}x$ is valid). In the case of 3 products, the approximation does not work well, which is quite understandable as the approximation is based on the Central Limit Theorem. As the number of products further increases, my approximations become more accurate and the gap becomes negligible when $n \geq 10$. The

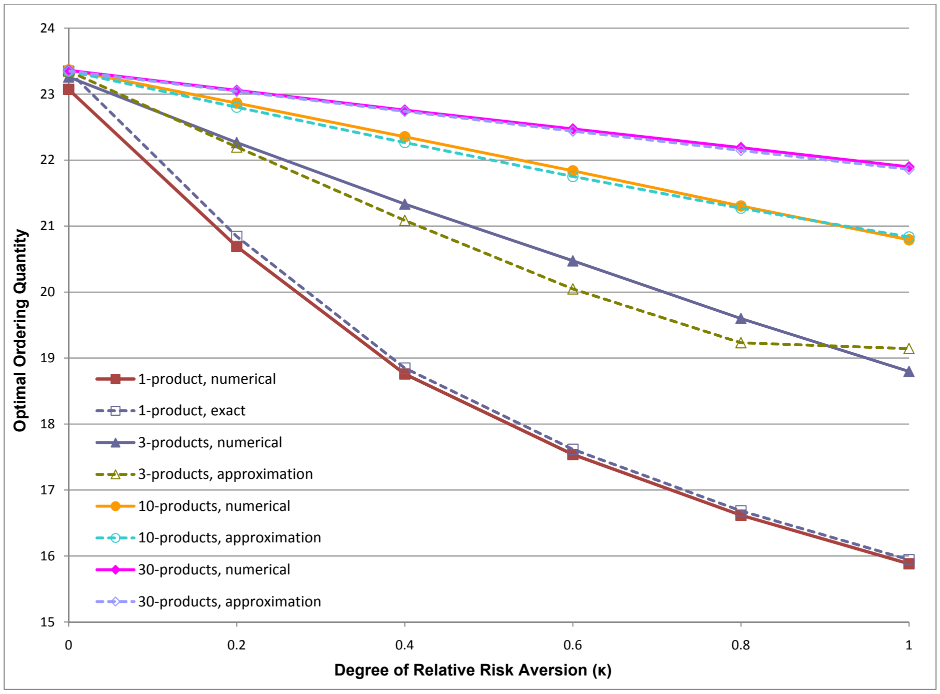

Figure 6.2: Identical Products with Independent Demands – Approximate or exact solution vs. sample-based LP solutions.

numerical study also shows that the order quantities should decrease as the degree of risk-aversion increases, which confirms Proposition 4. Also, I notice that as the number of products increases, the error of the risk neutral solution decreases.

For independent but heterogenous products, I tested the accuracy of the approximations on 30 randomly generated problems, 10 for each number of products $n = 3, 10, 30$. For each problem, I calculated the sample-based LP solution and an approximate solution by the continuation method (see §4.4.4). At each value of $\kappa = 0.2, 0.4, 0.6, 0.8, 1$, I made just one step of the iterative method.

Table 6.4 shows that the continuation method is much more stable and accurate than the straightforward approximation starting from the risk-neutral solution (the iterative method with $\nu = 1$), especially for smaller numbers of products, when the difference between risk-neutral solution and risk-averse solution is larger (e.g., κ is larger). For $n = 30$ both methods work very well.

For each instance that the continuation method can generate a feasible solution,

λ	Number of Products					
	3		10		30	
	Straightforward	Continuation	Straightforward	Continuation	Straightforward	Continuation
0.2	1	1	0	0	0	0
0.4	2	1	0	0	0	0
0.6	5	2	0	0	0	0
0.8	10	2	1	0	0	0
1	10	3	5	0	0	0

Table 6.4: Comparison of the Degree of Stability between Straightforward and Continuation Methods with respect to Number of Products and Degree of Risk Aversion.

I compute the absolute percentage error of the approximate solution relative to the sample-based LP solution, which is defined by the absolute difference between the approximate solution and the sample-based LP solution over the sample-based LP solution. For comparison purposes, I also compute the absolute percentage error of the risk-neutral solution relative to the sample-based LP solution. Then for each value of n and κ, I compute the average and maximum percentage error over all the solutions generated. The average (and maximum) percentage errors of the risk-neutral solutions and of the solutions obtained by the continuation method are displayed in Figure 6.3 (Figure 6.4, respectively).

By Figures 6.3 and 6.4, I first see that in all cases and both in terms of the average errors and maximum errors, my approximations outperform the risk-neutral solutions. Furthermore, in most cases, the improvements brought by my approximations are significant. Indeed, the approximations cut the errors of the risk-neutral solutions often by 3 to 6 times, although only one step of the continuation method was made at each κ. Second, I observe that the approximations are quite accurate for all cases of $n = 10$ and $n = 30$. However, the approximations do not work well at $n = 3$, which is similar to what I observed in the identical products case. Finally, I observe that the average and maximum errors of the risk-neutral solutions are decreasing in n which confirms with Proposition 6.

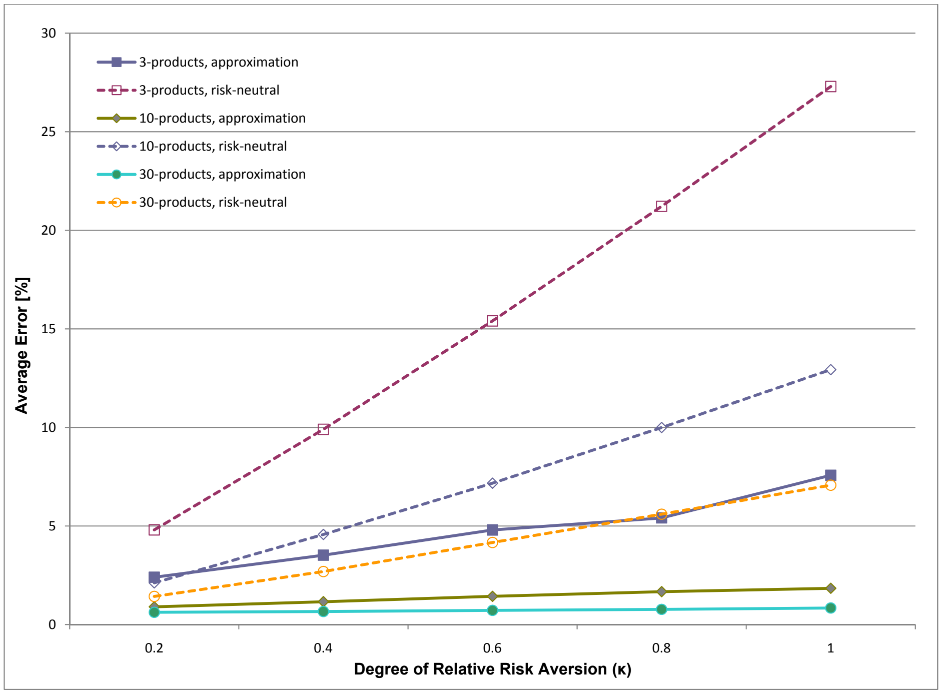

Figure 6.3: Heterogeneous Products with Independent Demands – Average percentage error of approximate solutions and risk-neutral solutions.

6.3.2 Impact of Dependent Demands under Risk Aversion

The objective of this subsection is to study the impact of demand dependence on the optimal order quantity under risk aversion. For this purpose, I first consider a simple system with two-identical products, then a system with two heterogenous products. The numerical results here are obtained by sample-based LP.

I choose the following cost parameters for the system with two identical products: $r_1 = r_2 = 15$, $c_1 = c_2 = 10$ and $s_1 = s_2 = 7$. I assume that demand follows bivariate lognormal distribution, which is generated by exponentiating a bivariate normal with the parameters $\mu_1 = \mu_2 = 3$, $\sigma_1 = \sigma_2 = 0.4724$ and a correlation coefficient of $-1, -0.8, -0.6, ..., 1$. Thus, the mean and standard deviation of each marginal distribution are 22.46 and 11.23 respectively, with CV $= 0.5$. The numerical results are summarized in Figure 6.5.

I draw the following insights by Figure 6.5: First, the numerical study confirms

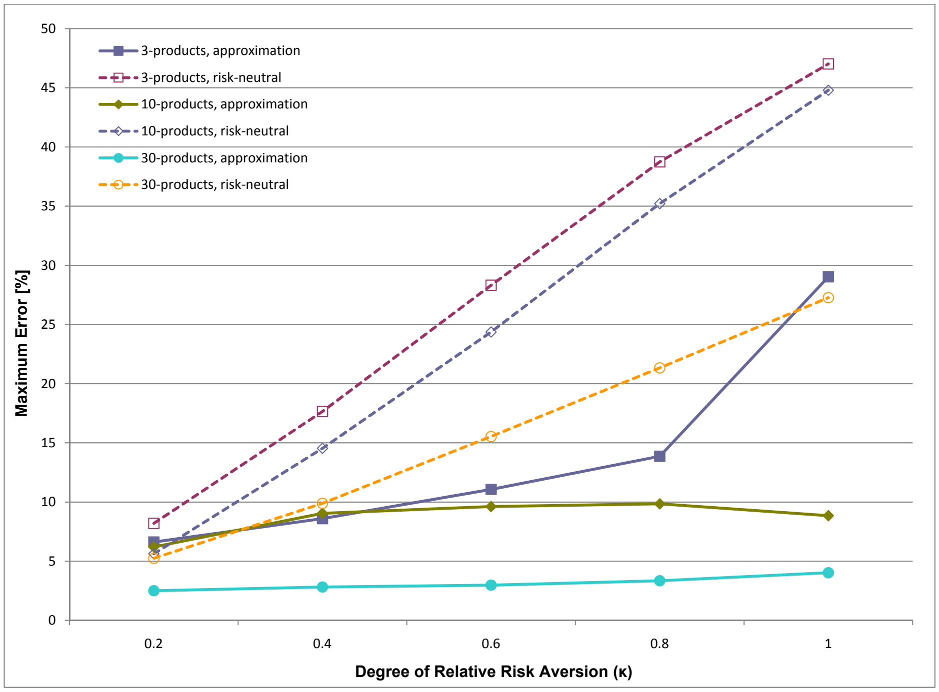

Figure 6.4: Heterogeneous Products with Independent Demands – Maximum percentage error of approximate solutions and risk-neutral solutions.

that demand dependence has no impact on the optimal order quantity for the risk-neutral newsvendor. Second, consistent with our analysis in §4.3, risk aversion reduces the optimal order quantity for independent demands. This is also true for positively correlated demands. But interestingly, this may not be true for strongly negatively correlated demands, where an increased risk aversion can result in a greater optimal order quantity. In order to explain the intuition behind these counterexamples, let's consider two-identical products with perfectly negatively correlated demands, D_1 and D_2: A larger order quantity, Q, can increase the strength of the negative correlation between the sales quantities of $\min(D_1, Q_1)$ and $\min(D_2, Q_2)$, and thus leads to smaller variability of the total sales, $\min(D_1, Q_1) + \min(D_2, Q_2)$. This intuition is implemented by my analysis of Case B at §4.6 when the risk-neutral solution for each product is higher than the median of marginal demand distribution, $\bar{r} > 2\bar{c}$. Third, consistent to our analysis in §4.5, negatively correlated demands result in higher optimal order quantities than independent demands under risk aversion, while positively correlated

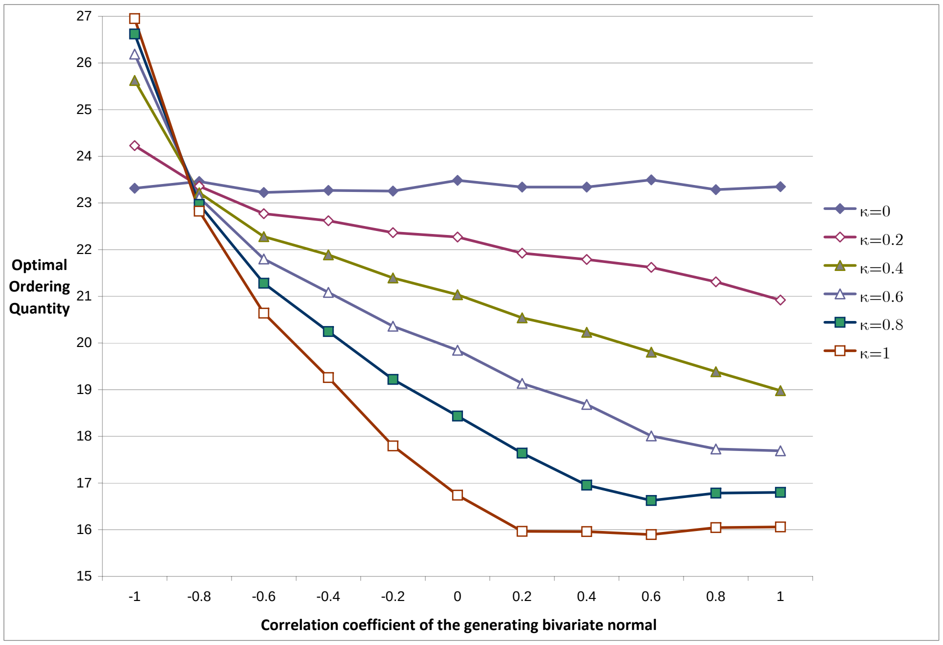

Figure 6.5: Identical Products with Dependent Demands – Impact of demand correlation under risk aversion κ.

demand leads to lower optimal order quantity under risk aversion. Finally, the impact of demand correlation is almost monotonic with small deviations generated by random sample errors.

Economically, these observations imply that if the firm is risk-averse, then demand dependence/correlation can have a significant impact on its optimal order quantities. These observations confirm with the intuition that stronger positively (negatively) correlated demands indicate higher (lower) risk, and therefore lead to lower (higher) order quantities.

For heterogenous products, I consider a simple system with two products and the following parameters: $r_1 = 15, c_1 = 10, s_1 = 7$ and $r_2 = 30, c_2 = 10, s_2 = 4$. The demand follows bivariate lognormal which is generated by exponentiating a bivariate normal with $\mu_1 = \mu_2 = 3$, $\sigma_1 = 0.4724$, $\sigma_2 = 1.26864$ and a correlation coefficient of $-1, -0.8, -0.6, ..., 1$. The marginal demand distribution of product 1 (2) has a mean 22.46 (44.913), a standard deviation 11.23 (89.826), and a cv 0.5 (2). Thus, product 1 is less risky and also less profitable than product 2.

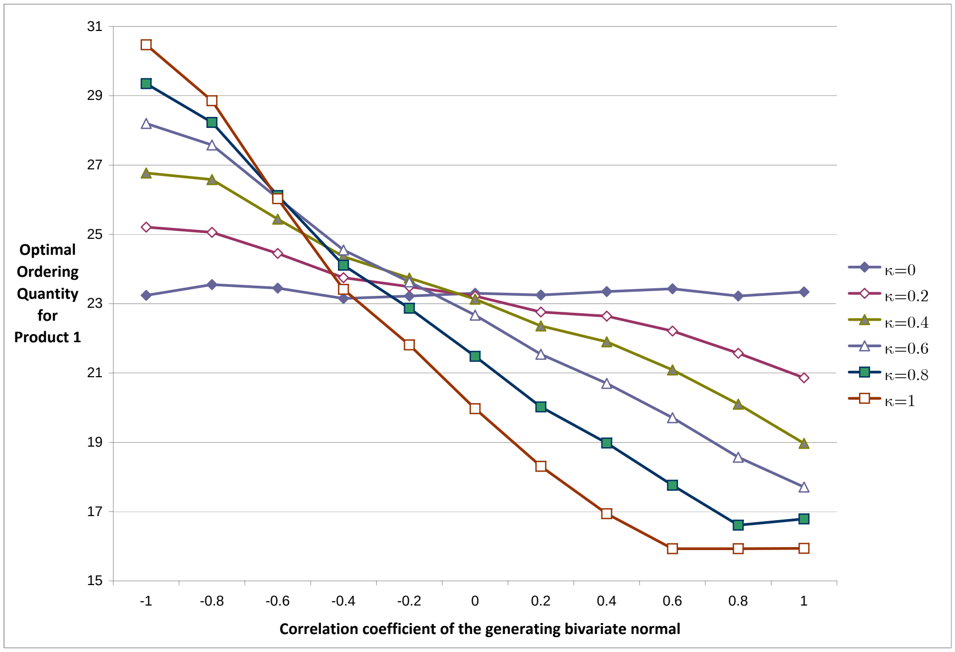

Figure 6.6: Heterogenous Products with Dependent Demands– Impact of demand correlation under risk aversion κ for the product with low risk and low profit.

My numerical study shows that for product 1 with low risk and low profit, the impact of demand correlation is similar to that for identical products; see Figure 6.6. However, for product 2 with high risk and high profit, the optimal ordering quantity always decreases in κ, but not in correlation except minor random sample errors; see Figure 6.6.

The economic implication is that for heterogenous products, the impact of demand correlation under risk aversion can be product-specific. Specifically, as the firm becomes more risk-averse, it should always order less for the more risky and more profitable products (due to its high risk). However, for the less risky and less profitable products, while it should order less when demands are positively correlated, it may order more when demands are strongly negatively correlated.

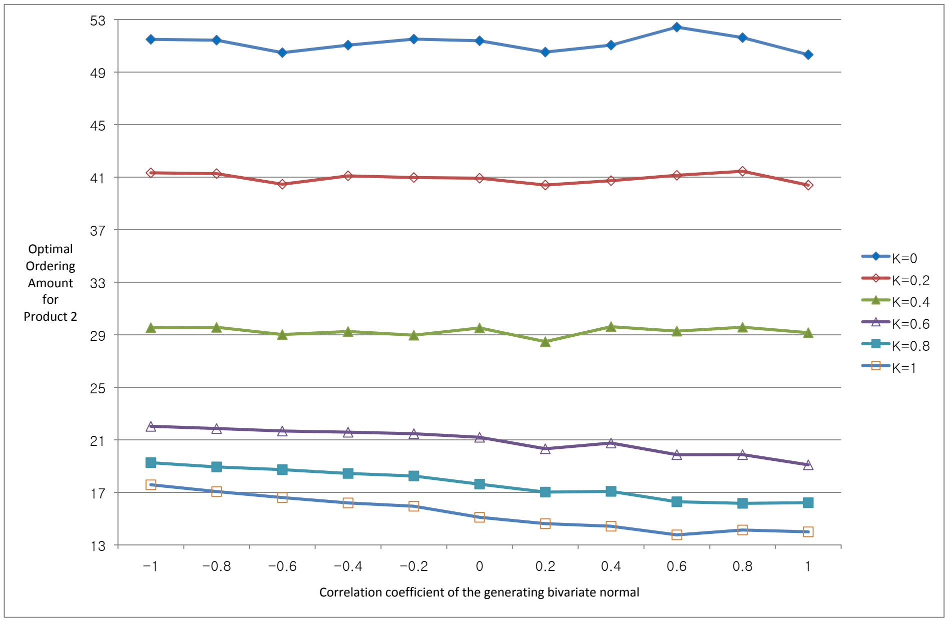

Figure 6.7: Heterogenous Products with Dependent Demands – Impact of demand correlation under risk aversion κ for the product with high risk and high profit.

6.4 Numerical Analysis for the Multi-Product Model under Exponential Utility Function

The objective of this section is two-fold. First, I demonstrate the accuracy and the convergence rate of the approximations. Second, I confirm the Proposition 9 and provide additional insights on the interplay between demand dependence and risk aversion.

In all examples considered I also apply sample-based optimization to solve the resulting stochastic programming problems similar to §6.3. In all our examples I used $T = 1,000$. For the empirical distribution, the corresponding optimization problem (5.2) has an equivalent nonlinear programming formulation. For each $j = 1, \ldots, n$ and $t = 1, \ldots, T$, I introduce the variable u_{jt} to represent the salvaged number of product j

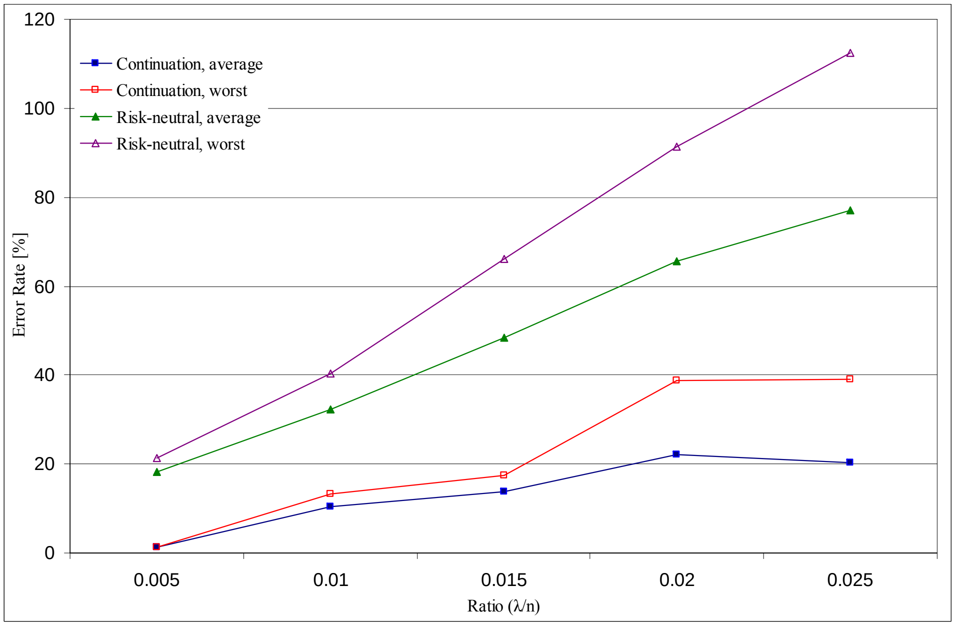

Figure 6.8: Independent Products – Average and maximum percentage errors of approximate solutions and risk-neutral solutions.

in scenario t. Then I obtain the formulation as follows:

$$\max \quad \frac{1}{T}\sum_{t=1}^{T}\left[e^{-\frac{\lambda}{n}\sum_{j=1}^{n}\Pi_{jt}}\right] \tag{6.2}$$

$$\begin{aligned}
\text{subject to} \quad & \Pi_{jt} = (r_j - c_j)x_j - (r_j - s_j)u_{jt}, \quad j = 1,\ldots,n; \quad t = 1,\ldots,T,\\
& x_j - d_{jt} \le u_{jt}, \quad j = 1,\ldots,n; \quad t = 1,\ldots,T,\\
& u_{jt} \ge 0, \quad j = 1,\ldots,n; \quad t = 1,\ldots,T,\\
& x_j \ge 0, \quad j = 1,\ldots,n.
\end{aligned}$$

6.4.1 Accuracy of Approximation

In this subsection, I test the accuracy of the closed-form approximations of §5.2.2 on ten randomly selected problems. For each problem, I calculated sample-based nonlinear programming solution by CPLEX and an approximation solution by the continuation method at §5.2.3. At each value of $\frac{\lambda}{n} = 0.005, 0.01, 0.015, 0.02, 0.025$, I made just one step of the straightforward approximation starting from the risk-neutral solution and continuation method. Similar to the results §6.3.1, the continuation method is also

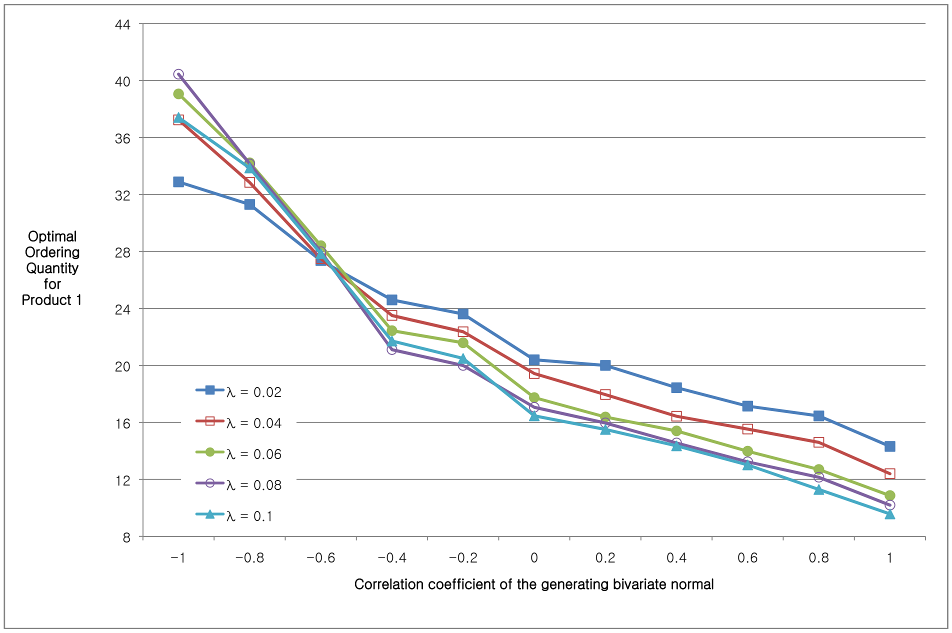

Figure 6.9: Impact of Demand Correlation under Risk Aversion for the Product with Low Risk and Low Profit.

much more stable and accurate than the straightforward approximation method starting from the risk-neutral solution especially for lower values of $\frac{\lambda}{n}$.

For each instance, I compute the average and maximum percentage errors of the approximate solution relative to the sample-based nonlinear programming solution for each value of $\frac{\lambda}{n}$. Then, for comparison purposes, I also compute the average and maximum percentage errors of the risk-neutral solution relative to the sample-based nonlinear programming solution. These errors are displayed in Figure 6.8. I see that for both of the average and maximum percentage errors, my approximations outperform the risk-neutral solutions by 3 to 5 times in all cases of the ratio. Then I observe that the average and maximum errors of risk-neutral solution and my approximation are decreasing in $\frac{\lambda}{n}$. These results are well in harmony with those at the case with coherent measures of risk.

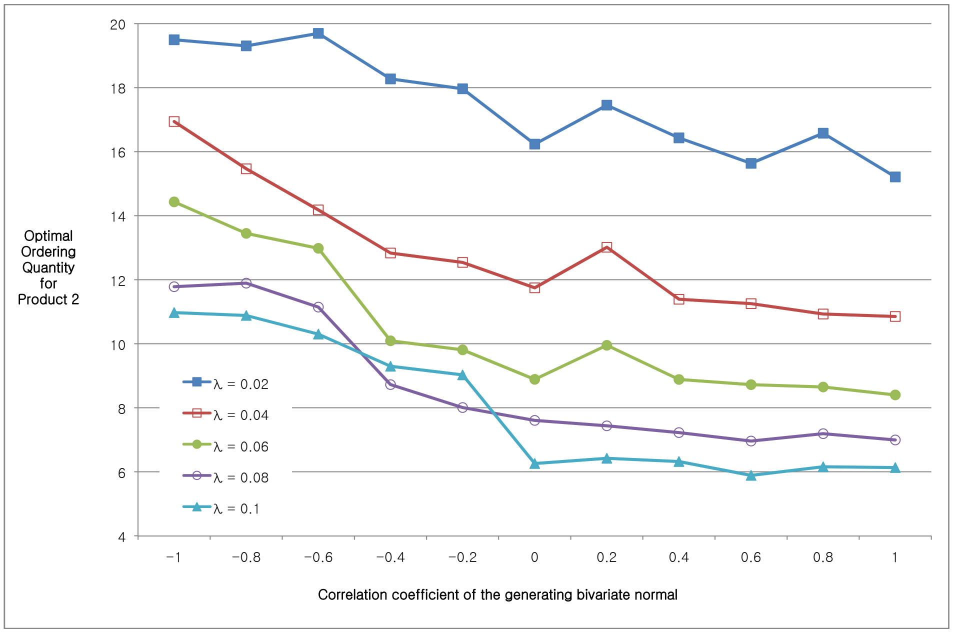

Figure 6.10: Impact of Demand Correlation under Risk Aversion for the Product with High Risk and High Profit.

6.4.2 Impact of Dependent Demands under Risk Aversion

The objective of this section is to study the impact of demand correlation on the optimal ordering amount under risk aversion. For this purpose, I consider a two-product system and the numerical results are obtained by the sample nonlinear programming problem. I choose the following parameters for the system: $r_1 = 15, c_1 = 10, s_1 = 7$ and $r_2 = 30, c_2 = 10, s_2 = 4$ with $\lambda = 0.02, 0.04, 0.06, 0.08, 0.1$. I also assume that demand follows bivariate lognormal distribution, which is generated by exponentiating a bivariate normal with the parameters $\mu_1 = \mu_2 = 3$ and $\sigma_1 = 0.4724, \sigma_2 = 1.2684$ to achieve the desirable coefficients of variance (CV) of 0.5 and 2. From this setting, product 1 (product 2) represents less (more) profitable and less (more) risky. The numerical results are summarized in Figures 6.9 and 6.10 and I can draw the following insights.

First, consistent with my analysis, risk aversion reduces the optimal order quantity for independent demands. This is also true for positively correlated demands. However, this may not be true for strongly negatively correlated demands, which is consistent with the intuition provided at section §6.3.2 and my analysis of the special cases at section §5.4. Second, consistent to my analysis, negatively correlated demands result in higher optimal order quantities than independent demands under risk aversion, while positively correlated demands lead to lower optimal order quantities under risk aversion. Third, for heterogenous products, the impact of demand correlation under risk aversion can be very different depending on the product heterogeneity. Forth, the impact of demand correlation is almost monotone with small deviations due to random sample errors. Last, all the above insights and economic implications are very similar to those with coherent measures of risk. One reason is that, in problem formulation the two models at Chapter 4 and 5 are only different from the risk measures used and the two risk measures share Convexity and Monotonicity axioms. Although many properties are alike and overlapped, the two models have similarities and differences, which are summarized at Chapter 7.

Chapter 7
Conclusion

I study single- and multi-product risk-averse newsvendor models under two risk measures, general coherent measures of risk and exponential utility function.

I formulate single-product models with the two risk measures, respectively. I study the impact of risk aversion on the optimal ordering quantity. Then I obtain closed-form optimal solutions under general coherent measures of risk and closed-form approximation under an exponential utility function. For the impact of risk aversion, the optimal ordering quantity decreases as the degree of risk aversion increases under both the two risk measures. This phenomenon can be explained that big ordering amount may increase the chance of getting higher revenue, but also increase the risk of being left as salvage items. Such trade-off relationship is very natural and the latter effect becomes more important as a newsvendor becomes more risk-averse. Thus, the more risk-averse newsvendor is, the less the optimal ordering quantity is, consequently. This is well consistent with our insights and typical findings from literature.

The multi-product newsvendor problem with coherent measures of risk does not decompose into independent problems, one for each product. The portfolio of products has to be considered as a whole. My analytical results focus on the impacts of risk aversion and demand dependence on the optimal order quantity. When product demands are independent, I analyze the asymptotic behavior of the optimal risk-averse solution with respect to the number of products and simple and accurate approximations of the optimal order quantities for a large number of products. For dependent demands case, I derive analytical and numerical insights how risk aversion and demand dependence interact to each other and affect to the optimal solution. My numerical examples confirm the accuracy of these approximations and enriches our understanding of the interplay of demand dependence and risk aversion.

The multi-product newsvendor with exponential utility function can be decomposable if the product demands are independent, otherwise not decomposable. When product demands are independent, I obtain closed-form approximations for sufficiently small ratios of the number of products to the degree of risk aversion and prove asymptotic behaviors of the solution with respect to the ratio of the degree of risk aversion to the number of products. For dependent demands case, I study the exponential utility function model by the similar way done with coherent measures of risk. My numerical examples also confirm the accuracy of the approximations and add understanding the interplay of demand dependence and risk aversion in a similar way done with coherent measures of risk. These results are in harmony with chapter 4 where coherent measures of risk are used.

1. Product Decomposability

   Under coherent measures of risk, the multi-product problem is not decomposable to each product, even if product demands are independent. For exponential utility function, when product demands are independent, the multi-product model can be separated into each single-product model and is equivalent to multiple individual single-product models.

2. The Monotonicity of the Impact of Risk Aversion

   Under coherent measures of risk, the monotonicity of the impact of risk aversion is proved only under the assumptions of independent demands and identical products. However, under exponential utility function, the same proposition is proved under the assumption of independent demands, even if the products considered are heterogenous.

3. Asymptotic Behaviors

   When the degree of risk aversion is zero, the exponential utility function approach does not work. However, when the degree of risk aversion converges to zero, the corresponding risk-neutral solution asymptotically optimal. In another extreme case, when the degree of risk aversion goes to infinity, the optimal solution converges to zero in the exponential utility function. On the other hand, under coherent measures of risk, the multi-product model is always defined at the zero degree of risk aversion where the corresponding risk-neutral solution is optimal regardless of demand dependency. For coherent measures of risk, the degree of risk aversion always has a finite value because it is defined only in a limited range to satisfy all the four axioms.

4. Closed-Form Approximations

   In the exponential utility function, the degree of risk aversion and the number of products are not considered separately, but affect the solution via the ratio of the degree of risk aversion to the number of products $\left(\frac{\lambda}{n}\right)$. Then, I obtain a closed-form approximation for sufficiently small $\left(\frac{\lambda}{n}\right)$. On the other hand, under coherent measures of risk, these two factors are clearly separated. Thus, a closed-form approximation is always obtained for sufficiently large number of products regardless of the degree of risk aversion. This is also consistent with that, under coherent measures of risk, the degree of risk aversion is only defined at a certain limited range of finite values.

5. Dependent Demands Case

   When product demands are dependent, the multi-product models are not generally decomposable under both models and demand dependence has significant impacts on optimal order quantities. For the impact of demand correlation under risk aversion, I am able to prove that one tends to decrease (or increase) the order quantity in case of positively (or negatively) dependent demands relative to the case of independent demands with exponential utility function. However, with coherent measures of risk, the same proposition is proved only in two-identical product systems. In addition, the impact of demand correlation under risk aversion can be very different in both models when the heterogeneity increases between products. For the impact of risk aversion, the optimal order quantity decreases in risk aversion for both models with independent or positively correlated demands cases. When the demand has highly negative correlations, risk-averse solutions may be equal to, higher than or lower than risk-neutral solution in both models. In §4.6 and §5.4, I discuss three special cases of perfectly negative demand correlation in two-identical product systems with each risk measure.

It is appropriate to conclude my dissertation by comparing the multi-product risk-averse newsvendors to the risk-averse portfolio optimization problem. In a portfolio problem, there are n assets with random returns $R_1, \ldots, R_n$ and the objective is to determine optimal investment quantities $x_1, \ldots, x_n$ to obtain the best desirable characteristics of the total portfolio return $P(x, R) = R_1 x_1 + \cdots + R_n x_n$. In the classical mean–variance approach of Markowitz (1952 and 1959), the mean of the return and its variance are used to find efficient portfolio allocations. See also Elton, Gruber, Brown and Goetzmann (2006). In more modern approaches (refer to Konno and Yamazaki (1991), Mansini, Ogryczak and Speranza (2003), Ruszczyński and Vanderbei (2003) and Miller and Ruszczyński (2008)) more general mean–risk models and coherent measures of risk are used, similarly to problem (5.2). There are, however, fundamental structural differences which make the multi-product newsvendor problem significantly different from the financial portfolio problem.

The most important difference is that the portfolio return $P(x, R)$ is *linear* with respect to the decision vector x, while the newsvendor profit $\Pi(x, D)$ is *concave* and *nonlinear* with respect to the order quantities x. This leads to the following different properties of the problems.

- The risk-neutral portfolio problem has no solution, unless the total amount invested is restricted (e.g., to 1), in which case the optimal solution is to invest everything in the asset(s) having highest expected returns. On the contrary, the risk-neutral newsvendor problem always has a solution because of natural limitations of the demand only except the trivial case of zero degree of risk aversion with exponential utility function.

- The effect of using risk measures in the portfolio problem is a diversification of the solution, which otherwise would remain completely non-diversified. In the newsvendor problems the use of risk measures results in changes of the already diversified risk-neutral solution, by ordering more of products having less variable or negatively correlated demands and less of products having more variable or positively correlated demands. The optimal order quantity for each product is unlikely to be zero due to risk aversion, because very small amounts will almost always be sold and thus they introduce very little risk.

- In the portfolio problem, independently of the number of assets considered, the risk-neutral solution remains structurally different from the risk-averse solution. On the contrary, in the newsvendor models, the number of products affects to the optimal order quantity with coherent measures of risk. For exponential utility function, the number of products affects to the optimal solution via the form of $\left(\frac{\lambda}{n}\right)$. For the asymptotic behaviors with coherent measures of risk, the risk-neutral solution converges to the optimal solution under risk aversion as the number of products goes to infinity. On the other hand, when $\left(\frac{\lambda}{n}\right)$ converges to zero, the risk-neutral solution is asymptotically optimal under risk aversion with exponential utility function.

Finally it is worth stressing that the nonlinearity of the newsvendor profit $\Pi(x, D)$ is the source of formidable technical difficulties in the analysis of the composite functions (4.5), which involves two nondifferentiable functions.

References

[1] Acerbi, C. (2004) Coherent Representation of Subjective Risk Aversion. in: G. Szegö. *Risk Measures for the 21st Century*, Wiley (Chichester), 147-207.

[2] Acerbi, C., and D. Tasche (2002). On the coherence of expected shortfall. *Journal Of Banking and Finance*, **26(7)**, 1487-1503.

[3] Ağrali, S. and A. Soylu (2006). Single Period Stochastic Problem with CVaR Risk Constraint. Presentation, University of Florida.

[4] Agrawal, V. and S. Seshadri (2000a). Impact of Uncertainty and Risk Aversion on Price and Order Quantity in the Newsvendor Problem. *Manufacturing and Service Operations Management*, **2(4)**, 410-423.

[5] Agrawal, V. and S. Seshadri (2000b). Risk Intermediation in Supply Chains. *IIE Transactions*, **32**, 819-831.

[6] Ahmed, S., U. Çakmak and A. Shapiro (2007). Coherent Risk Measures in Inventory Problems. *European Journal of Operational Research*, **182**, 226-238.

[7] Ahmed, S., D. Filipovic and G. Svindland (2008). A Note on Natural Risk Statistics. *Operations Research Letters*, **36**, 662-664.

[8] Allais, M. (1953). Le Comportement de l'homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'école Américaine. *Econometrica*, **21**, 503-546.

[9] Anvari, M. (1987). Optimality Criteria and Risk in Inventory Models: The Case of the Newsboy Problem. *The Journal of the Operational Research Society*, **38(7)**, 625-632.

[10] Arnold, B. C. (1980). Majorization and the Lorenz Order: A Brief Introduction. *Lecture Notes in Statistics*, **43**, Springer-Verlag.

[11] Artzner, P., F. Delbaen, J. Eber and D. Heath (1999). Coherent Measures of Risk. *Mathematical Finance*, **9(3)**, 203-228.

[12] Aviv, Y. and A. Federgruen (2001). Capacitated Multi-Item Inventory Systems with Random and Fluctuating Demands: Implications for Postponement Strategies. *Management Science*, **47(4)**, 512-531.

[13] Bassok, Y., R. Anupindi and R. Akella (1999). Single-period Multi-Product Inventory Models with Substitution. *Operations Research*, **47(4)**, 632-642.

[14] Bawa, V.S. (1982). Stochastic Dominance: A Research Bibliography. *Management Science*, **28**, 698-712.

[15] Bonnans, J. F. and A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York

[16] Bouakiz, M. and M. Sobel (1992). Inventory Control with an Exponential Utility Criterion. *Operations Research*, **40(3)**, 603-608.

[17] Charnes, A., W. W. Cooper and G. H. Symonds (1958). Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil, *Management Science*, **4**, 235-263.

[18] Chen, F. and A. Federgruen (2000). Mean-Variance Analysis of Basic Inventory Models. Working paper, Columbia University.

[19] Chen, X., M. Sim, D. Simchi-Levi and P. Sun (2007). Risk Aversion in Inventory Management. *Operations Research*, **55**, 828-842.

[20] Choi, S. and A. Ruszczyński (2008). A Risk-Averse Newsvendor with Law Invariant Coherent Measures of Risk. *Operations Research Letters*, **36**, 77-82.

[21] Choi, S., A. Ruszczyński and Y. Zhao (2009). A Multi-Product Risk-Averse Newsvendor with Law Invariant Coherent Measures of Risk. Working Paper, Rutgers University.

[22] Choi, S. and A. Ruszczyński (2009). A Risk-Averse Newsvendor with Exponential Utility Function. Working Paper, Rutgers University.

[23] Chung, K. H. (1990). Risk in Inventory Models: The Case of the Newsboy Problem – Optimality Conditions. *The Journal of the Operational Research Society*, **41(2)**, 173-176.

[24] Decroix, G. and A. Arreola-Risa (1998). Optimal Production and Inventory Policy for Multiple Products Under Resource Constraints. *Management Science*, **44**, 950-961.

[25] Delbaen, F. (2002). Coherent Risk Measures on General Probability Space. *Advance in Finance and Stochastics*, Springer, Berlin, 1-37.

[26] Dentcheva, D. and A. Ruszczyński (2003). Optimization with Stochastic Dominance Constraints. *SIAM Journal on Optimization*, **14(2)**, 548-566.

[27] Dentcheva, D. and A. Ruszczyński (2004). Semi-infinite probabilistic optimization: First order stochastic dominance constraints. *Optimization*, **53**, 583-601.

[28] Dentcheva, D. and A. Ruszczyński (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking and Finance*, **30**, 433-451.

[29] Dowd, K. (1997). Beyond Value at Risk. *The Science of Risk Management*, Wiley, New York.

[30] Eeckhoudt, L., C. Gollier and H. Schlesinger (1995). The Risk-averse (and Prudent) Newsboy. *Management Science*, **41(5)**, 786-794.

[31] Ellsberg, D. (1961). Risk, Ambiguity and Savage Axioms. *Quarterly Journal of Economics*, **75(4)**, 643-669.

[32] Elton, E. J., M. J. Gruber, S. J. Brown, and W. N. Goetzmann (2006) , *Modern Portfolio Theory and Investment Analysis*, John Wiley & Sons, New York.

[33] Esary, J. D., F. Proschan & D. W. Walkup (1967). Association of Random Variables, with Applications. *Annals of Mathematical Statistics*, **38**, 1466-1474.

[34] Evans, R. (1967). Inventory Control of a Multiproduct System with a Limited Production Resource. *Naval Research Logistics*, **14(2)**, 173-184.

[35] Federgruen, A. and P. Zipkin (1984). Computational issues in an infinite-horizon, multi-echelon inventory model. *Operations Research*, **32**, 818-836.

[36] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 3rd ed. Wiley, New York, NY.

[37] Fishburn, P. C. (1964). *Decision and Value Theory*, Wiley, New York, NY.

[38] Föllmer, H. and A. Schied (2002). Convex Measures of Risk and Rrading Constraints. *Finance and Stochastics*, **6**, 429-447.

[39] Föllmer, H. and A. Schied (2004). *Stochastic Finance: An Introduction in Discrete Time*, de Gruyter, Berlin.

[40] Gan, X., S.P. Sethi and H. Yan (2004). Coordination of Supply Chains with Risk-Averse Agents. *Production and Operations Management*, **14(1)**, 80-89.

[41] Gan, X., S.P. Sethi and H. Yan (2005). Channel Coordination with a Risk-Neutral Supplier and a Downside-Risk-Averse Retailer. *Production and Operations Management*, **13(2)**, 135-149.

[42] Gaur, V. and S. Seshadri (2005). Hedging Inventory Risk Through Market Instruments. *Manufacturing and Service Operations Management*, **7(2)**, 103-120.

[43] Gotoh, J. and Y. Takano (2007). Newsvendor Solutions via Conditional Value-at-risk Minimization. *European Journal of Operational Research*, **179**, 80-96.

[44] Hadar, J and W. Russell (1969). Rules for Ordering Uncertain Prospects. *The American Economic Review*, **59**, 25-34.

[45] Hadley, G. and T. M. Whitin (1963). *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, New Jersey.

[46] Hanoch, G., and H. Levy (1969). The Efficiency Analysis of Choices Involving Risk. *Review of Economic Studies*, **36**, 335-346.

[47] Hardy, G.H., J.E. Littlewood and G. Polya (1934). *Inequalities*, Cambridge University Press, Cambridge, MA

[48] Heyde, C.C., S.G. Kou and X.H. Peng (2006). What is a Good Risk Measure: Bridging the Gaps Between Data, Coherent Risk measures and Insurance Risk Measures. Working paper, Columbia University

[49] Heyman, D.P. and M.J. Sobel (1984). *Stochastic Models in Operations Research: Volume 2. Stochastic Optimization*, McGraw Hill Company.

[50] Ignall, E. and D.P. Veinott Jr. (1969). Optimality of Myopic Inventory Policies for several Substitute Products. *Management Science*, **15**, 284-304.

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

[52] Klein Haneveld, W. K. (1986). Duality in Stochastic Linear and Dynamic Programming. *Lecture Notes in Economics and Mathematical Systems*, **274**, Springer-Verlag, New York.

[53] Klein Haneveld, W. K. and M. H. Van Der Vlerk (2002). Integrated Chance Constraints: Reduced Forms and an Algorithm, Research Report 02A33, SOM, University of Groningen, The Netherlands. Also available as *Stochastic Programming* E-Print Series 2004-14.

[54] Konno, H. and H. Yamazaki (1991). Mean–Absolute Deviation Portfolio Optimization model and its Application to Tokyo Stock Market. *Management Science*, **37**, 519-531.

[55] Kusuoka, S. (2003). On Law Invariant Coherent Risk Measures. *Advances in Mathematical Economics*, **3**, 83-95.

[56] Kusuoka, S. and Y. Morimoto (2004). Homogeneous Law Invariant Coherent Multiperiod Value Measures and Their Limits. Working Paper, Graduate School of Mathematical Science, University of Tokyo, Tokyo, Japan.

[57] Lau, H.S. (1980). The Newsboy Problem under Alternative Optimization Objectives. *The Journal of the Operational Research Society*, **31(6)**, 525-535.

[58] Lau, H.S. and H.L. Lau (1999). Manufacturer's Pricing Strategy and Return Policy for a Single-Period Commodity. *European Journal of Operational Research*, **116**, 291-304.

[59] Lehmann, E. (1955), Ordered Families of Distributions, *Annals of Mathematical Statistics*, **26**, 399-419.

[60] Levy, H. (1992). Stochastic Dominance and Expected Utility: Survey and Analysis *Management Science*, **38**, 555-593.

[61] Levy, H. (2006). *Stochastic Dominance: Investment Decision Making Under Uncertainty*, 2nd Edition, Springer, New York.

[62] Lorenz, M. O. (1905). Methods of Measuring Concentration of Wealth. *Journal of American Statistical Association*, **9**, 209-219.

[63] Mahajan, S. and G.J. van Ryzin (2001). Inventory Competition Under Dynamic Consumer Substitution. *Operations Research*, **49**, 646-657.

[64] Mansini, R., W. Ogryczak and M. G. Speranza MG (2003). On LP Solvable Models for Portfolio Selection. *Informatica*, **14**, 37-62.

[65] Markowitz, H.M. (1952). Portfolio Selection. *Journal of Finance*, **7**, 77-91.

[66] Markowitz, H.M. (1959). *Portfolio Selection*, Wiley, New York, NY.

[67] Marshall, A. W. and I. Olkin (1979). *Inequalities: Theory of Majorization and Its Applications*, Academic Press, San Diego

[68] Martinez-de-Albeniz, V. and D. Simchi-Levi (2006). Mean-Variance Trade-offs in Supply Contracts. *Naval Research Logistics*, **53**, 603-616.

[69] Miller, N., and A. Ruszczynski (2008). Risk-Adjusted Probability Measures in Portfolio Optimization with Coherent Measures of Risk. *European Journal of Operational Research*, **191**, 193-206.

[70] Miller, L. B. and H. Wagner (1965). Chance-Constrained Programming with Joint Constraints. *Operations Research*, **13**, 930-945.

[71] Müller, A., and D. Stoyan (2002). *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons, Chichester.

[72] Noyan, N. (2006). Optimization with First-Order Stochastic Dominance Constraints. PhD Dissertation, RUTCOR (RUTgers Center of Operations Research), Rutgers University

[73] Ogryczak, W. and A. Ruszczyński (1999). From Stochastic Dominance to Mean-Risk Models. *European Journal of Operational Research*, **116**, 33-50.

[74] Ogryczak, W. and A. Ruszczyński (2001). On Consistency of Stochastic Dominance and Mean-Semideviation Models. *Mathematical Programming*, **89**, 217-232.

[75] Ogryczak, W. and A. Ruszczyński (2002). Dual Stochastic Dominance and Related Mean-Risk Models. *SIAM Journal on Optimization*, **13(1)**, 60-78.

[76] Özler, A., B. Tan and F. Karaesmen (2009). Multi-Product Newsvendor Problem with Value-at-Risk Constraints. *International Journal of Production Economics*, **117**, 244-255.

[77] Pflug, G. Ch. (2000), Some remarks on the Value-at-Risk and the Conditional Value-at-Risk, in: S. Uryasev (Ed.). *Probabilistic Constrained Optimization - Methodology and Applications*, Kluwer Academic Publishers, 272-281.

[78] Porteus, E.L. (1990). Stochastic Inventory Theory, in: D.P. Heyman and M.J. Sobel (Eds.). *Handbook in OR and MS*, Vol. 2, Elsevier Science Publishers B.V., Chapter 12, 605-652.

[79] Pospišil, L. and J. Večer and M. Xu (2008). Tradeable Measures of Risk. Working Paper, Columbia University

[80] Prékopa, A. (1970). On Probabilistic Constrained Programming. *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, New Jersey, 113-138.

[81] Prékopa, A. (1973). Contributions to the Theory of Stochastic Programming. *Mathematical Programming*, **4**, 202-221.

[82] Prékopa, A. (2003). Probabilistic Programming, In: Ruszczyński, A., and A. Shapiro (eds.) *Stochastic Programming*, Elsevier, Amsterdam.

[83] Quirk, J. P. and R. Saposnik (1962). Admissibility and Measurable Utility Functions. *Review of Economic Studies*, **29**, 140-146.

[84] Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic Processes and Their Applications*, **112**, 185-200.

[85] Rockafellar, R.T. and S.P. Uryasev (2000). Optimization of Conditional Value-at-Risk. *The Journal of Risk*, **2**, 21-41.

[86] Rockafellar, R.T. and S.P. Uryasev (2002). Conditional Value-at-Risk for General Loss Distribution. *The Journal of Banking and Finance*, **26**, 1443-1471.

[87] Rothschild, M. and J. E. Stiglitz (1969). Increasing risk: I. A Definition. *Journal of Economic Theory*, **2**, 225-243.

[88] Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica*, **20**, 431-449.

[89] Ruszczyński, A. and A. Shapiro (2005). Optimization of Risk Measures. in: G. Calafiore and F. Dabbene (Eds.). *Probabilistic and Randomized Methods for Design under Uncertainty*, Springer-Verlag, London, 117-158.

[90] Ruszczyński, A. and A. Shapiro (2006a). Optimization of Convex Risk Functions. *Mathematics of Operations Research*, **31(3)**, 433-452.

[91] Ruszczyński, A. and A. Shapiro (2006b). Conditional Risk Mappings. *Mathematics of Operations Research*, **31(3)**, 544-561.

[92] Ruszczyński, A. and R. J. Vanderbei (2003). Frontiers of Stochastically Nondominated Portfolios. *Econometrica*, **71(4)**, 1287-1297.

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

[94] Shapiro, A. (2003). Monte Carlo Sampling Method. in: A. Ruszczyński, A. Shapiro (Eds.). *Stochastic Programming*, Elsevier, Amsterdam.

[95] Shapiro, A. (2008). Asymptotics of Minimax Stochastic Programs. *Statistics and Probability Letters*, **78**, 150-157.

[96] Shapiro, A., D. Dentcheva, and A. Ruszczyński (2009). *Lectures on Stochastic Programming: Modeling and Theory*, SIAM (Society for Industrial and Applied Mathematics), Philadelhpia.

[97] Sharpe. W. (1971). A Linear Programming Approximation for the General Portfolio Analysis Problem. *Journal of Financial and Quantitative Analysis*, **6(5)** 1263-1275.

[98] Smith, A. A. and N. Agrawal (2000). Management of Multi-Item Retail Inventory System with Demand Substitution. *Operations Research*, **48(1)**, 50-64.

[99] Tomlin, B. and Y. Wang (2005). On the Value of Mix Flexibility and Dual Sourcing in Unreliable Newsvendor Networks. *Manufacturing and Service Operations Management*, **7(1)**, 37-57.

[100] Tong, Y. L. (1980). *Probability Inequalities in Multivariate Distributions*. Academic Press, New York.

[101] Tsay, A. A. (2002). Risk Sensitivity in Distribution Channel Partnerships: Implications for Manufacturer Return Policies. *Journal of Retailing*, **78**, 147-160.

[102] Tverski, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, **5**, 297-323.

[103] van Mieghem, J. A. (2003). Capacity management, investment and hedging: Review and recent developments. *Manufacturing Service Operations Management*, **5(4)**, 269-302.

[104] van Mieghem, J. A. (2007). Risk Mitigation in Newsvendor Networks: Resource Diversification, Flexibility, Sharing, and Hedging. *Management Science*, **53(8)**, 1269-1288.

[105] van Ryzin, G.J. and S. Mahajan (1999). On the Relationship Between Inventory Cost and Variety Benefits in Retail Assortments. *Management Science*, **45**, 1496-1509.

[106] Veinott, A. (1965). Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science*, **12**, 206-222.

[107] von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

[108] Wang, S.S., V.R. Young and H.H. Panjer (1997). Axiomatic Characterization of Insurance Prices. *Insurance: Mathematics and Economics*, **21**, 173-183.

[109] Whitmore, G.A. and M.C. Findlay (Eds.) (1978). *Stochastic Dominance: An Approach to Decision-Making Under Risk*, Heath, Lexington, MA

Vita

Sungyong Choi

2004 M.S. in Probability and Statistics, Michigan State University.

2002 Supply Officer (Rank: Lieutenant, Junior Grade), The First Fleet, South Korean Navy.

1999 Research Scientist, The Institute for Business Research and Education, South Korea.

1998 M.B.A., Korea University, South Korea.

1994 B.S. in Management Science, Korea Advanced Institute of Science and Technology, South Korea.

Publication

2008 Choi, S. and A. Ruszczyński. A Risk-Averse Newsvendor with Law Invariant Coherent Measures of Risk. *Operations Research Letters*, **36**, 77-82.

Working Papers

2009 Choi, S., A. Ruszczyński and Y. Zhao. A Multi-Product Risk-Averse Newsvendor under Coherent Risk Measures. In-preparation for the 3^{rd} round review at *Operations Research* for publication.

2009 Choi, S. and A. Ruszczyński. A Multi-Product Risk-Averse Newsvendor with Exponential Utility Function. In-preparation for the 2^{nd} round review at *European Journal of Operational Research* for publication.