A Multiproduct Risk-Averse Newsvendor with Law-Invariant Coherent Measures of Risk*

Sungyong Choi † Andrzej Ruszczyński ‡ Yao Zhao §

August 1, 2010

Abstract

We consider a multi-product risk-averse newsvendor under the law-invariant coherent measures of risk. We first establish several fundamental properties of the model regarding the convexity of the problem, the symmetry of the solution and the impact of risk aversion. Specifically, we show that for identical products with independent demands, increased risk aversion leads to decreased orders. For a large but finite number of heterogenous products with independent demands, we derive closed-form approximations for the optimal order quantities. The approximations are as simple to compute as the classical risk-neutral solutions. We also show that the risk-neutral solution is asymptotically optimal as the number of products tends to be infinity, and thus risk aversion has no impact in the limit. For a risk-averse newsvendor with dependent demands, we show that positively (negatively) dependent demands lead to a lower (higher) optimal order quantities than independent demands. Using a numerical study, we examine the convergence rates of the approximations and develop additional insights on the interplay between dependent demands and risk aversion.

Keywords and Phrases: Multiple products, newsvendor, risk averse, coherent risk measure, diversification, portfolio.

1 Introduction

1.1 Motivation

The multi-product newsvendor model is a classical model in the inventory control literature. In this model, there are multiple products to be sold in a single selling season. 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. The firm's objective is to determine the optimal order quantity for each product so as to maximize a certain performance measure. This model finds its applications in many manufacturing, distribution and retailing firms that handle short life-cycle products.

*This paper was submitted on January 28, 2008 and revised on January 14, 2009, September 29, 2009, and March 15, 2010. Forthcoming at *Operations Research* in June 2, 2010.

†Division of Systems and Engineering Management, Nanyang Technological University, sungyongchoi@gmail.com, http://pegasus.rutgers.edu/~sungyong

‡Department of Management Science and Information Systems, Rutgers University, rusz@business.rutgers.edu, http://www.rusz.rutgers.edu

§Department of Supply Chain Management and Marketing Sciences, Rutgers University, yaozhao@andromeda.rutgers.edu, http://zhao.rutgers.edu

The literature of the multi-product newsvendor model focuses mainly on risk-neutral performance measures, when the firm maximizes the expected total profit. Under such a measure, the model decomposes into multiple single-product newsvendor models, unless resource constraints are imposed or demand substitution is allowed (see [30, 64]). [47] surveys various newsvendor models.

Our aim is to replace the risk-neutral performance measure by measures taking risk aversion into account. Such a model is generally not decomposable, and one needs to consider all products simultaneously, as a portfolio. In this paper, we lay the foundations of the multi-product newsvendor model under coherent measures of risk and we derive its basic properties. They provide insight into the impact of risk aversion on the multi-product newsvendor with either independent or dependent demands. Moreover, we study asymptotic properties of the solution as the number of products tends to infinity, and we develop simple yet accurate approximations of risk-averse solutions, which allow fast computation of large-scale problems.

Below, we first review the literature of risk measures and their recent applications in supply chain inventory management. Then, we summarize our model and main results.

1.2 Risk Measures

The risk-neutral inventory models provide the best decision *on average*. This may be justified by the Law of Large Numbers. However, we cannot always rely on repeated similar chances. The first few outcomes may turn out to be very bad and entail unacceptable losses. [56] provide experimental evidence suggesting that inventory managers may be risk-averse for high-value products. Because of these reasons, attempts to overcome the drawbacks of the expected value optimization have a long history and there exist four typical approaches to model decision making under risk. They are expected utility theory, stochastic dominance, chance constraints and mean-risk analysis. These approaches are related and consistent, to some extent.

The expected utility theory of [65] derives, from simple axioms, the existence of a nondecreasing *utility function*, which transforms the observed outcomes in a nonlinear way. The decision maker optimizes, instead of the expected outcome, the expected value of the utility function. In the maximization context, when the outcome represents profit, risk-averse decision makers have concave and nondecreasing utility functions.

The second approach is based on the theory of *stochastic dominance*, developed in statistics and economics (see [38, 29] and references therein). Stochastic dominance relations are partial orders on the space of distributions, and thus allow for pairwise comparison of different solutions. 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 V is preferred to random outcome Y in terms of a stochastic dominance relation if and only if expected utility of V is preferred to expected utility of Y for all utility functions in a certain class, called the generator of the relation. In particular, the second-order stochastic dominance corresponds to all concave nondecreasing utility functions, and is thus well suited to model risk-averse preferences. For an overview of these issues, see [43, 39]. 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 (see [18]), and also utilized as a reference standard whether a particular solution approach is appropriate (see [44, 55]).

The third approach specifies constraints on probabilities of unfavorable events. [49] 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 stochastic dominance relation of the first-order, and thus it is related to the expected utility theory, but there is no equivalence. In finance, chance constraints are known under the name of Value-at-Risk (VaR) constraints. Chance constraints sometimes lead to non-convex formulations of the resulting optimization problems.

The fourth approach, 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 the maximization context, we select from the universe of all possible solutions those that are *efficient*: for a given value of the mean they minimize the risk, or equivalently, for a given value of risk they maximize 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.

In the context of portfolio optimization, [40] used the variance of the return as the risk functional. It is easy to compute, and it reduces the financial portfolio selection problem to a parametric quadratic programming problem. One can, however, construct simple counterexamples that show the imperfection of the variance as the risk measure: it treats over-performance equally as under-performance, and more importantly it may suggest a portfolio which is stochastically dominated by another portfolio.

To overcome the drawbacks of the mean-variance analysis, the general theory of *coherent measures of risk* was suggested by [7] and extended to general probability spaces by [17]. For further generalizations, see [23, 24, 34, 52, 53]. Dynamic versions for a multi-period case were analyzed, among others, by [50, 35, 13, 54]. In this theory, an integrated performance measure is proposed, comprising both the mean and variability measures, and four axioms (Convexity, Monotonicity, Translation Equivariance and Positive Homogeneity; see §3 for a precise definition) are imposed to guarantee consistency with intuition about rational risk-averse decision making. Coherent measures of risk are extensions of the mean-risk analysis. It is known that coherent measures of risk are consistent with the 1st and 2nd order stochastic dominance relations (see [58]).

Expected utility models and coherent risk measures share the properties of convexity and consistency with stochastic dominance. In addition, the coherent risk measures satisfy the axioms of Translation Equivariance and Positive Homogeneity. For a multi-product newsvendor, the Translation Equivariance axiom implies that adding a constant gain is equivalent to changing the vendor's performance measure by the same amount; the Positive Homogeneity axiom guarantees that one obtains the same solution when considering the total profit or the profit rate (i.e., average profit per product), and when one changes the currency in which the profit is calculated. Under expected utility theory, these two axioms typically do not hold; see, e.g., the exponential utility function in [32].

For inventory systems where the initial endowment effect is significant, i.e., when the initial wealth could affect the decision of a risk-averse manager, or when constant demand for some products could affect order quantities of other products, an expected utility model may be preferred to a model with a coherent risk measure, because the latter ignores the endowment effect. In newsvendor models, where we are mainly concerned about the overage and underage costs associated with random demand, and in other problems, where risk is primarily associated with uncertainty, coherent risk measures may capture risk preferences better. The following arguments speak in favor of coherent measures of risk: (i) Translation Equivariance allows us to properly rank risky alternatives by excluding the impact of constant gains or losses (see [7]). (ii) The Positive Homogeneity axiom ensures that our attitude to risk will not change when the unit system is changed (e.g., from dollars to cents). More importantly, this axiom indicates no diversification effect when demands are completely correlated. To see this, we note that the subadditivity property, $\rho(X + Y) \leq \rho(X) + \rho(Y)$, implies $\rho(nX) \leq n\rho(X)$. However $\rho(nX) < n\rho(X)$ would imply diversification effect even when the random demands are completely correlated. To avoid this counter-intuitive effect, we are left with $\rho(nX) = n\rho(X)$ which is the Positive Homogeneity axiom.

Several modifications and extensions of coherent measures of risk have been suggested in the literature, including *convex measures of risk, insurance risk measures, natural risk statistic* and *tradeable measures of risk*. We point out that all these risk measures ignore the initial endowment effect, and a coherent risk measure has certain attractive features as compared to these measures, making it worth considering.

[23] 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 our 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 widely used. [66] suggest insurance risk measures which are law-invariant, and satisfy the axioms of conditional state independence, monotonicity, comonotone additivity and continuity. [31] propose the natural risk statistics, which is also law invariant, and in which the convexity axiom is required only for comonotone random variables. [5] 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. [48] propose *tradeable measures of risk*. They argue that the proper risk measures should be constructed by historically realized returns. Comparing to the coherent measures of risk, these risk measures appear 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.

1.3 Risk-Averse Inventory Models

In recent years, risk-averse inventory models have received increasing attention in the supply chain management literature. Table 1 classifies the literature by inventory models and risk measures. Because there is no research to date directly applying stochastic dominance to this field, we drop it from the table.

	Single-product, single-period	Single-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)
Coherent Measures of Risk	Gotoh & Takano (2007), Ahmed *et al.* (2007), Choi & Ruszczyński (2008), Chen *et al.* (2009)	Ahmed *et al.* (2007)	Ağrali & Soylu (2006), Tomlin & Wang (2005), this paper	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 VaR	Lau (1980)	None	None	Gan *et al.* (2005)

Table 1: Summary of the literature on risk-averse inventory models.

Most work to date dealt with single-product inventory models. For newsvendor models, research focused on finding the optimal solution under a risk-averse measure, and studying the impact of the degree of risk aversion (among 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. [11] review results in this direction.

For multi-product risk-averse newsvendor models, [59] 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.

Newsvendor networks are studied by [63], with many products and many resources under mean-variance and utility function approaches. The 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 most 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.

Finally, [1] conduct a numerical investigation on a two-product newsvendor model under the risk measure of CVaR. Assuming a discretized multi-variate 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. [37] study a manufacturer's pricing strategy and return policy under the mean-variance risk measure. [3] introduce a risk-neutral intermediaries to offer mutually beneficial contracts to risk-averse retailers. [61] studies how a manufacturer can use return policies to share risk under the mean-standard deviation measure. [25] study Pareto-optimality for suppliers and retailers under various risk-averse measures. [26] design coordination schemes of buyback and risk-sharing contracts in a supply chain under a Value-at-Risk constraint. For a review of the literature on risk aversion in capacity investment models and on operational hedging, see [62].

1.4 Our Model and Main Results

This paper considers a multi-product risk-averse newsvendor using a law-invariant coherent risk measure (see §2-3). As we argued in §1.2, coherent risk measures can be more attractive than the expected utility theory in the multi-product newsvendor problem due to their properties of Translation Equivariance and Positive Homogeneity.

The model presents a considerable challenge, both analytically and computationally, because the objective function cannot be decomposed by product and we have to look at the totality of all products as a portfolio. 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 [59] study a two-product system under CVaR, their focus is on the design of material flow topology and thus is different from ours.

We 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 study the asymptotic behavior of the system as the number

of products tends to infinity and obtain fast approximation for large-size problems.

This work contributes to literature in the following ways: We first establish several fundamental properties for the model (§4), e.g., the convexity of the model, the symmetry of the solution, and the impact of risk aversion. We then consider large but finite number of independent heterogenous products, for which we develop closed-form approximations (§5) which are exact in the single-product case. The approximations are as simple to compute as the risk-neutral solutions. We also show that under certain regularity conditions, the 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 demands under risk aversion poses a substantial analytical challenge. By utilizing the concept of associated random variables, we prove in §6 that in a risk-averse two-product model with positively dependent demands the optimal order quantities are lower than for independent demands, while for negatively dependent demands the optimal order quantities are higher. Using a sample-based optimization, we conduct in §7 a numerical study, which demonstrates that the approximations converge quickly to the optimal solutions as the number of products increases. It also provides additional insights into the impact of dependent demands. Specifically, we identify counterexamples to show that increased risk aversion can lead to greater optimal order quantities for strongly negatively dependent demands. In §8, we summarize the paper and compare the multi-product risk-averse newsvendor model to the financial portfolio problem.

2 Problem Formulation

Given products $j = 1, \dots, n$, let $x = (x_1, x_2, \dots, x_n)$ be the vector of ordering quantities and let $D = (D_1, \dots, D_n)$ be the demand vector. We also define $r = (r_1, \dots, r_n)$ to be the price vector, $c = (c_1, \dots, c_n)$ to be cost vector, and $s = (s_1, \dots, s_n)$ 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 $\bar{r}_j = r_j - s_j$, we can write the profit function as follows:

$$\Pi(x, D) = \sum_{j=1}^{n} \Pi_j(x_j, D_j), \tag{2.1}$$

where

$$\begin{aligned} \Pi_j(x_j, D_j) &= -\bar{c}_j x_j + \bar{r}_j \min\{x_j, D_j\} \\ &= (r_j - c_j)x_j - (r_j - s_j)(x_j - D_j)^+, \quad j = 1, \dots, n. \end{aligned} \tag{2.2}$$

We 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.

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

$$\max_{x \geq 0} \mathbb{E}[\Pi(x, D)]. \tag{2.3}$$

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. However, as we have mentioned it in the introduction, this formulation is inappropriate, if we are concerned with few (or just one) realizations and the Law of Large Numbers cannot be invoked.

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

$$\min_{x\geq 0} \rho[\Pi(x, D)], \tag{2.4}$$

where $\rho[\cdot]$ is a law-invariant coherent measure of risk, and $\Pi(x, D)$ represents the profit of the newsvendor, as defined in Eq. (2.1). It is worth stressing that problem (2.4) cannot be decomposed into independent subproblems, one for each product. Thus, it is necessary to consider the portfolio of products as a whole.

3 Coherent Measures of Risk

We present a formal definition of the coherent measures of risk following the abstract approach of [52, 53]. Let $(\Omega, \mathscr{F})$ be a certain measurable space. In our case, Ω is the probability space on which D is defined. An uncertain outcome (in our case, $\Pi(x, D)$) is represented by a measurable function $V : \Omega \to \mathbb{R}$. We specify the vector space $\mathscr{Z}$ of possible functions; in our case it is sufficient to consider $\mathscr{Z} = \mathscr{L}_\infty(\Omega, \mathscr{F}, P)$. Indeed, for a fixed order quantity x, the function $\omega \to \Pi(x, D(\omega))$ is bounded. For any V and $Y \in \mathscr{Z}$, we write $V \geq Y$ if $V \geq Y$ w.p.1.

In the minimization context, a *coherent measure of risk* is a functional $\rho : \mathscr{Z} \to \mathbb{R}$ satisfying the following axioms:
Convexity: $\rho(\alpha V + (1-\alpha)Y) \leq \alpha\rho(V) + (1-\alpha)\rho(Y)$, for all $V, Y \in \mathscr{Z}$ and all $\alpha \in [0, 1]$;
Monotonicity: If $V, Y \in \mathscr{Z}$ and $V \geq Y$, then $\rho(V) \leq \rho(Y)$;
Translation Equivariance: If $a \in \mathbb{R}$ and $V \in \mathscr{Z}$, then $\rho(V + a) = \rho(V) - a$;
Positive Homogeneity: If $t \geq 0$ and $V \in \mathscr{Z}$, then $\rho(tV) = t\rho(V)$.

A coherent measure of risk $\rho(\cdot)$ is called *law-invariant*, if the value of $\rho(V)$ depends only on the distribution of V, that is, $\rho(V_1) = \rho(V_2)$ if V_1 and V_2 have identical distributions.

Important examples of law-invariant coherent measures of risk are obtained from mean–risk models of form:

$$\rho(V) = -\mathbb{E}[V] + \lambda r[V], \tag{3.5}$$

where $\lambda > 0$ and $r[\cdot]$ is a variability measure of the random outcome V. Popular examples of functionals $r[\cdot]$ are the *semideviation* of order $p \geq 1$:

$$\sigma_p[V] = \mathbb{E}\big[\{(\mathbb{E}[V] - V)^+\}^p\big]^{\frac{1}{p}}, \tag{3.6}$$

and *weighted mean-deviation from quantile*:

$$r_\beta[V] = \min_{\eta\in\mathbb{R}} \mathbb{E}[\max\left((1-\beta)(\eta - V), \beta(V - \eta)\right)], \quad \beta \in (0, 1). \tag{3.7}$$

The optimal η in the problem above is the β-quantile of V. Optimization models with functionals (3.6) and (3.7) were considered in [44, 45, 46]. In the maximization context, from the practical point of view, it is most reasonable to consider $\beta \in (0, 1/2]$, because then $r_\beta[V]$ penalizes the left tail of the distribution of V much higher than the right tail.

The functional $\rho[\cdot]$ defined in Eq. (3.5), 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 (3.5) is a coherent measure of risk, if $\lambda \in [0, 1/\beta]$. All these results can be found in [53].

The mean-deviation from quantile $r_\beta[\cdot]$ is connected to the *Average Value-at-Risk* (AVaR), also known as *expected shortfall* or CVaR in [51], as follows,

$$\mathrm{AVaR}_\beta(V) = -\max_{\eta\in\mathbb{R}} \left\{\eta - \frac{1}{\beta}\mathbb{E}\big[(\eta - V)^+\big]\right\} = -\mathbb{E}[V] + \frac{1}{\beta} r_\beta[V]. \tag{3.8}$$

All these relations can be found in [46, 55], and [24] (with obvious adjustments for the sign of V). The relation (3.8) allows us to interpret $\mathrm{AVaR}_\beta(V)$ as a mean–risk model.

One of the fundamental results in the theory of law-invariant measures is the theorem of [34]: *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*

$$\rho[V] = \sup_{\mu \in \mathscr{M}} \int_0^1 \mathrm{AVaR}_\beta[V]\, \mu(d\beta). \tag{3.9}$$

Using identity (3.8) we can rewrite $\rho[V]$ as follows:

$$\rho[V] = -\mathbb{E}[V] + \sup_{\mu \in \mathscr{M}} \int_0^1 \frac{1}{\beta} r_\beta[V]\, \mu(d\beta). \tag{3.10}$$

This means that every problem (2.4) with a coherent law-invariant measure of risk is a mean–risk model, with the risk functional

$$\varkappa_{\mathscr{M}}[V] = \sup_{\mu \in \mathscr{M}} \int_0^1 \frac{1}{\beta} r_\beta[V]\, \mu(d\beta). \tag{3.11}$$

To illustrate the impact of scaling (the unit system) on risk measurement, we compare solutions of a single-product risk averse newsvendor model under the coherent risk measure (Eqs. (3.5) and (3.7)), the entropic exponential utility function $\frac{1}{\lambda} \ln \mathbb{E}\left[e^{-\lambda \Pi(x;D)}\right]$, which is an example of a convex measure of risk, and the mean-variance risk measure. We first set the unit of profit measurement to one dollar and choose parameters for each risk measure so that they yield an identical solution, i.e., the optimal order quantity, $\hat{x}^{RA} = 20.78$. Then we change the unit from a dollar to 30 cents, 10 cents, 3 cents, and 1 cent. We observe that while $\hat{x}^{RA}$ remains the same under the coherent risk measure, it changes from 20.78, to 17.09 (17.7), 12.29 (14.45), 7.29 (11.42) and 4.86 (9.56) under the entropic exponential (the mean-variance) measure. We refer to the Appendix for full details.

4 Analytical Results for Independent Demands

In this section, we provide several analytical results for the multi-product newsvendor model under coherent risk measures. We assume independent demands.

Proposition 1. *If $\rho[\cdot]$ is a coherent measure of risk, then $\rho[\Pi(x, D)]$ is a convex function of x.*

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

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

Using the monotonicity axiom, we obtain

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

The second inequality follows by the axiom of convexity.

Observe that we did not use the axiom of positive homogeneity, and thus Proposition 1 extends to more general convex measures of risk. We next prove the intuitively clear statement that identical products should be ordered in equal quantities under coherent measures of risk.

Proposition 2. *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 the optimal solutions of problem* (2.4) *is a vector with equal coordinates,* $\hat{x}_1^{\mathrm{RA}} = \hat{x}_2^{\mathrm{RA}} = \cdots = \hat{x}_n^{\mathrm{RA}}$.

Proof. Proof An optimal solution exists, because with no loss of generality we can assume that x is bounded by some large constant, and $\rho[\Pi(x, D)]$ is continuous with respect to x (see [53]).

Let us 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 1 and by the law-invariance of $\rho[\cdot]$ we obtain

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

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.

Note that Proposition 2 only requires symmetric joint demand distribution, but not independent demands. To study the impact of the degree of risk aversion, let us first focus on a specific variability functional – the weighted mean-deviation from quantile, given by Eq. (3.7). The corresponding measure of risk has the form,

$$\rho[V] = -\mathbb{E}[V] + \lambda r_\beta[V]. \tag{4.12}$$

By Eq. (3.8), we can write

$$\rho[V] = -(1 - \lambda\beta)\mathbb{E}[V] + \lambda\beta \mathrm{AVaR}_\beta[V]. \tag{4.13}$$

We consider the problem

$$\min_{x \geq 0} \Big\{ -\mathbb{E}[\Pi(x, D)] + \lambda r_\beta[\Pi(x, D)] \Big\}. \tag{4.14}$$

Proposition 3. *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* (4.14) *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* (4.14) *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. Proof In view of Proposition 2, we can assume that the coordinates of $\hat{x}^{\mathrm{RA}_i}$ are equal, $i = 1, 2$. Our argument extends that in Theorem 1 in [15] from the single-product case to the multi-product case.

Since all coordinates of the solutions are assumed equal, with a slight abuse of notation we consider a scalar decision variable x and we simplify Eqs. (2.1)–(2.2) to

$$\Pi(x, D) = nx(r - c) - (r - s) \sum_{j=1}^{n} (x - D_j)^+ = -n(c - s)x + (r - s) \sum_{j=1}^{n} \min(x, D_j). \tag{4.15}$$

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

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

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 Eq. (3.8) we obtain:

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

We shall differentiate both terms of the right hand side with respect to x. We 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, we 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, we suppress this dependence here for the ease of exposition. We consider two cases.

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

If $\hat{\eta}$ is unique, we 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 [8, Theorem 4.13]). Differentiating we obtain

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

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

$$P[\{Z_x < \hat{\eta}\} \cap \{D_j > x\}] = P[Z_x < \hat{\eta} | D_j > x] P[D_j > x] \le P[Z_x < \hat{\eta}] P[D_j > x] = \beta P[D_j > x].$$

The inequality holds true because

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

Thus

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

We conclude that

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

If $\hat{\eta}$ is not unique, we 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. We observe that the event $\{Z_x < \hat{\eta}\}$ does not change, and we conclude that the right derivative is non-negative.

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

As Z_x has an atom at nx, and $Z_x = nx$, for sufficiently small x we can just substitute $\hat{\eta} = nx$ in Eq. (3.8):

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

Taking derivative with respect to x, we 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, we consider the left derivative here, because if $\hat{\eta}(x) = nx$ then $\hat{\eta}(y) = ny$ for all $y < x$, and we arrive at the same conclusion.

We can extend the monotonicity property to all law-invariant coherent measures of risk. Observe that our assumption about continuous distribution of the demand implies that the probability space is nonatomic. Consider the problem

$$\min_{x \geq 0} \Big\{ - \mathbb{E}[\Pi(x, D)] + \lambda \varkappa_{\mathscr{M}}[\Pi(x, D)] \Big\}, \tag{4.16}$$

where $\varkappa_{\mathscr{M}}[V]$ is given by Eq. (3.11).

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* (4.16) *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* (4.16) *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. Proof As in the proof of Proposition 3, 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 Eq. (3.11) does not change this property. Therefore, Proposition 4 holds true also for the mean–risk model with the risk functional $r[\cdot] = \varkappa_{\mathscr{M}}[\cdot]$.

Finally, we discuss the impact of the shift in mean demand on the optimal order quantities under general coherent measures of risk. For this purpose, we consider identical products and demands with identical probability distribution except that $\mu_j = \mathbb{E}[D_j]$, $j = 1, \ldots, n$, may be different. Without loss of generality, we 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, by Proposition 2 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. We 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 by the Translation Equivariance axiom, it is easy to see that $\hat{x} = \tilde{x} + h$ is the solution of the problem

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

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

5 Asymptotic Analysis and Closed-Form Approximations

5.1 Asymptotic Optimality of Risk-Neutral Solutions

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

We 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. We assume that $s_j < c_j < r_j$, and that all these quantities are uniformly bounded for $j = 1, 2 \ldots$.

Consider the risk-neutral optimal order quantities

$$\hat{x}^{\mathrm{RN}}_{D_j} = \bar{F}^{-1}_{D_j}\Big(\frac{\bar{c}_j}{\bar{r}_j}\Big), \quad j = 1, 2, \ldots. \tag{5.17}$$

We assume that the following conditions are satisfied:

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

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

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

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

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

$$\min_{x_1,\ldots,x_n} \rho\Big[\frac{1}{n}\sum_{j=1}^{n} \Pi_j(x_j, D_j)\Big]. \tag{5.18}$$

Observe that in problem (5.18) we consider the average profit per product, rather than the total profit, as in problem (2.4). The reason is that we intend to analyze properties of the optimal value of this problem as $n \to \infty$ and we want the limit of the objective value of problem (5.18) to exist. Owing to the Positive Homogeneity axiom, problems (5.18) and (2.4) are equivalent.

We denote by $\hat{\rho}_n$ the optimal value of problem (5.18). We also introduce the following notation,

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

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

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

Proposition 5. *Assume that $\rho[\cdot]$ is a law-invariant coherent measure of risk and the space $(\Omega, \mathscr{F}, P)$ is nonatomic. Then*

$$\limsup_{n\to\infty} \frac{1}{\bar{s}_n}\left(\rho\Big[\frac{1}{n}\sum_{j=1}^{n} \Pi_j(\hat{x}^{\mathrm{RN}}_j, D_j)\Big] - \hat{\rho}_n\right) \le \rho[\mathscr{N}]. \tag{5.19}$$

Proof. Proof Denote $Z^n = \frac{1}{n}\sum_{j=1}^n \bar{r}_j \min(\hat{x}_j^{\mathrm{RN}}, D_j)$. We 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 $\{\bar{r}_j \min(\hat{x}_j^{\mathrm{RN}}, D_j)\}$, $j = 1, 2\ldots$, satisfies the Lindeberg condition (see, e.g, [22, p. 262]). We can therefore apply the Central Limit Theorem for non-identical independent random variables, to conclude that

$$\frac{Z^n - \bar{\mu}_n}{\bar{s}_n} \xrightarrow{\mathscr{D}} \mathscr{N}. \tag{5.20}$$

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 (5.18), we have

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

by Eq. (3.10) and because $\hat{x}_j^{\mathrm{RN}}$ maximizes $\mathbb{E}[\Pi_j(x_j, D_j)]$. Combining the last two relations we 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 we obtain

$$\frac{1}{\bar{s}_n}\left(\rho\Big[\frac{1}{n}\sum_{j=1}^n \Pi_j(\hat{x}_j^{\mathrm{RN}}, D_j)\Big] - \hat{\rho}_n\right) \leq \rho\left[\frac{Z^n - \bar{\mu}_n}{\bar{s}_n}\right]. \tag{5.21}$$

Let $\Phi_n(\cdot)$ be the cdf of $(Z^n - \bar{\mu}_n)/\bar{s}_n$. By Eq. (5.20), $\Phi_n \to \Phi$ 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, we 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 (5.21) tends to $\rho[\mathscr{N}]$ as $n \to \infty$. Passing to the limit in inequality (5.21), we obtain inequality (5.19).

Conditions (i) and (ii) imply that $\bar{s}_n = O(1/\sqrt{n})$, and thus it follows by inequality (5.19) that

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

Asymptotically, the difference between the optimal value of problem (5.18) 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.

5.2 Adjustments in the Mean–Deviation from Quantile Model

In this subsection, we develop closed-form approximations to the optimal risk-averse solution when the number of products is moderately large. Our idea is to use the risk-neutral solution as the starting point, and to calculate an appropriate correction to account for risk aversion.

We first consider the mean–deviation from quantile model in which the measure of variability is defined at Eq. (3.7). Recall that the corresponding mean–risk model in Eq. (4.12) is equivalent to the minimization of a combination of the mean and the Conditional Value-at-Risk, as in Eq. (4.13). We then consider the general coherent risk measure in §5.3. We finally discuss several iterative methods that are based on the approximations in §5.4.

Similarly to §5.1, we 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_x^n on x). Using Eqs. (2.1)–(2.2), we 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,

$$\begin{aligned}\rho[\bar{\Pi}(x, D)] &= \frac{1}{n}\sum_{j=1}^n \bar{c}_j x_j + \big(- \mathbb{E}[Z_x^n] + \lambda r_\beta(Z_x^n)\big) \\ &= \frac{1}{n}\sum_{j=1}^n \bar{c}_j x_j + \Big(\mathbb{E}[Z_x^n](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}} \Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x^n)^+]\Big\}\Big).\end{aligned} \tag{5.22}$$

Similarly to the proof of Proposition 3, let $\hat{\eta}$ be the maximizer in Eq. (5.22), 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, we consider two cases.

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 Eq. (5.22), we observe again that

$$\frac{\partial\rho[\bar{\Pi}(x, D)]}{\partial x_j} = \frac{\bar{c}_j}{n} + \frac{\bar{r}_j(\lambda\beta - 1)}{n} P[D_j > x_j] - \frac{\bar{r}_j\lambda}{n} P[\{Z_x^n < \hat{\eta}\} \cap \{D_j > x_j\}]. \tag{5.23}$$

Here we used [8, Theorem 4.13] to avoid differentiating with respect to $\hat{\eta}$.

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

$$\begin{aligned}P[\{Z_x^n < \hat{\eta}\} \cap \{D_j > x_j\}] &= P[Z_x^n < \hat{\eta} | D_j > x_j] P[D_j > x_j] \\ &= P\Big[\frac{1}{n}\sum_{k\neq j}^n \bar{r}_k \min(x_k, D_k) < \hat{\eta} - \frac{\bar{r}_j x_j}{n}\Big] \cdot P[D_j > x_j].\end{aligned} \tag{5.24}$$

Suppose $x_j \geq x_{\min}$, $j = 1, 2, \ldots,$. Owing to conditions (i) and (ii), exactly as in §5.1, 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 by $\hat{\eta} \simeq \bar{\mu}_n + z_\beta \bar{s}_n$, 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 $\mathcal{N}$ the standard normal random variable we

obtain:

$$P\Big[\frac{1}{n}\sum_{k\neq j}^{n}\bar{r}_k\min(x_k,D_k)<\hat{\eta}-\frac{\bar{r}_j x_j}{n}\Big]\simeq P\Bigg[\mathscr{N}<\frac{-\bar{r}_j(x_j-\mu_j)+z_\beta\sqrt{\sum_{k=1}^{n}\bar{r}_k^2\sigma_k^2}}{\sqrt{\sum_{k\neq j}\bar{r}_k^2\sigma_k^2}}\Bigg]$$
$$=P\Bigg[\mathscr{N}<\frac{-\bar{r}_j(x_j-\mu_j)}{\sqrt{n-1}\gamma_{nj}}+z_\beta\sqrt{1+\frac{\bar{r}_j^2\sigma_j^2}{(n-1)\gamma_{nj}^2}}\Bigg], \tag{5.25}$$

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, we conclude that γ_{nj} is bounded from above and below for all j and n.

This estimate can be put into Eq. (5.24) and thus Eq. (5.23) can be approximated as follows:

$$\frac{\partial\rho[\bar{\Pi}(x,D)]}{\partial x_j}\simeq\frac{\bar{c}_j}{n}+\frac{\bar{r}_jP[D_j>x_j]}{n}\Bigg(\lambda\beta-1-\lambda P\Bigg[\mathscr{N}<\frac{-\bar{r}_j(x_j-\mu_j)}{\sqrt{n-1}\gamma_{nj}}+z_\beta\sqrt{1+\frac{\bar{r}_j^2\sigma_j^2}{(n-1)\gamma_{nj}^2}}\Bigg]\Bigg). \tag{5.26}$$

Our next step is to approximate the probability on the right-hand side of Eq. (5.26). To this end, we derive its limit and calculate a correction to this limit for a finite n. When $n\to\infty$ we have

$$P\Bigg[\mathscr{N}<\frac{-\bar{r}_j(x_j-\mu_j)}{\sqrt{n-1}\gamma_{nj}}+z_\beta\sqrt{1+\frac{\bar{r}_j^2\sigma_j^2}{(n-1)\gamma_{nj}^2}}\Bigg]\to\beta \tag{5.27}$$

and thus

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

This means that the conditions of the risk-averse solution

$$\frac{\partial\rho[\bar{\Pi}(x,D)]}{\partial x_j}=0,\quad j=1,2,\ldots,n, \tag{5.28}$$

approaches that of the risk-neutral solution in Eq. (5.17). Thus the risk-neutral solution will be used as the base value, to which corrections will be calculated.

We can estimate the difference between the probability in Eq. (5.27) 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 $1/\sqrt{n-1}$, we obtain

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

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 $P[\mathscr{N}<z_\beta]=\beta$ and that its derivative at $z=z_\beta$ is the standard normal density at z_β, we get

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

with

$$\delta_{nj}^{\mathrm{RN}}=\frac{e^{-z_\beta^2/2}}{\sqrt{2\pi}}\frac{\bar{r}_j(\hat{x}_j^{\mathrm{RN}}-\mu_j^{\mathrm{RN}})}{\sqrt{n-1}\gamma_{nj}^{\mathrm{RN}}},\quad j=1,\ldots,n. \tag{5.29}$$

These estimates can be substituted to Eq. (5.26) for the derivative and yield

$$\frac{\partial \rho[\bar{\Pi}(x,D)]}{\partial x_j} \simeq \frac{\bar{c}_j}{n} + \frac{\bar{r}_j}{n}(-1 + \lambda \delta_{nj}^{\mathrm{RN}}) P[D_j > x_j]. \tag{5.30}$$

Using the above approximations of the derivatives in Eq. (5.28), we obtain the first-order approximation of the risk-averse solution:

$$\hat{x}_j^{\mathrm{RA}} \simeq \bar{F}_{D_j}^{-1}\Big[\frac{\bar{c}_j}{\bar{r}_j(1-\delta_{nj}^{\mathrm{RN}}\lambda)}\Big], \quad j = 1, 2, \ldots, n. \tag{5.31}$$

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

We 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,

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

Equating the right-hand side to 0, we get

$$\hat{x}_j^{\mathrm{RA}} = \bar{F}_{D_j}^{-1}\Big(\frac{\bar{c}_j + \bar{r}_j\lambda(1-\beta)}{\bar{r}_j(1+\lambda(1-\beta))}\Big). \tag{5.32}$$

Note that the solution in Case (ii) does not depend on the number of products, n. 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 $j = 1, 2, \ldots, n$. 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 $\delta_n^{\mathrm{RN}} = \dfrac{e^{-z_\beta^2/2}}{\sqrt{2\pi}} \dfrac{\hat{x}^{\mathrm{RN}} - \mu_x^{\mathrm{RN}}}{\sqrt{n-1}\sigma_x^{\mathrm{RN}}}$, 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}, respectively. Equating the right hand side to 0, we obtain

$$\hat{x}_1^{\mathrm{RA}} \simeq \bar{F}_{D_1}^{-1}\Big(\frac{\bar{c}}{\bar{r}(1-\delta_n^{\mathrm{RN}}\lambda)}\Big), \quad j = 1, \ldots, n. \tag{5.33}$$

Eq. (5.33) is similar to Eq. (5.31) except that the terms $\bar{c}_j$, $\bar{r}_j$ and $\delta_{nj}^{\mathrm{RN}}$ are now identical for all j. In Case (ii), Eq. (5.32) reduces to

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

In the special case of a single-product, by Eq. (5.23) in Case (i) we obtain

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

where $Z_x = \min(x, D)$. Observe that in Case (i), $P[\{Z_x < \hat{\eta}\} \cap \{D > x\}] = P[Z_x < \hat{\eta} | D > x]P[D > x] = 0$. 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

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

This special case solution is also obtained by [28]. 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.

5.3 General Law-Invariant Coherent Measures of Risk

So far our analysis focused on a special risk measure, weighted mean-deviation from quantile, given in Eq. (3.7). We now generalize the results to any law-invariant coherent risk measure $\rho[\cdot]$.

Consider problem (4.16) where $\varkappa_{\mathscr{M}}[V]$ is given by Eq. (3.11). By Kusuoka theorem, for nonatomic spaces, every law-invariant coherent measure of risk has such representation. Then Eq. (5.22) can be replaced by

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

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 Eq. (5.30),

$$\frac{\partial\rho[\bar{\Pi}(x,D)]}{\partial x_j} \simeq \frac{\bar{c}_j}{n} + \frac{\bar{r}_j}{n}\left(-1 + \lambda\int_0^1 \delta_{nj}^{\mathrm{RN}}(\beta)\,\hat{\mu}(d\beta)\right) P[D_j > x_j]. \tag{5.34}$$

We denote here the quantity given in Eq. (5.29) by $\delta_{nj}^{\mathrm{RN}}(\beta)$, to stress its dependence on β. Let us approximate $\hat{\mu}$ by the measure $\hat{\mu}^{\mathrm{RN}}$, obtained for the risk-neutral solution $\hat{x}^{\mathrm{RN}}$. Equating the approximate derivatives in Eq. (5.34) to zero, we obtain an approximate solution:

$$\hat{x}_j^{\mathrm{RA}} \simeq \bar{F}_{D_j}^{-1}\left(\frac{\bar{c}_j}{\bar{r}_j\Big(1 - \lambda\int_0^1 \delta_{nj}^{\mathrm{RN}}(\beta)\,\hat{\mu}^{\mathrm{RN}}(d\beta)\Big)}\right), \quad j = 1, 2, \ldots, n. \tag{5.35}$$

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 5.

In the single-product problem, we obtain

$$\rho[\bar{\Pi}(x,D)] = \bar{c}x + \sup_{\mu\in\mathscr{M}} \int_0^1 \Big(\mathbb{E}[Z_x](\lambda\beta - 1) - \lambda\beta \max_{\eta\in\mathbb{R}} \Big\{\eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)^+]\Big\}\Big)\mu(d\beta). \tag{5.36}$$

Assuming that $\hat{\mu}$ is the unique maximizer in (5.36), we obtain

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

Similarly to the model with mean-deviation from quantile case, $P[\{Z_x < \hat{\eta}\} \cap \{D > x\}] = P[Z_x < \hat{\eta} | D > x]P[D > x] = 0$. Thus,

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

Therefore, the closed-form exact solution for general coherent measures of risk is given by:

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

5.4 Iterative Methods

So far, we discussed 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. We explain the simplest idea for the approximation developed in §5.2; the same idea applies to general coherent measures of risk discussed in §5.3.

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 we calculate $\hat{x}^{(1)}$ by applying Eq. (5.31). In the iteration $\nu = 1, 2, \ldots$, we use $\hat{x}^{(\nu)}$ instead of $\hat{x}^{\mathrm{RN}}$ in our approximation, calculating:

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

Finally, Eq. (5.31) 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 when the risk aversion coefficient $\kappa = \lambda\beta$ is close to zero or the number of products is very large. We must point out that the iterative method does not guarantee convergence to the optimal risk-averse solution. One reason is that our approximation in Eq. (5.31) may result in infeasible solutions as the term $\dfrac{\bar{c}_j}{\bar{r}_j(1-\delta_{nj}^{(\nu)}\lambda)}$ can be negative or greater than 1 (due to approximation). When this occurs less likely, we say that the approximation is more stable. Generally, the approximation is more stable for larger number of products and smaller κ. To improve stability, we propose a more accurate method called the *continuation method*. In this approach, we apply the iterative method for a small value of κ, starting from the risk-neutral solution. Then we increase κ a little, and we apply the iterative method again, but starting from the best solution found for the previous value of κ. In this way, we gradually increase κ, until we reach the risk aversion coefficients which are of interest (usually, between 0 and 1). The stability of the iterative and continuation methods is summarized in §7.2.

6 Impact of Dependent Demands

In this section, we provide some insights on the impact of dependent demands. Due to significant analytical challenges, we focus on a two-product system and the mean-deviation from quantile model.

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 demands entail larger (smaller) variability and thus increase (decrease) risk, as compared to independent demands. Thus, one tends to decrease (increase) the order quantity in case of positively (negatively) dependent demands relative to the case of independent demand.

To characterize the impact of demand dependence on the optimal order quantity under the coherent risk measure, we 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 [19]; see [60] for a review.

Definition 1. *The random variables $D_1, D_2, \ldots, D_n$ are associated, if $Cov[f(D), g(D)] \geq 0$, or, equivalently, $\mathbb{E}[f(D)g(D)] \geq \mathbb{E}[f(D)]\mathbb{E}[g(D)]$, 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 1. (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) *If $D_1, D_2, \ldots, D_n$ are associated, then for all $(y_1, y_2, \ldots, y_n) \in R^n$*

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

We refer to [60] for proofs.

Association is closely related to correlation. By [60, p. 99], a set of multi-variate normal random variables is associated if their correlation matrix has the structure l ([60, p. 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 we can represent the demands as having one common factor:

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

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

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

$$\rho(\Pi(x, D)) = 2\overline{c}x + \overline{r}\rho(Z_x), \tag{6.37}$$

$$\rho(Z_x) = E(Z_x)(\lambda\beta - 1) - \lambda\beta \max_{\eta \in \mathbb{R}} \left\{ \eta - \frac{1}{\beta}\mathbb{E}[(\eta - Z_x)^+] \right\}, \tag{6.38}$$

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 3, we obtain

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

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, we focus on the second term, which is affected by the dependence of D_1 and D_2. We have

$$\frac{d\mathbb{E}[(\hat{\eta} - Z_x)^+]}{dx} = -\sum_{j=1}^{2} P[\{Z_x < \hat{\eta}\} \cap \{D_j > x\}] = -2P[\min\{x, D_2\} < \hat{\eta} - x, D_1 > x]. \tag{6.39}$$

Consider three cases of (D_1, D_2), with the same marginal distributions of D_1 and D_2. In case 1, (D_1, D_2) are associated random variables, and we use $\hat{\eta}_P$ to denote 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 Z_x; In case 3, $(D_1, -D_2)$ are associated random variables with $\hat{\eta}_N$ as the β-quantile of Z_x. We also let x_P^*, x_I^* and x_N^* be the optimal order quantities in cases 1, 2, and 3, respectively.

Proposition 6. *If $\hat{\eta}_P \le \hat{\eta}_I \le \hat{\eta}_N < 2x$, then*

$$x_P^* \le x_I^* \le x_N^*. \tag{6.40}$$

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

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

$$\begin{aligned}
&P[\min\{x, D_2\} < \hat{\eta}_P - x, D_1 > x] = P[D_2 < \hat{\eta}_P - x, D_1 > x] \\
&= P[D_1 > x] - P[D_2 \ge \hat{\eta}_P - x, D_1 > x] \le P[D_1 > x] - P[D_2 \ge \hat{\eta}_P - x]P[D_1 > x] \\
&= P[D_2 < \hat{\eta}_P - x]P[D_1 > x] \le P[D_2 < \hat{\eta}_I - x]P[D_1 > x].
\end{aligned}$$

The first inequality follows by Lemma 1 part (iv). The second inequality follows by $\hat{\eta}_P \le \hat{\eta}_I$. Note that the last term corresponds to independent (D_1, D_2). Thus, by Eq. (6.39), 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 $x_P^* \le x_I^*$.

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

$$\begin{aligned}
P[D_2 < \hat{\eta}_N - x, D_1 > x] &= P[-D_2 > -\hat{\eta}_N + x, D_1 > x] \ge P[-D_2 > -\hat{\eta}_N + x]P[D_1 > x] \\
&= P[D_2 < \hat{\eta}_N - x]P[D_1 > x] \ge P[D_2 < \hat{\eta}_I - x]P[D_1 > x].
\end{aligned}$$

The first inequality follows by Lemma 1 part (iv). The second inequality follows by $\hat{\eta}_I \le \hat{\eta}_N$. Note that the last term corresponds to independent (D_1, D_2). Thus, by Eq. (6.39), associated $(D_1, -D_2)$ have the derivatives $d\rho(Z_x)/dx$ no larger than independent (D_1, D_2), which implies that $x_I^* \le x_N^*$.

The condition $\hat{\eta}_P \le \hat{\eta}_I \le \hat{\eta}_N$ holds when $Y_1 = \min\{x, D_1\}$ and $Y_2 = \min\{x, D_2\}$ follow bivariate normal distribution and $\beta \le 0.5$. 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 $c - s$.

7 Numerical Examples

The objective of this section is two-fold. First, we study the accuracy and the convergence rates of the approximations. Second, we provide insights (in addition to the analysis in §§4-6) on the impact of demand dependence and risk aversion. We first introduce the sample-based optimization method.

7.1 Sample-Based Optimization

In all examples considered we apply sample-based optimization to solve the resulting stochastic programming problems. We 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 we replace the original demand distribution by the empirical distribution based on the sample, that is, we 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 [57]). In all our examples we used $T = 10,000$.

For the empirical distribution, the corresponding optimization problem (4.14) 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 ($0 \le \kappa \le 1$). We obtain the formulation

$$\begin{aligned}
\max \quad & (1-\kappa)\sum_{j=1}^{n}\Big[(r_j - c_j)x_j - (r_j - s_j)\sum_{t=1}^{T} p_t w_{jt}\Big] + \kappa\Big(\eta - \frac{1}{\beta}\sum_{t=1}^{T} p_t u_t\Big) \qquad (7.41)\\
\text{subject to} \quad & \sum_{j=1}^{n}\Big[(r_j - c_j)x_j - (r_j - s_j)w_{jt}\Big] + u_t \ge \eta, \quad t = 1, \ldots, T,\\
& x_j - d_{jt} \le w_{jt}, \quad j = 1, \ldots, n; \quad t = 1, \ldots, T,\\
& w_{jt} \ge 0, \quad j = 1, \ldots, n; \quad t = 1, \ldots, T,\\
& u_t \ge 0, \quad t = 1, \ldots, T,\\
& x_j \ge 0, \quad j = 1, \ldots, n.
\end{aligned}$$

To explain this formulation, 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 we maximize with respect to η the last term in problem (7.41), 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 we used Eq. (3.8). Therefore, Eq. (7.41) is equal to $(1-\kappa)\mathbb{E}[\Pi(x, D)] - \kappa\mathrm{AVaR}_\beta[\Pi(x, D)]$.

7.2 Accuracy of Approximations

In this subsection, we assess the accuracy of the closed-form approximations of §5. We first consider identical products, then non-identical products.

For identical products, we assume that all products have identical cost structure, and iid demands. We set $r = 15$, $c = 10$ and $s = 7$. We 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.

We choose the number of products, n, to be 1, 3, 10 and 30, and we 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 §5.4). The sample-based LP solutions can take hours to solve, especially for

large n and T. For instance, with $n = 30$ and a sample size of 10,000, the running time by CPLEX 9.0 at an Intel Pentium 4 PC is 32,607 seconds for identical products and 50,889 seconds for heterogenous products. In contrast, the approximate solution can be obtained within one or two seconds. We use $\beta = 0.5$, that is, we are concerned with the shortfall below the median.

In our numerical study of identical products, we set the optimal order quantities for different products to be identical by Proposition 2. In model (7.41) all variables x_j are replaced by a single variable x. The corresponding results are illustrated in Figure 1, where on the horizontal axis we display the relative risk aversion parameter $\kappa = \lambda\beta$. The term "exact", "numerical" and "approximation" represent the solution obtained by the exact calculation, the sample-based LP, and the closed-form approximation, respectively.

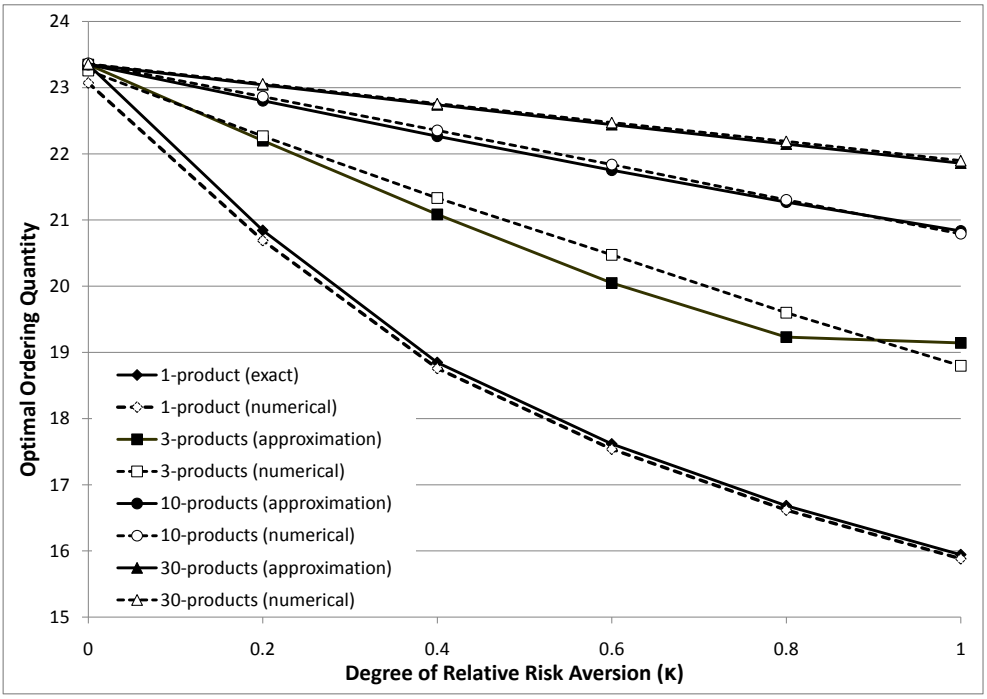

Figure 1: Identical products with independent demands – Approximate or exact solutions *vs.* sample-based solutions. The terms "exact," "numerical," and "approximation" refer to exact solutions, solutions of the sample-based model, and closed-form approximations, respectively.

Figure 1 shows that our analytical solution is very close to the numerical solution when $n = 1$. This is obvious as our solution is exact for the single-product case (here the case $\hat{\eta} = x$ is valid). In the case of a 3-product model, 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 increases, our approximations become more accurate and the gap becomes negligible when $n \geq 10$. We also observe that the order quantities decrease as the degree of risk-aversion increases, which confirms Proposition 3; and as the number of products increases, the error of the risk neutral solution decreases (consistent with Proposition 5).

For independent but heterogenous products, we tested the accuracy of the approximations on 30 randomly generated problems, 10 for each number of products $n = 3, 10, 30$. At each value of $\kappa = 0.2, 0.4, 0.6, 0.8, 1$, we calculated the sample-based LP solution and an approximate solution by the continuation method with $\nu = 1$. Our numerical study shows that the continuation method is much more stable and accurate than 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 in which the continuation method can generate a feasible solution, we 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, we 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, we 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 Figures 2 and 3, respectively).

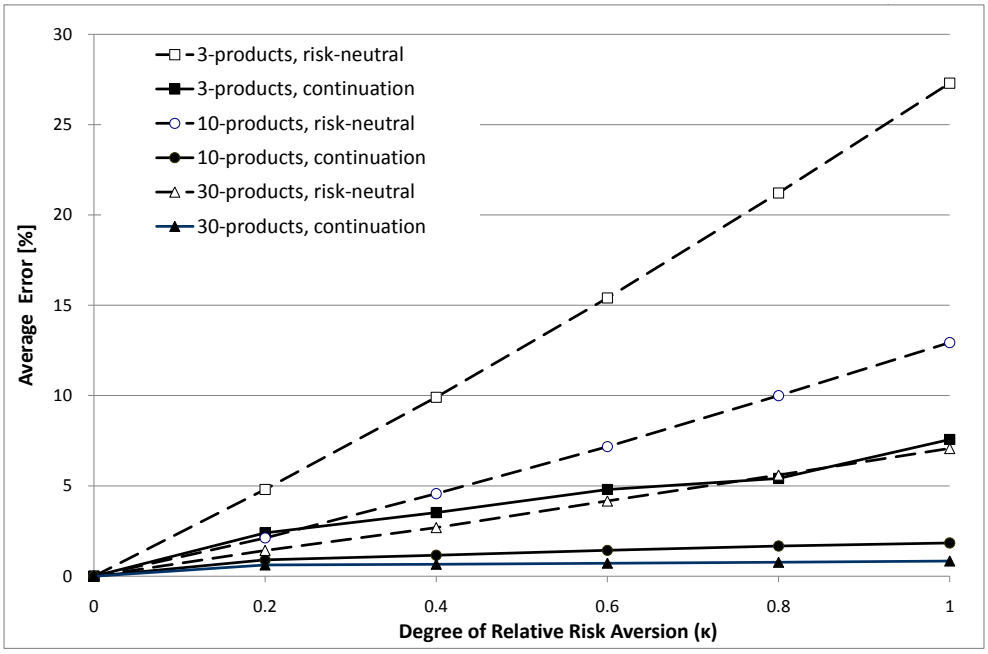

Figure 2: Heterogeneous products with independent demands – The average percentage error of the approximate solutions and risk-neutral solutions.

In all cases, in terms of the average and maximum errors, our approximation outperforms the risk-neutral solution. Furthermore, in most cases, the improvement brought by our approximation is significant. Often, the approximation cuts the error of the risk-neutral solution by 3 to 6 times, although only one step of the continuation method was made at each κ. Second, we observe that the approximation is quite accurate for all cases of $n = 10$ and $n = 30$. However, the approximation does not work well for $n = 3$, which is similar to what we observed in the identical products case. Finally, we observe that the average and maximum errors of the risk neutral solutions are decreasing in n, as established in Proposition 5.

7.3 Impact of Dependent Demands under Risk Aversion

We first consider a simple system with two identical products, then a system with two heterogenous products. The numerical results here are obtained by the sample-based LP.

We choose the following cost parameters for the system with two identical products: $r_1 = r_2 = 15$, $c_1 = c_2 = 10$ and $s_1 = s_2 = 7$. We assume that demand follows bivariate lognormal distribution, which is generated by exponentiating a bivariate normal with the parameters $\mu_1 = \mu_2 = 3$, $\sigma_1 = \sigma_2 = 0.4724$

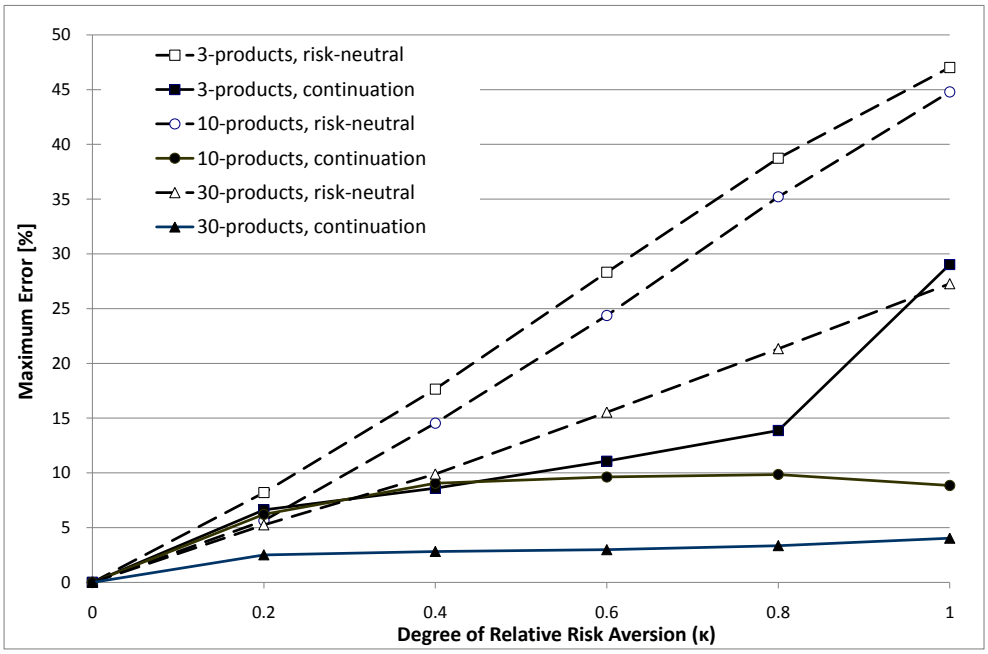

Figure 3: Heterogeneous products with independent demands – The maximum percentage error of the approximate solutions and risk-neutral solutions.

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 4.

Consistent with our analysis in §4, risk aversion reduces optimal order quantities for independent or positively correlated demands, relative to the risk-neutral solution. But interestingly, this observation may not hold for strongly negatively correlated demands, where increased risk aversion can result in a greater optimal order quantity. 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, increases negative correlation between the sales $\min(D_1, Q)$ and $\min(D_2, Q)$, and thus leads to smaller variability of the total sales $\min(D_1, Q) + \min(D_2, Q)$.

Figure 4 also shows that consistent with our analysis in §6, negatively correlated demands result in higher optimal order quantities than independent demands under risk aversion, while positively correlated demand leads to lower optimal order quantity under risk aversion. Indeed, the impact of demand correlation is almost monotonic with small deviations due to random sample errors.

These observations imply that if the firm is risk-averse, then demand dependence can have a significant impact on its optimal order quantities. They agree with the intuition that stronger positively (negatively) correlated demands indicate higher (lower) risk, and therefore lead to lower (higher) order quantities. More interestingly, while in most cases, the order quantity decreases in the degree of risk aversion, it can increase when the demands are strongly negatively correlated.

For heterogenous products, we 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 is bivariate lognormal 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 distributions of products 1 and 2 have means 22.46 and 44.913, standard deviations 11.23 and 89.826, and cv's 0.5 and 2, respectively. Intuitively, product 1 is

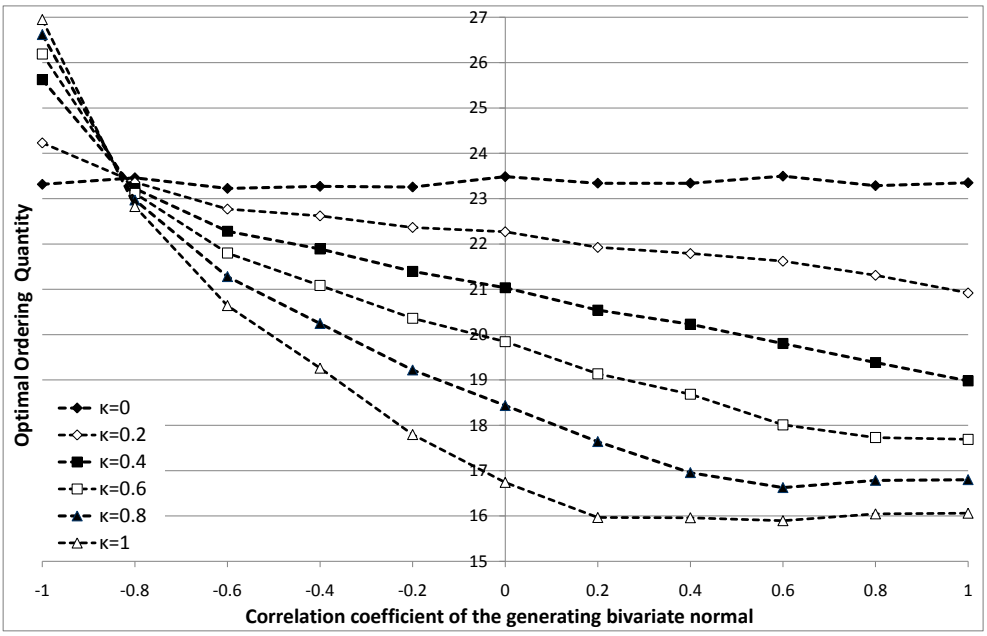

Figure 4: Identical products with dependent demands – The impact of demand correlation and risk aversion κ.

less risky and less profitable than product 2.

Our numerical study shows that for product 1, the impact of demand correlation is similar to that for identical products; see Figure 5. For product 2, however, the optimal ordering quantity always decreases in κ but not in correlation, see Figure 6.

The 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 of the more risky and more profitable products. 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.

For more details on the numerical study, we refer the readers to [14].

8 Conclusions

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. Our analytical results focus on the impact of risk aversion and demand dependence on the optimal order quantities. We analyze the asymptotic behavior of the optimal risk-averse solution and derive in Eq. (5.31) and in Eq. (5.35), simple and accurate approximations of the optimal order quantities for a large number of products with independent demands, and for general law-invariant coherent measures of risk. Our numerical study confirms the accuracy of these approximations for the numbers of products as small as 10, and enriches our understanding of the interplay of demand dependence and risk aversion.

It is perhaps appropriate to conclude this paper by comparing the multi-product risk-averse newsvendor problem (2.4) to the risk-averse portfolio optimization problem. In a portfolio problem, we have n assets with random returns $R_1, \ldots, R_n$ and our objective is to determine investment quantities $x_1, \ldots, x_n$ to

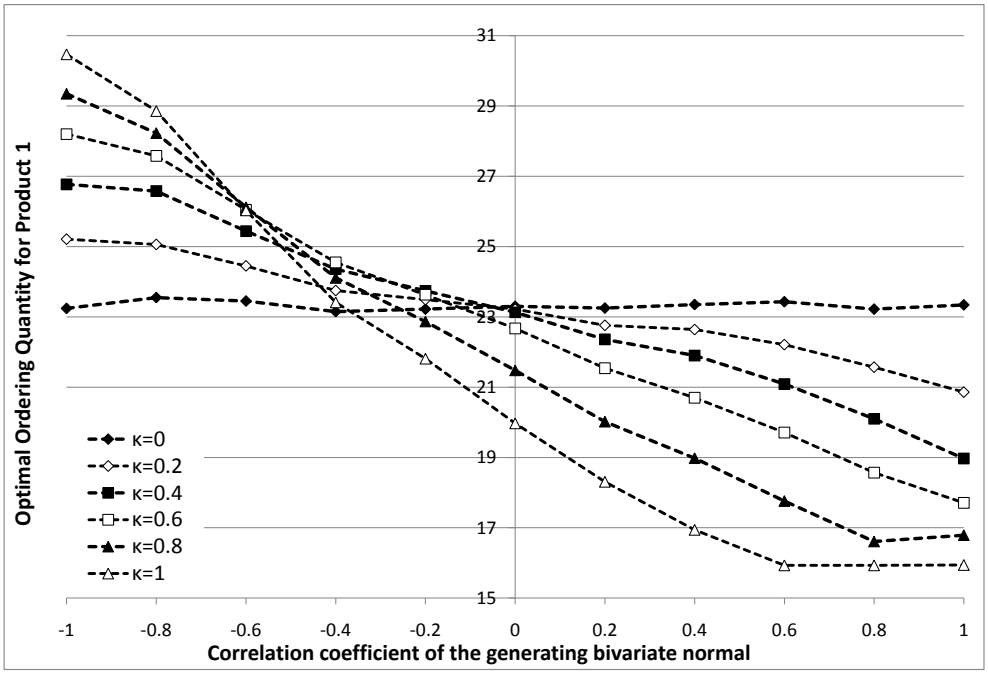

Figure 5: Heterogenous products with dependent demands – The impact of demand correlation and risk aversion κ for the product with low risk and low profit.

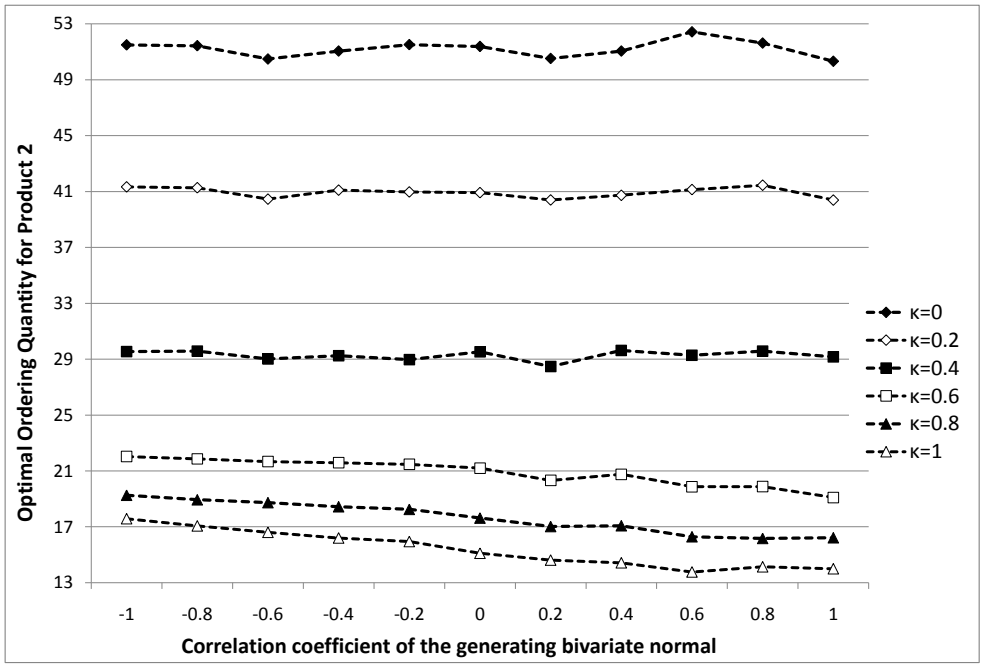

Figure 6: Heterogenous products with dependent demands – The impact of demand correlation and risk aversion κ for the product with high risk and high profit.

obtain 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 [40], the mean of the return and its variance are used to find efficient portfolio allocations. See also [21]. In more modern approaches (e.g., [33, 55, 42]) more general mean–risk models and coherent measures of risk are used, similarly to problem (2.4). 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 we restrict the total amount invested (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.

- 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 problem 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. Products are unlikely to be eliminated because of 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 problem the risk-neutral solution is asymptotically optimal under risk aversion, when the number of independent products approaches infinity.

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 function (2.4), which involves two nondifferentiable functions.

Acknowledgement: The second author was partially supported by the NSF Awards DMS-0603728 and CMII-0965689. The third author is supported in part by the NSF Award CMMI-0747779. The authors are very grateful to the Associate Editor and three anonymous Referees for their comments and suggestions that lead to significant improvements of the manuscript.

References

[1] Ağrali, S. and A. Soylu. 2006. Single Period Stochastic Problem with CVaR Risk Constraint. Presentation, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL.

[2] Agrawal, V. and S. Seshadri. 2000. Impact of Uncertainty and Risk Aversion on Price and Order Quantity in the Newsvendor Problem. *Manufacturing and Service Operations Management*, **2(4)**, 410-423.

[3] Agrawal, V. and S. Seshadri. 2000b. Risk Intermediation in Supply Chains. *IIE Transactions*, **32**, 819-831.

[4] Ahmed, S., U. Çakmak and A. Shapiro. 2007. Coherent Risk Measures in Inventory Problems. *European Journal of Operational Research*, **182**, 226-238.

[5] Ahmed, S., D. Filipović and G. Svindland. 2008. A note on natural risk staistics, *Operations Research Letters*, **36**, 662-664.

[6] 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.

[7] Artzner, P., F. Delbaen, J. Eber and D. Heath. 1999. Coherent Measures of Risk. *Mathematical Finance*, **9(3)**, 203-228.

[8] Bonnans, J. F. and A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, NY.

[9] Bouakiz, M. and M. Sobel. 1992. Inventory Control with an Exponential Utility Criterion. *Operations Research*, **40(3)**, 603-608.

[10] Chen, F. and A. Federgruen. 2000. Mean-Variance Analysis of Basic Inventory Models. Working paper, Division of Decision, Risk and Operations, Columbia University, New York, NY.

[11] Chen, X., M. Sim, D. Simchi-Levi and P. Sun. 2007. Risk Aversion in Inventory Management. *Operations Research*, **55**, 828-842.

[12] Chen, Y., M. Xu and Z. Zhang. 2009. A Risk-Averse Newsvendor Model Under the CVaR Criterion. *Operations Research*, **57(4)**, 1040-1044.

[13] Cheridito, P., F. Delbaen and M. Kupper. 2006. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, **11**, 57-106.

[14] Choi, S. 2009. Risk-Averse Newsvendor Models. PhD Dissertation, Department of Management Science and Information Systems, Rutgers University, Newark, NJ.

[15] Choi, S. and A. Ruszczyński. 2008. A Risk-Averse Newsvendor with Law Invariant Coherent Measures of Risk. *Operations Research Letters*, **36**, 77-82.

[16] 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.

[17] Delbaen, F. 2002. Coherent Risk Measures on General Probability Space. *Advance in Finance and Stochastics*, Springer, Berlin, Germany, 1-37.

[18] Dentcheva, D. and A. Ruszczyński. 2003. Optimization with Stochastic Dominance Constraints. *SIAM Journal on Optimization*, **14(2)**, 548-566.

[19] Esary, J. D., F. Proschan, and D. W. Walkup. 1967. Association of Random Variables, with Applications. *Annals of Mathematical Statistics*, **38**, 1466-1474.

[20] Eeckhoudt, L., C. Gollier, and H. Schlesinger. 1995. The Risk-averse (and Prudent) Newsboy. *Management Science*, **41(5)**, 786-794.

[21] 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, NY.

[22] Feller, W. 1971. *An Introduction to Probability Theory and Its Applications*, Vol. 2, 3rd ed., Wiley, New York, NY.

[23] Föllmer, H. and A. Schied. 2002. Convex Measures of Risk and Trading Constraints. *Finance and Stochastics*, **6**, 429-447.

[24] Föllmer, H. and A. Schied. 2004. *Stochastic Finance: An Introduction in Discrete Time*, de Gruyter, Berlin, Germany.

[25] 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.

[26] 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.

[27] Gaur, V. and S. Seshadri. 2005. Hedging Inventory Risk Through Market Instruments. *Manufacturing and Service Operations Management*, **7(2)**, 103-120.

[28] Gotoh, J. and Y. Takano. 2007. Newsvendor Solutions via Conditional Value-at-risk Minimization. *European Journal of Operational Research*, **179**, 80-96.

[29] Hadar, J and W. Russell. 1969. Rules for Ordering Uncertain Prospects. *The American Economic Review*, **59**, 25-34.

[30] Hadley, G. and T.M. Whitin. 1963. *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, NJ.

[31] 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, Department of Statistics, Columbia University, New York, NY.

[32] Howard, R.A. 1988. Decision analysis: practice and promise. *Management Science*, **34**, 679-695.

[33] Konno, H. and H. Yamazaki. 1991. Mean–Absolute Deviation Portfolio Optimization model and its Application to Tokyo Stock Market. *Management Science*, **37**, 519-531.

[34] Kusuoka, S. 2003. On Law Invariant Coherent Risk Measures. *Advances in Mathematical Economics*, **3**, 83-95.

[35] Kusuoka, S. and Y. Morimoto. 2007. Homogeneous Law Invariant Coherent Multiperiod Value Measures and Their Limits. *The Journal of Mathematical Sciences*, University of Tokyo, **14(2)**, 117-156.

[36] Lau, H.S. 1980. The Newsboy Problem under Alternative Optimization Objectives. *The Journal of the Operational Research Society*, **31(6)**, 525-535.

[37] 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.

[38] Lehmann, E. 1955. Ordered Families of Distributions. *Annals of Mathematical Statistics*, **26**, 399-419.

[39] Levy, H. 2006. *Stochastic Dominance: Investment Decision Making Under Uncertainty*, 2nd Edition, Springer, New York, NY.

[40] Markowitz, H.M. 1959. *Portfolio Selection*, Wiley, New York, NY.

[41] Martinez-de-Albeniz, V. and D. Simchi-Levi. 2006. Mean-Variance Trade-offs in Supply Contracts. *Naval Research Logistics*, **53**, 603-616.

[42] Miller, N., and A. Ruszczyński. 2008. Risk-Adjusted Probability Measures in Portfolio Optimization with Coherent Measures of Risk. *European Journal of Operational Research*, **191**, 193-206.

[43] Müller, A., and D. Stoyan. 2002. *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons, Chichester, UK.

[44] Ogryczak, W. and A. Ruszczyński. 1999. From Stochastic Dominance to Mean-Risk Models. *European Journal of Operational Research*, **116**, 33-50.

[45] Ogryczak, W. and A. Ruszczyński. 2001. On Consistency of Stochastic Dominance and Mean-Semideviation Models. *Mathematical Programming*, **89**, 217-232.

[46] Ogryczak, W. and A. Ruszczyński. 2002. Dual Stochastic Dominance and Related Mean-Risk Models. *SIAM Journal on Optimization*, **13(1)**, 60-78.

[47] 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, Armsterdam.

[48] Pospišil, L., J. Večer and M. Xu. 2008. Tradeable Measures of Risk. Working Paper, Department of Statistics, Columbia University, New York, NY.

[49] Prékopa, A. 2003. Probabilistic Programming. In: Ruszczyński, A. and A. Shapiro (eds.) *Stochastic Programming*, Elsevier, Amsterdam, Netherlands.

[50] Riedel, F. 2004. Dynamic coherent risk measures. *Stoch. Processes and Their Appl.*, **112**, 185-200.

[51] Rockafellar, R.T. and S.P. Uryasev. 2000. Optimization of Conditional Value-at-Risk. *The Journal of Risk*, **2**, 21-41.

[52] 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, UK, 117-158.

[53] Ruszczyński, A. and A. Shapiro. 2006a. Optimization of Convex Risk Functions. *Mathematics of Operations Research*, **31(3)**, 433-452.

[54] Ruszczyński, A. and A. Shapiro. 2006b. Conditional Risk Mappings. *Mathematics of Operations Research*, **31(3)**, 544-561.

[55] Ruszczyński, A. and R. J. Vanderbei. 2003. Frontiers of Stochastically Nondominated Portfolios. *Econometrica*, **71(4)**, 1287-1297.

[56] 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.

[57] Shapiro, A. 2008. Asymptotics of Minimax Stochastic Programs. *Statistics and Probability Letters*, **78**, 150-157.

[58] Shapiro, A., D. Dentcheva, and A. Ruszczyński. 2009. *Lectures on Stochastic Programming: Modeling and Theory*, SIAM (Society for Industrial and Applied Mathematics), Philadelhpia, PA.

[59] 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.

[60] Tong, Y. L. 1980. *Probability Inequalities in Multivariate Distributions*, Academic Press, New York, NY.

[61] Tsay, A. A. 2002. Risk Sensitivity in Distribution Channel Partnerships: Implications for Manufacturer Return Policies. *Journal of Retailing*, **78**, 147-160.

[62] van Mieghem, J. A. 2003. Capacity management, investment and hedging: Review and recent developments. *Manufacturing Service Operations Management*, **5(4)**, 269-302.

[63] van Mieghem, J. A. 2007. Risk Mitigation in Newsvendor Networks: Resource Diversification, Flexibility, Sharing, and Hedging. *Management Science*, **53(8)**, 1269-1288.

[64] 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.

[65] von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

[66] Wang, S. S., V. R. Young and H. H. Panjer. 1997. Axiomatic Characterization of Insurance Prices. *Insurance: Mathematics and Economics*, **21**, 173-183.

Appendix

We compare solutions of a single-product newsvendor model under the coherent measure of risk, the entropic exponential utility function, and mean-variance risk measure. The entropic exponential utility function is an example of a convex measure of risk which is not coherent and is equivalent to an exponential utility function by a "certainty equivalent" operator.

We select parameters for each risk measure so that they have the same optimal solution when the unit of profit measurement is one dollar. Specifically, we set $r = 15$, $c = 10$ and $s = 7$ (in dollars) for all three risk measures. Demand follows a lognormal distribution with $\mu = 3$ and $\sigma = 0.4724$. This demand distribution is used in all instances. For the coherent measure of risk, we set $\beta = 0.5$ and $\lambda = \lambda_1 = 0.2$. By the sample-based LP method, the optimal solution is $\hat{x}^{\mathrm{RA}_1} = 20.7824$. For the entropic exponential utility function model, defined as

$$\min_{x\geq 0} \frac{1}{\lambda_2} \ln \mathbb{E}\Big[e^{-\lambda_2 \Pi(x;D)}\Big], \tag{8.42}$$

we set $\lambda_2 = 0.0072$, which results in a sample-based solution $\hat{x}^{\mathrm{RA}_2} = 20.7786$. For the mean-variance model, defined as

$$\min_{x\geq 0} -\mathbb{E}[\Pi(x;D)] + \lambda_3 \mathbb{V}\mathrm{ar}[\Pi(x;D)], \tag{8.43}$$

we set $\lambda_3 = 0.0037$, which results in a sample-based solution $\hat{x}^{\mathrm{RA}_3} = 20.7918$. Then we change the unit of r (price), c (cost) and s (salvage value) from dollar to 30 cents, 10 cents, 3 cents and 1 cent while keeping all other parameters unchanged. Our results are summarized in Table 2.

As we can see from this table, while the numerical solution under a coherent measure of risk is invariant with respect to the unit system, it varies significantly under other risk measures.

	Unit of Profit Measurement				
	1 Dollar	30 Cents	10 Cents	3 Cents	1 Cent
Coherent	20.7824	20.7824	20.7824	20.7824	20.7824
Entropic Exponential	20.7786	17.0952	12.2944	7.2879	4.8568
Mean-Variance	20.7918	17.6962	14.4454	11.4197	9.5603

Table 2: The impact of rescaling on solutions - a coherent measure of risk, entropic exponential utility function and a mean–variance model.