# Asymmetry and Ambiguity in Newsvendor Models

Karthik Natarajan* Melvyn Sim† Joline Uichanco‡

Submitted: July 14, 2008, First Revision: June 1, 2009

**Abstract**

The traditional decision making framework for the newsvendor model is to assume a distribution of the underlying demand. However, the probability distribution is itself often ambiguous. A more conservative approach is to assume that the distribution belongs to a set parameterized by a few known moments. An ambiguity-averse newsvendor would then choose to maximize the worst-case expected profit over this set of probability distributions. The earliest model of this type assumed that only the mean and variance of demand are known. While efforts have been made to model asymmetry by including skewness information, closed-form expressions remain difficult to find. In this paper, we propose to model asymmetry through the knowledge of a parameter: normalized semivariance ($s$). We provide a closed-form expression for the optimal order quantity with only mean, variance and semivariance information. This leads to simple rule-of-thumb ordering policies that guides the newsvendor in ordering above or below the mean depending on whether the critical fractile is above or below $\frac{1}{2}(1+s)$. We show that knowledge of asymmetry thorough this parameter significantly reduces the gap between the best- and worst-case bounds on the expected profit. We extend our results to the setting where the newsvendor is risk-averse and uses the Conditional Value-at-Risk measure. Assuming that only the mean and variance of demand are known, we develop a closed form optimal ordering quantity for the ambiguity-averse risk-averse newsvendor. The knowledge of asymmetry reduces the degree of conservatism even more significantly in this case. The model can also be applied to various extensions such as bounded support information and multiple demand partitions. In all these extensions, the problem can be reduced to solving a second-order cone program (SOCP).

## 1 Introduction

The single-period newsvendor problem is the foundation of many operations management models ranging from inventory control to the design of contracts for coordination of supply chains. In the classical

*Email: matkbn@nus.edu.sg. National University of Singapore, Department of Mathematics, NUS Risk Management Institute, Singapore-MIT Alliance.

†Email: dscsimm@nus.edu.sg. NUS Business School, NUS Risk Management Institute, Singapore-MIT Alliance. The research of the author was partially supported by NUS academic research grant R-314-000-068-122.

‡Email: uichanco@mit.edu. Massachusetts Institute of Technology, Operations Research Center. The research of the author was partially supported by the Singapore-MIT Alliance.

version, a newsvendor decides before the sales period how many units of a product to order. The actual demand occurs during the sales period and is satisfied as much as possible with the units on hand. The newsvendor incurs a cost $c$ for each ordered unit, and sells each unit for a price $p$. One common practice is to assume stochastic demand that follow a well-known distribution, such as Gaussian, Poisson or lognormal (Ridder et al. [21]). Alternatively, Song et al. [26] propose to fit demand data to the best distribution without forcing it to be Gaussian or Poisson. Based on the assumed demand distribution, the newsvendor then orders at a level that maximizes the expected profit, which is the $1 - \frac{c}{p}$ quantile of the demand distribution. However, in many practical situations, it may be impossible to elicit the exact distribution of the uncertain demand. For certain products, such as the antiviral drug, Tamiflu, there is often not enough stationary demand data to ascertain the form of the distribution. Without the exact distribution, one would not be able to determine the exact expected profit. How does the newsvendor decide on an ordering policy in this case?

Savage [23] supported the idea of subjective expected utility in which he argued that a rational decision maker would rank uncertain monetary outcomes based on taking expectation over a subjective probability measure on outcomes. Hence, where it is impossible to determine the exact demand distribution, Savage's approach would legitimize the practice of picking a subjective demand distribution and the newsvendor would order at the level that maximizes the subjective expected profit. However, Ellsberg [6] vehemently objected to Savage's notion of subjective expected utility. In the famous Ellsberg's paradox experiment, subjects play a game that rewards $1000 for picking a red ball from a box. They are told to pick either from the first box with 50 red and 50 black balls, or from the second box where the distribution of balls is unknown. Most subjects strictly prefer the first box in which the probability of winning the prize is 50%. In the second part of the experiment, subjects are rewarded the same amount for picking a black ball instead of red. Again, subjects strictly prefer the first box. Ellsberg argued that the experiment findings are inconsistent with the paradigm of subjective expected utility. Under this, individuals who strictly prefer the first box may perceive that in the second box, red balls are fewer in number than black ones. In doing so, they should prefer the second box in the second experiment. Indeed, Ellsberg's paradox can be resolved by considering a worst-case expected utility where the decision maker is ambiguity averse (see Gilboa and Schmeidler [9]). Föllmer and Schied [7] also argue that the worst-case or robust approach models ambiguity-averse behavior, which otherwise cannot be explained under the traditional expected utility framework.

We illustrate ambiguity aversion now in the context of inventory problems. Suppose a manager is only certain of a few properties of the demand. Rather than committing to a potentially wrong distribution and getting an uncertain (possibly low) payoff, he would assume that nature works against him and order accordingly. Doing so would at least assure him of a known minimum level of expected profit. The natural question that arises is how we can describe uncertainty in demand that encompasses the notion of distributional ambiguity. The most common approach is to specify a family of demand distributions having a set of known moments. The information about the moments may come from estimates using past realizations or some prediction by industry experts. The ambiguity averse or robust approach optimizes the worst-case objective (e.g. expected profit or risk) over the parametric family. Another version, called minimax regret, minimizes the maximum opportunity cost from not making the optimal decision (Savage [23]). Most research that uses the ambiguity averse approach in newsvendor models describes the distribution set by some known mean and variance (see Scarf [24]; Lo [13]; Gallego and Moon [8]; Perakis and Roels [18]). Due to their second-order nature, closed-form expressions for the optimal bounds have been found for most of these robust models. It is important to note that minimax regret captures a subset of individuals with specific behavioral choices driven by the fear of losing out on the best possible deal. In this study, we focus on the subset of individuals who are ambiguity averse and are driven by the fear of uncertainty.

A weakness of the mean-variance robust models is that it permits a very fairly large family of demand distributions. To scope down the size of the family and improve its resolution, the next logical extension is to impose on some asymmetry information in describing the demand distribution. We emphasize the importance of modeling asymmetry when the distribution is unknown through the following example. Suppose that the unit cost is \$200 and the unit price is \$300. Moreover, the demand mean is known to be 100 units with a standard deviation of 50, and is very strongly positively skewed. If the manager ignores asymmetry completely and assumes that the distribution is Gaussian, he would order 78 units. What is the smallest possible profit the manager obtains under any distribution satisfying the known information? In Section 2, we show that we can generate a positively skewed distribution which assures an expected profit of \$7,800 (see Figure 2.5). Instead, if he orders 96 units, the smallest possible profit scenario would be almost \$9,400. That is, by ignoring asymmetry information, he loses an assured profit of \$1,600.

Recent extensions attempt to incorporate asymmetry and other features into the model by assuming knowledge of higher moments, such as skewness or kurtosis (Jansen et al. [12]; De Schepper and Heijnen [25]; He et al. [10]; Zuluaga et al. [31]). Since the higher moments result in a problem with third and fourth order form, it is usually not easy to find a closed-form expression for the optimal bounds. Even if they are found, the expressions are complicated and provide limited managerial insight. An option is to incorporate first order measures of asymmetry such as the mode and median. These first order asymmetric measures are better understood and easier to predict by experts. However it forms a very weak measure of asymmetry as it does not look at the magnitude of the deviations above and below the mean. For example, it is possible for distributions have the median to the right or left of the mean and at the same time be right skewed (see von Hippel [28]).

Our approach in this paper is to represent asymmetry using a well-known measure: semivariance. The advantage of using semivariance in moment bounds is that asymmetry is introduced without needing to stray from the simplicity of second-order models. A consequence of this is that closed-form expressions are more easily derived yielding clearer insights into the optimal decision. The idea of using semivariance as a measure of asymmetry is not new. Semivariance as a measure of downside risk was put forward by Markowitz [15] in his seminal paper on portfolio selection. The semivariance of a random variable $\tilde{x}$ around it's mean $\mu$ is

$$\mathrm{E}\left((\mu-\tilde{x})_+^2\right),$$

where $x_+^2 = \max(0,x)^2$. The relative magnitude of semivariance compared to variance indicates how the deviations from the mean are split between the upper and lower parts of the distribution. Berck and Hihn [2] use semivariance to tighten Chebychev's inequality which is based on the mean and variance. They show that by using semivariance, a much sharper bound on the probability of the tail of a distribution can be found especially when the underlying distribution is asymmetric.

An important question remains: is it any easier to make accurate estimates of semivariance? Several approaches have been proposed to calculate semivariance. Methods by Sortino and Forsey [27] and Bond [3], for instance, calculate semivariance directly from some estimated density function. However, the simplest and most obvious approach is still through a sample-based calculation. Yet one common concern is that the volatility of the sample-based semivariance is so high as to make it impractical in applied work. In fact, Sortino and Forsey [27] are critical of using the sample-based semivariance, since they argue that it can easily over- or under-estimate the true semivariance due to its dependency on some target value. Instead, they suggest that fitting a continuous probability density function is superior to discrete sample calculations. However, in a recent paper by Bond and Satchell [4] which studies the statistical properties of the sample semivariance, it has been shown that sample semivariance is in fact less volatile than sample variance when the distribution is asymmetrical. Their results suggest that the major concern of practitioners against using sample semivariance is not valid. In our recent paper on robust portfolio optimization [16], we capture distributional asymmetry of stock returns using partitioned covariance, which is essentially the extension of semivariance to multivariate random variables. The computational tests over ten years of real financial data suggest that the robust portfolios that take into account of asymmetry yield efficient frontiers that dominate portfolios found using empirical distributions and portfolios that are mean-covariance robust.

It is crucial to make the distinction between "ambiguity-aversion" and "risk-aversion". Risk aversion implies that a person prefer to make a sure average amount as compared to the possibility of making either huge losses or huge profits. The concept of risk aversion is captured using a risk measure on the uncertain payoffs. On the other hand, ambiguity aversion refers to the preference of a sure payoff over something unsure but with a potentially higher payoff. As we saw earlier, ambiguity aversion implies making a decision under the worst-case distribution. It is thus possible to consider a newsvendor who is both risk averse and ambiguity averse by taking the risk measure with respect to the worst-case distribution. As a matter of fact, ambiguity aversion has already become an integral part of modern convex risk measures initiated by and popularized in the mathematical finance community (see Föllmer and Schied). Our contributions can be classified as follows:

1. *Modeling asymmetry through normalized semivariance:* We propose to model asymmetry of a distribution through a single parameter: normalized semivariance ($s$), whose magnitude corresponds to the degree of asymmetry. A positive or negative $s$ roughly relates to a positively or negatively skewed distribution. We show that under a risk-neutral setting, the robust mean-variance-semivariance policies have a simple closed-form expression. Moreover, many extensions of these models (e.g. risk-aversion, bounded support, multiple partition of demand) can be converted into second-order cone programs (SOCPs) that can easily be solved by modern solvers.

2. *Ordering rule-of-thumb for mean-variance (MV) and mean-variance-semivariance (MVS) robust policies:* Scarf [24] shows that a risk-neutral, ambiguity-averse newsvendor under the mean-variance setting would decide to order above or below the mean depending on whether critical fractile $1-\frac{c}{p}$ is above or below $\frac{1}{2}$. Under the MVS framework, we have an alternate rule-of-thumb: the newsvendor would order above or below the mean depending on whether this ratio is above or below $\frac{1}{2}(1+s)$. We also provide a rule-of-thumb for a risk-averse newsvendor whose risk parameter is $\alpha$ under a Conditional Value-at-Risk measure. If he is ambiguity-averse and only knows the mean and variance, he will choose to order above or below the mean depending on whether the critical fractile is above or below $\frac{1}{2(1-\alpha)}$.

3. *Reduction in best- and worst-case expected profit gap:* Under the MV setting, the best- and worst-case expected profit gap can be very large. However, we show that we can reduce this gap by including asymmetry information into the model, especially if the degree of asymmetry is large. We also show that in many cases, MV policies can be conservative due to the fact that the expected profits they achieve are much less than the optimal profit under the MVS setting.

4. *Less conservative ordering policies under the MVS model:* The risk neutral MV order policies are criticized to be conservative since they suggest ordering nothing for a wide range of $\frac{c}{p}$ ratios. We show that by including asymmetry information, the robust policies are much less conservative. This improvement is even more apparent for a risk-averse newsvendor, since for common risk preference parameters, the MV models suggest ordering nothing. However, under the MVS framework, the optimal order quantity is positive if the degree of asymmetry is large enough and known.

5. *Uncertainty in estimates of semivariance:* We provide an additional degree of freedom in the asymmetry model by assuming that the semivariance may not be known exactly. Instead, the range of values for the semivariance may be known. We show that this model can be converted into an SOCP.

The structure of the paper is as follows. We discuss the risk-neutral newsvendor model in Section 2 and introduce our asymmetry model. In Section 3, we consider a risk-averse newsvendor and develop methods to find optimal policies in this setting. In Appendix A, we generalize the model to multiple partitions and a piecewise linear objective. The proofs are provided in Appendix B.

## 2 Ambiguity-Averse, Risk-Neutral Newsvendor

Consider a newsvendor facing a random demand $\tilde{d}$ for the product observed during the sales period. He satisfies the demand as much as possible with the units he has preordered. Any unmet demand is assumed to be lost. Let $c$ be the unit ordering cost and $p$ the exogenously determined unit selling price. A standard assumption is $p > c$, since otherwise, the newsvendor will choose to order nothing.

If the random demand has a probability density function $f$, then for a given order quantity $q$, the newsvendor's expected profit is then

$$p\mathrm{E}_f\left(\min\{\tilde{d},q\}\right)-cq.$$

Here the expectation $\mathrm{E}_f(\cdot)$ is taken with respect to the known distribution $f$. A risk-neutral newsvendor would be concerned with finding an ordering policy that maximizes the expected newsvendor profit. Suppose, rather than having a complete knowledge of the demand distribution, all the newsvendor knows are some of its parameters (e.g., known moments). Instead of maximizing the expected profit under some assumed distribution, an ambiguity-averse newsvendor will take a conservative approach by maximizing the worst-case profit. Mathematically, the problem is

$$\max_{q\geq 0}\left\{\inf_{f\in\mathbb{F}}\ p\mathrm{E}_f\left(\min\{\tilde{d},q\}\right)-cq\right\},$$

where $\mathbb{F}$ is the parametric family of distributions satisfying the known information.

## 2.1 Mean-Variance (MV) Model

Scarf [24] addressed the robust newsvendor model when the parametric family of distributions consist of those with mean $\mu$, variance $\sigma^2$ and nonnegative support. The worst-case newsvendor profit under this setting is

$$\begin{aligned}\Pi^{MV}(q) \ \triangleq\ & \inf_{f}\ p\mathrm{E}_f\left(\min\{\tilde{d},q\}\right)-cq\\ & \text{s.t.}\ \ \mathrm{E}_f(\tilde{d})=\mu,\quad \mathrm{E}_f(\tilde{d}^2)=\mu^2+\sigma^2\\ &\qquad\ \mathrm{E}_f(1)=1,\quad f(\tilde{d})\geq 0,\quad \forall\tilde{d}\geq 0.\end{aligned}$$

For any distribution belonging in this set, he found through a lengthy mathematical argument the optimal lower bound for the expected newsvendor profit. In particular,

$$p\mathrm{E}_f\left(\min\{\tilde{d},q\}\right)-cq\geq\begin{cases} pq\frac{\mu^2}{\mu^2+\sigma^2}-cq, & \text{for } q\in\left[0,\frac{\mu^2+\sigma^2}{2\mu}\right],\\ p\left(\frac{\mu+q}{2}-\frac{1}{2}\sqrt{(q-\mu)^2+\sigma^2}\right)-cq, & \text{for } q\in\left[\frac{\mu^2+\sigma^2}{2\mu},\infty\right).\end{cases}$$

This bound is tight, in the sense that there exists a feasible distribution with mean $\mu$ and variance $\sigma^2$ whose expected profit is exactly equal to the lower bound. The worst-case distribution is one that has a positive mass at exactly two points. If $q\leq(\mu^2+\sigma^2)/(2\mu)$, the worst-case distribution has mass $\sigma^2/(\mu^2+\sigma^2)$ at 0 and mass $\mu^2/(\mu^2+\sigma^2)$ at $(\mu^2+\sigma^2)/\mu$. Otherwise, the worst-case two-point distribution is

$$\tilde{d}=\begin{cases} q-\sqrt{(q-\mu)^2+\sigma^2}, & \text{w.p. } \frac{1}{2}\left(1+\frac{q-\mu}{\sqrt{(q-\mu)^2+\sigma^2}}\right),\\ q+\sqrt{(q-\mu)^2+\sigma^2}, & \text{w.p. } \frac{1}{2}\left(1-\frac{q-\mu}{\sqrt{(q-\mu)^2+\sigma^2}}\right).\end{cases}$$

Gallego and Moon [8] reach the same conclusion, but with a more concise proof that invokes the use of Cauchy-Schwartz inequality.

It is straightforward to find that an optimal ordering policy $q^*_{MV}$ that maximizes the worst-case profit is given by:

$$q^*_{MV} = \begin{cases} 0, & \text{if } \frac{c}{p} \geq \frac{\mu^2}{\mu^2+\sigma^2}, \\ \mu + \frac{\sigma}{2}\frac{(p-2c)}{\sqrt{c(p-c)}}, & \text{if } \frac{c}{p} \leq \frac{\mu^2}{\mu^2+\sigma^2}. \end{cases}$$

Under this ordering policy, the worst-case expected profit is

$$\Pi^{MV}(q^*_{MV}) = \begin{cases} 0, & \text{if } \frac{c}{p} \geq \frac{\mu^2}{\mu^2+\sigma^2}, \\ (p-c)\mu - \sigma\sqrt{c(p-c)}, & \text{if } \frac{c}{p} \leq \frac{\mu^2}{\mu^2+\sigma^2}. \end{cases}$$

It might be interesting to look at the expected profit under the best-case scenario. A naïve upper bound for the expected newsvendor profit can be found by invoking Jensen's inequality. Thus,

$$p\mathrm{E}_f\left(\min\{\tilde{d}, q\}\right) - cq \leq \min\{p\mu - cq, (p-c)q\}.$$

In fact, this upper bound is tight under mean and variance information (see De Schepper and Heijnen [25]).

## 2.2 Mean-Variance-Semivariance (MVS) Model

We introduce asymmetry into the robust newsvendor model through a characterization of the lower partial moments of the demand distribution. We will be focusing on the partial moments taken with respect to the mean demand to derive a closed-form expression for the worst-case newsvendor profit. The first-order lower partial moment does not capture asymmetry, because regardless of the distribution $f$, it is always true that

$$\mathrm{E}_f\left((\mu - \tilde{d})_+\right) = \mathrm{E}_f\left((\tilde{d} - \mu)_+\right).$$

This relationship results from the partial moment being defined with respect to the mean. Hence, we simply focus on the second-order lower partial moment, or semivariance. It is entirely possible to include asymmetry information in the first partial moment if it is defined with respect to some value other than the mean. However, the resulting closed-form expression for the model with the first two lower partial moments is untidy at best and and does not give us much additional insights. Instead, we direct interested readers to Appendix A which shows how this model can be solved as an SOCP.

We introduce the notion of *normalized semivariance*, which we define as

$$s \triangleq \frac{\mathrm{E}_f\left((\tilde{d} - \mu)^2_+\right) - \mathrm{E}_f\left((\mu - \tilde{d})^2_+\right)}{\sigma^2}. \tag{2.1}$$

This measure is only defined for random variables with a strictly positive and finite variance. We can immediately see that the normalized semivariance must take values in the range of $-1$ to 1. Clearly, $s$ describes how the volatility of the demand is divided between the upper and lower parts of the distribution. Figure 2.1 shows examples of some common probability distributions and their semivariances. Normal and uniform distributions always have a normalized semivariance of zero. An exponential distribution always has $s = 4e^{-1} - 1 \approx 0.4715$. The $s$ value of a beta distribution can be positive or negative

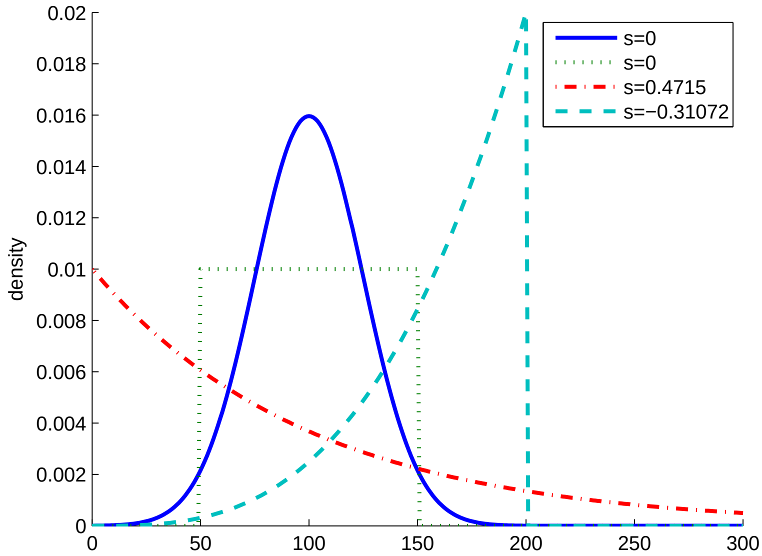


Figure 2.1: Some probability density functions and their normalized semivariances.

depending on its parameters. In fact, we can think of a value of $s = 0$ as a weaker form of distributional symmetry. All symmetric distributions (e.g. uniform, normal) must have a normalized semivariance of zero. However, the converse is not true. We can also think of a distribution with $s > 0$ as roughly positively skewed. Similarly, $s < 0$ implies that the distribution is roughly negatively skewed. We have mentioned that $s$ is in between -1 and 1. In fact, in the following proposition, we find a tighter bound for $s$ for nonnegative distributions.

**Proposition 2.1.** *If a nonnegative random variable with mean $\mu > 0$, standard deviation $\sigma > 0$ and normalized semivariance $s$, then*

$$\frac{\sigma^2 - \mu^2}{\sigma^2 + \mu^2} \leq s < 1. \tag{2.2}$$

*Moreover, given an triplet $(\mu, \sigma, s)$ that satisfies these conditions, we can also construct a nonnegative distribution with these moments. If the lower bound on $s$ is tight, then this distribution is unique.*

*Proof.* See Appendix B

Consider a newsvendor model where the exact demand distribution is unknown, but the mean $\mu$, variance $\sigma^2$ and normalized semivariance $s$ are known. For a given quantity $q$, the worst-case expected newsvendor profit is

$$\begin{aligned} \Pi^{MVS}(q) \quad \triangleq \quad \inf_f \quad & p\mathrm{E}_f\left(\min\{\tilde{d}, q\}\right) - cq \\ \text{s.t.} \quad & \mathrm{E}_f(\tilde{d}) = \mu, \quad \mathrm{E}_f\left((\tilde{d} - \mu)^2\right) = \sigma^2, \\ & \mathrm{E}_f\left((\tilde{d} - \mu)_+^2\right) - \mathrm{E}_f\left((\mu - \tilde{d})_+^2\right) = s\sigma^2, \\ & \mathrm{E}_f(1) = 1, \quad f(\tilde{d}) \geq 0, \quad \forall \tilde{d} \geq 0. \end{aligned}$$

We can consider $f$ to be an infinite dimensional vector indexed by $\tilde{d} \in \Re^+$, such that $f(\tilde{d}) : \Re^+ \mapsto \Re^+$. We assume that the conditions for Proposition 2.1 are satisfied so that the moment problem is well-defined. In fact, as we show in Theorem 2.1, we can find a closed-form expression for $\Pi^{MVS}(q)$. The proof of the theorem, which we relegate to Appendix B, is quite involved since it consists of constructing various forms of the dual feasible solutions. For each dual solution, we find a corresponding primal feasible distribution that achieves the same objective value. Note that in the theorem, the domain of $q$ is partitioned into five different regions. Each of the regions implies a particular form of the dual feasible solution which is optimal. A distribution that gives the worst-case expected profit is in fact one with at most three support points.

**Theorem 2.1.** *Consider a newsvendor problem specified by a unit cost $c$ and unit price $p$. Suppose the family of nonnegative demand distributions is specified by a known mean $\mu$, standard deviation $\sigma$, and normalized semivariance $s$. The worst-case expected profit, $\Pi^{MVS}(q)$, is given by*

$$\begin{cases} (p-c)q - \frac{p(1-s)\sigma^2}{2\mu^2}q, & \textit{for } \text{(i): } q \in \left[0, \frac{\mu}{2}\right], \\ (p-c)q - \frac{p(1-s)\sigma^2}{8(\mu-q)}, & \textit{for } \text{(ii): } q \in \left[\frac{\mu}{2}, \mu - \frac{\sigma}{2}\sqrt{\frac{1-s}{1+s}}\right], \\ p\left(\frac{(1-s)}{2}q + \frac{(1+s)}{2}\mu - \frac{\sigma}{2}\sqrt{1-s^2}\right) - cq, & \textit{for } \text{(iii): } q \in \left[\mu - \frac{\sigma}{2}\sqrt{\frac{1-s}{1+s}}, \mu + \frac{\sigma}{2}\sqrt{\frac{1+s}{1-s}}\right], \\ p\mu - cq - \frac{p(1+s)\sigma^2}{8(q-\mu)}, & \textit{for } \text{(iv): } q \in \left[\mu + \frac{\sigma}{2}\sqrt{\frac{1+s}{1-s}}, \mu + \frac{\mu(1+s)}{2(1-s)}\right], \\ \frac{p}{2}\left(\mu + bq - \sqrt{(bq-\mu)^2 - (1-b)^2\mu^2 + \frac{(1+s)\sigma^2 b}{2}}\right) - cq, & \textit{for } \text{(v): } q \in \left[\mu + \frac{\mu(1+s)}{2(1-s)}, \infty\right), \end{cases}$$

*where*

$$b = 1 - \frac{(1-s)\sigma^2}{2\mu^2}.$$

*Moreover, among the set of nonnegative distributions parameterized by $(\mu, \sigma, s)$, there exists a distribution with at most three support points that attains this bound.*

*Proof.* See Appendix B.

Note that $\Pi^{MVS}(q)$ is a concave function of $q$. This follows from observing that for any realization of the demand, the profit is a concave function of $q$. Linearity of expectations and the infimum operator maintains the concavity of the function $\Pi^{MVS}(\cdot)$. Using Theorem 2.1, we can find the order quantity $q^*_{MVS}$ that maximizes $\Pi^{MVS}(\cdot)$. Theorem 2.2 provides the optimal policy that depends on the magnitude of the ratio of unit cost to unit price $\frac{c}{p}$.

**Theorem 2.2.** *Consider a newsvendor problem specified by a unit cost $c$ and unit price $p$. Suppose the family of nonnegative demand distributions specified by a known mean $\mu$, standard deviation $\sigma$, and normalized semivariance $s$. An ordering policy $q^*_{MVS}$ that maximizes the worst-case expected profit*

$\Pi^{MVS}(\cdot)$ *is*

$$q^*_{MVS} = \begin{cases} 0, & \text{if } 1 - \frac{(1-s)\sigma^2}{2\mu^2} \le \frac{c}{p}, \\ \mu - \frac{\sigma}{2}\sqrt{\frac{(1-s)p}{2(p-c)}}, & \text{if } \frac{1}{2}(1-s) \le \frac{c}{p} < 1 - \frac{(1-s)\sigma^2}{2\mu^2}, \\ \mu + \frac{\sigma}{2}\sqrt{\frac{(1+s)p}{2c}}, & \text{if } \frac{1}{2}\frac{(1-s)^2}{(1+s)}\frac{\sigma^2}{\mu^2} \le \frac{c}{p} < \frac{1}{2}(1-s), \\ \frac{\mu}{b} + \frac{(pb-2c)}{2b}\sqrt{\frac{(1+s)\sigma^2 b - 2(1-b)^2\mu^2}{2c(pb-c)}}, & \text{if } \frac{c}{p} < \frac{1}{2}\frac{(1-s)^2}{(1+s)}\frac{\sigma^2}{\mu^2}, \end{cases}$$

*where*

$$b = 1 - \frac{(1-s)\sigma^2}{2\mu^2}.$$

*The worst-case expected newsvendor profit attained by the policy is*

$$\Pi^{MVS}(q^*_{MVS}) = \begin{cases} 0, & \text{if } 1 - \frac{(1-s)\sigma^2}{2\mu^2} \le \frac{c}{p}, \\ (p-c)\mu - \frac{\sigma}{2}\sqrt{2p(p-c)(1-s)}, & \text{if } \frac{1}{2}(1-s) \le \frac{c}{p} < 1 - \frac{(1-s)\sigma^2}{2\mu^2}, \\ (p-c)\mu - \frac{\sigma}{2}\sqrt{2pc(1+s)}, & \text{if } \frac{1}{2}\frac{(1-s)^2}{(1+s)}\frac{\sigma^2}{\mu^2} \le \frac{c}{p} < \frac{1}{2}(1-s), \\ \left(p - \frac{c}{b}\right)\left(\mu - \sqrt{\frac{c((1+s)\sigma^2 b - 2(1-b)^2\mu^2)}{2(pb-c)}}\right), & \text{if } \frac{c}{p} < \frac{1}{2}\frac{(1-s)^2}{(1+s)}\frac{\sigma^2}{\mu^2}. \end{cases}$$

*Proof.* See Appendix B.

**Corollary 2.1.** *Let $q^*_{MVS}$ be the optimal policy that maximizes the worst-case expected profit $\Pi^{MVS}(\cdot)$. Then we have the following:*

*(i) $q^*_{MVS}$ is decreasing in $\frac{c}{p}$,*
*(ii) $\Pi^{MVS}(q^*_{MVS})/p$ is decreasing in $\frac{c}{p}$,*
*(iii) $\Pi^{MVS}(q^*_{MVS})/\mu$ is decreasing in the coefficient of variation $\frac{\sigma}{\mu}$.*

Corollary 2.1 describes how the MVS policy changes as the parameter of the models change. Note that $\Pi^{MVS}(q)$ is the lower bound on the expected profit if the newsvendor orders $q$ units. Therefore, $\Pi^{MVS}(q^*_{MVS})$ is the guarantee on the expected profit if the MVS policy is used. Note that as the distribution is more and more spread out (higher coefficient of variation), then this guarantee decreases.

Another quantity of interest is the best-case bound for the expected newsvendor profit. We have seen that under the mean-variance framework, the best-case bound is the Jensen's bound. Under the MVS framework, the best-case profit is given by

$$\begin{aligned} \sup_f \quad & p\mathrm{E}_f\left(\min\{\tilde{d}, q\}\right) - cq \\ \text{s.t.} \quad & \mathrm{E}_f(\tilde{d}) = \mu, \quad \mathrm{E}_f\left((\tilde{d}-\mu)^2\right) = \sigma^2, \\ & \mathrm{E}_f\left((\tilde{d}-\mu)_+^2\right) - \mathrm{E}_f\left((\mu-\tilde{d})_+^2\right) = s\sigma^2, \\ & \mathrm{E}_f(1) = 1, \quad f(\tilde{d}) \ge 0, \quad \forall \tilde{d} \ge 0. \end{aligned}$$

Unlike the worst-case profit, the best-case bound is not necessarily a concave function of $q$. In Appendix A, we show a technique of casting this problem into an equivalent SOCP formulation if the value of $q$ is fixed. In fact, this technique can be applied to many variations to the newsvendor problem, such as if the demand has a bounded support, or if the newsvendor is risk-averse (Section 3). Different asymmetry models can also be handled, such as multiple partitions of the distribution. This is especially useful if, aside from partitioning at the mean, we also include partitions one standard deviation away from the mean in both directions. This could give us a more complete picture of asymmetry than just semivariance. In all these variations of the model, the best- and worst-case objective can be found by solving second order conic programs.

## 2.3 Comparing MV and MVS policies

Figure 2.2 illustrates how the ordering policies induced by the MV model and the MVS model are related. The values $q^*_{(ii)}, q^*_{(iv)}, q^*_{(v)}$ are defined in the proof of Theorem 2.2 (the subscript refers to the region where the optimum lies). Note that the optimum never occurs only in Region (iii), since the function is linear in $q$ in this region. Due to Proposition 2.1, we always have

$$\frac{1}{2}(1-s) \leq \frac{\mu^2}{\mu^2+\sigma^2} \leq 1 - \frac{(1-s)\sigma^2}{2\mu^2}.$$

Based on the diagram, we see that for any given $(\mu, \sigma, s)$ triplet, the difference between the two policies is clearly only dependent on the ratio $\frac{c}{p}$. Also observe that the MVS policy is less conservative than the MV policy in the sense that it recommends ordering nothing for a smaller range of $\frac{c}{p}$ values. This degree of conservatism also decreases as $s$ approaches 1. In fact, when $s$ is approximately 1, the optimal policy is to almost always order approximately $\mu$ units.

Figures 2.3 plots the optimal ordering quantity under the two models as a function of the cost to price ratio. We can see from Figure 2.3 that the MVS ordering policy follows the same general trend as the MV policy over the $\frac{c}{p}$ range. Scarf [24] observes that if $\frac{c}{p} < \frac{1}{2}$, the MV model suggests stocking more than the mean demand. If $\frac{c}{p} > \frac{1}{2}$, the policy is to stock less than the mean demand. On the other hand, under the MVS model, we have the following rule-of-thumb: stock less than the mean demand if $\frac{c}{p} > \frac{1}{2}(1-s)$, and more otherwise. We can also see in this figure that the MVS policy is less conservative since $q^*$ is zero only for a small region. The jumps we observe in the plot of $q^*$ indicate that there is in fact a range of quantities that maximize the worst-case profit.

Figure 2.4 plots the optimal policies as the degree of asymmetry increases. The MV policy remains static over all degrees of asymmetry. On the other hand, the MVS policy is adaptive to changes in the degree of asymmetry. Note that $q^*_{MVS}$ is not a monotone function of $s$, or of the coefficient of variation. However, $\Pi^{MVS}(q^*_{MVS})/\mu$ is decreasing in the coefficient of variation (Corollary 2.1).

Figure 2.5 plots the different bounds on the newsvendor profit as a function of $q$. The four different plots correspond to four different values of the normalized semivariance $s$. The normalized semivariance of the upper left plot is the smallest possible value in which the model is still well-defined (by Proposition 2.1). MV Worst and MV Best refer to the worst- and best-case expected profits under the given mean and variance information. MVS Worst and MVS Best are the worst- and best-case expected

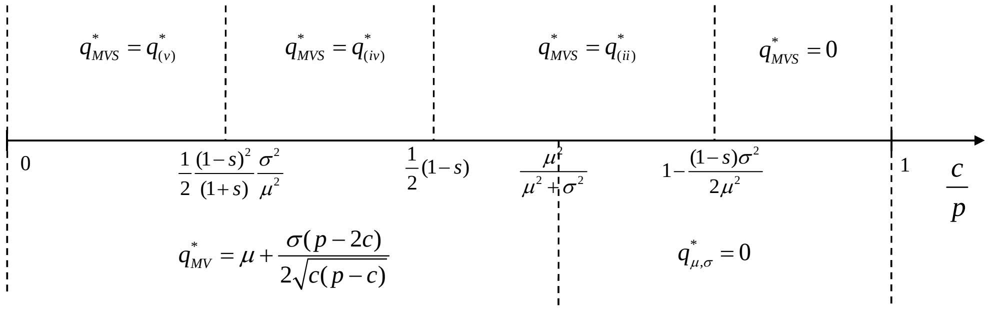


Figure 2.2: Optimal policies under different ranges of the unit cost to unit price ratio $c/p$.

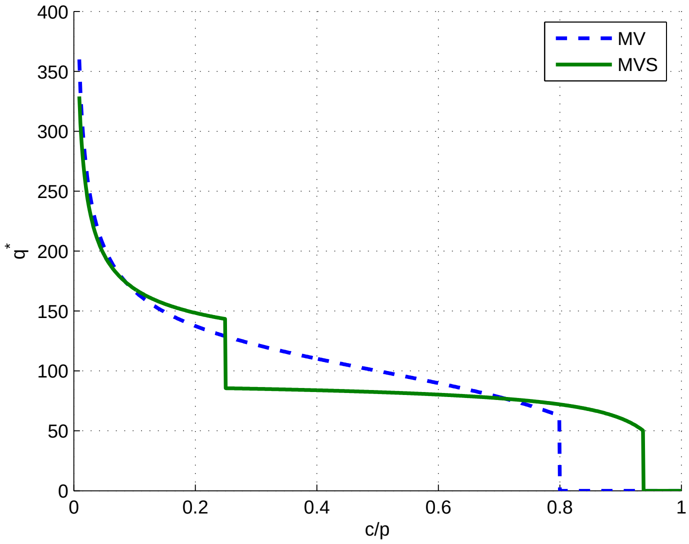


Figure 2.3: Sensitivity of the optimal order quantity to the unit cost to unit price ratio $c/p$ ($p = 1, \mu = 100, \sigma = 50, s = 0.5$).

profits with the additional semivariance information. MVS Best is plotted from the solution of multiple SOCP problems. The other three bounds are found in closed-form. Since the feasible distribution set of the MVS model is a subset of the one in the MV model, then clearly its optimal bounds must be contained between the MV bounds. We can see that, unlike the static MV bounds, the MVS bounds are highly dependent on the asymmetry. Moreover, the difference between the MVS bounds becomes small as $s$ approaches its upper and lower limits. One way to view this is that, at the limits of $s$, the feasible distribution set becomes more restrictive. For instance, at $s = -0.6$, the feasible set consists of a single distribution (see Proposition 2.1), which explains why the MVS upper and lower bounds are equal.

We observe from Figure 2.5 that the gap between the best- and worst-case expected profit can be large. In other words, there is a large uncertainty in the expected profit if the MV worst-case strategy is followed. However, the MVS framework manages to reduce this gap, thereby decreasing the uncertainty

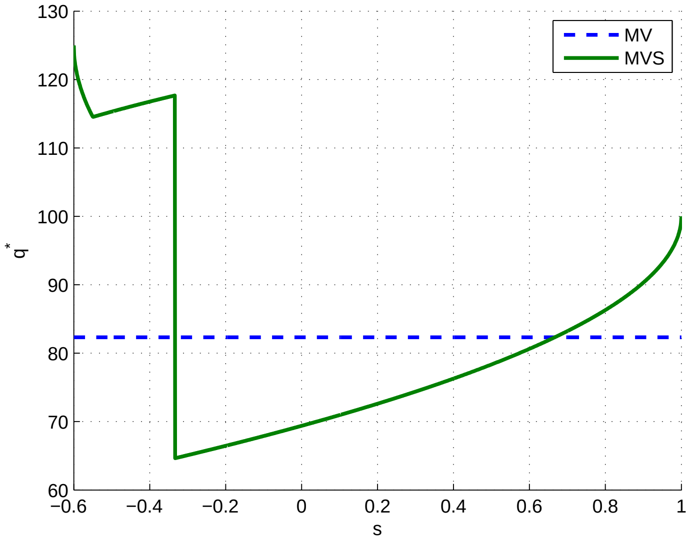


Figure 2.4: Sensitivity of the optimal order quantity and the worst-case bounds against the normalized semivariance $s$ $(p = 3, c = 2, \mu = 100, \sigma = 50)$.

in the expected profit. This figure also illustrates the conservativeness of the MV robust strategy. Consider the case when $s = 0$. There exists a distribution (satisfying the values for $\mu, \sigma, s$) under which, if the newsvendor follows the MV strategy (i.e., order about 80 units), he receives approximately \$30. On the other hand, by following the MVS strategy (i.e., order about 70 units), then the worst expected profit he can get is approximately \$39. That is, the loss of optimality of the MV policy can potentially be at least 23%.

## 2.4 Uncertainty in the Semivariance Estimate

It is conceivable, that the semivariance itself may not be known exactly. If there is uncertainty the semivariance estimate, then it might be better to assume that it belongs in a range of values. For instance, if the demand distribution is known to exhibit positive skewness, then we can instead assume that $s \geq 0$, rather than assuming an exact value for $s$. Specifying a range of semivariances increases the uncertainty set of distributions. Therefore, the resulting optimal bounds would be closer to the mean-variance bounds than if the semivariance is known. However, we will see that even if the range of values of $s$ is large, the MV policy can still be suboptimal.

Suppose $s$ is known to belong in the nonempty range $[s_l, s_h]$. Then for a given order quantity $q$, the

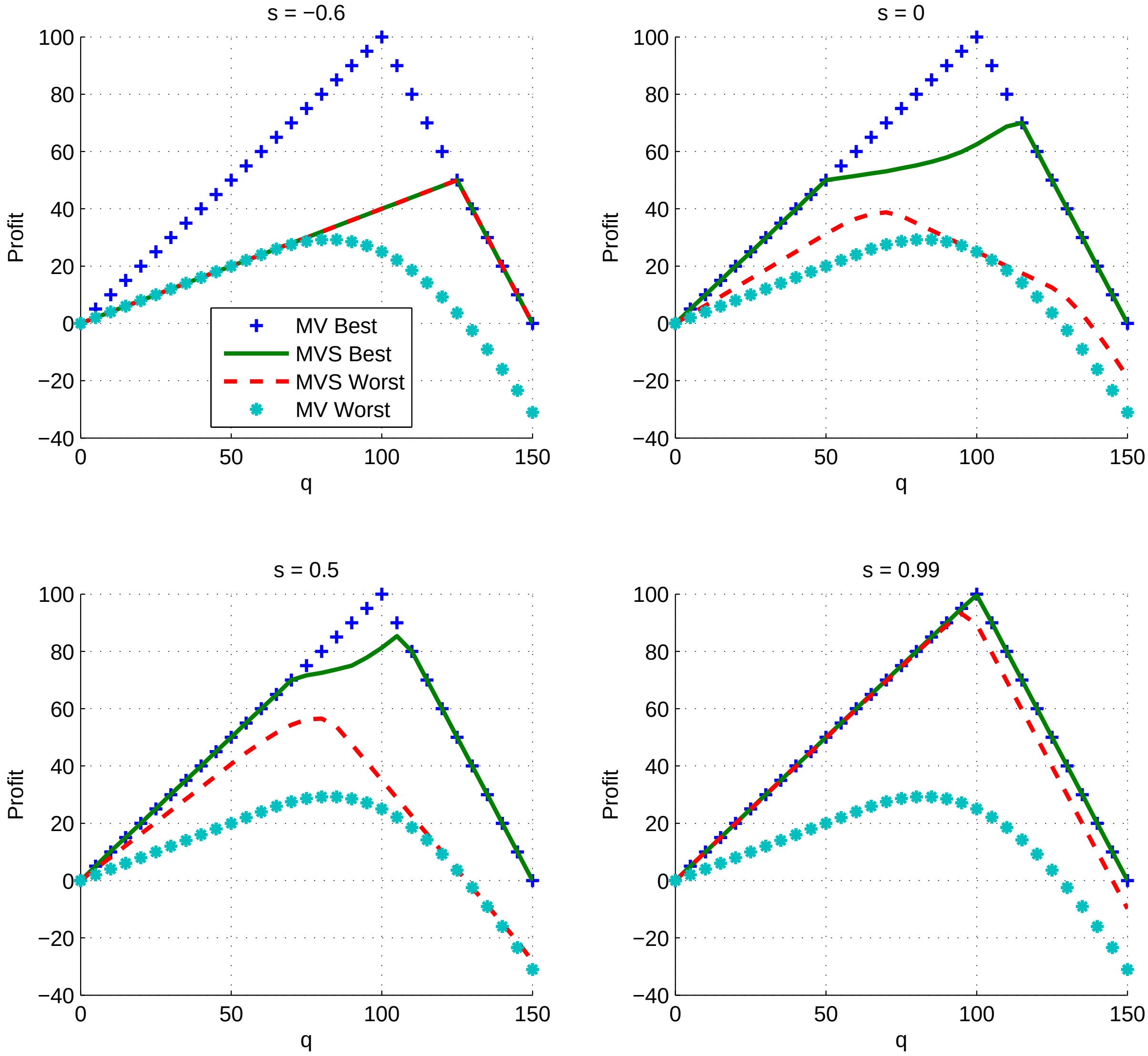


Figure 2.5: Bounds on the expected newsvendor profit ($p = 3, c = 2, \mu = 100, \sigma = 50$).

worst-case expected profit is

$$\begin{aligned}
\tilde{\Pi}^{MVS}(q) \quad \triangleq \quad & \inf_{f,s} \quad p\mathrm{E}_f\left(\min\{\tilde{d}, q\}\right) - cq \\
& \text{s.t.} \quad \mathrm{E}_f(\tilde{d}) = \mu, \quad \mathrm{E}_f\left((\tilde{d}-\mu)^2\right) = \sigma^2, \\
& \qquad \mathrm{E}_f\left((\tilde{d}-\mu)_+^2\right) - \mathrm{E}_f\left((\mu-\tilde{d})_+^2\right) = s\sigma^2, \\
& \qquad \mathrm{E}_f(1) = 1, \quad f(\tilde{d}) \geq 0, \quad \forall \tilde{d} \geq 0, \\
& \qquad s_l \leq s \leq s_h.
\end{aligned}$$

An ambiguity-averse newsvendor will choose a quantity $q$ that maximizes the worst-case expected profit. Using the technique in Appendix A, we can solve this problem as an SOCP. The worst-case expected profit can be expressed as

$$\tilde{\Pi}^{MVS}(q) = \inf_{s_l \le s \le s_h} \Pi^{MVS}(q; s)$$

where $\Pi^{MVS}(q; s)$ is the worst-case expected profit when the semivariance is $s$. For a fixed $q$, $\Pi^{MVS}(q; s)$ is a convex function of $s$. To see this, using Theorem 2.2, for $0 \le q \le \mu$, we express $\Pi^{MVS}(q; s)$ as

$$\Pi^{MVS}(q; s) = (p - c)q - \frac{p(1-s)\sigma^2}{2\mu^2} q.$$

If $\frac{\mu}{2} \le q \le \mu$, then

$$\Pi^{MVS}(q; s) = \begin{cases} (p-c)q - \frac{p(1-s)\sigma^2}{8(\mu-q)}, & \text{for } s \ge \frac{\sigma^2 - 4(\mu-q)^2}{\sigma^2 + 4(\mu-q)^2}, \\ p\left(\frac{(1-s)}{2} q + \frac{(1+s)}{2}\mu - \frac{\sigma}{2}\sqrt{1-s^2}\right) - cq, & \text{for } s < \frac{\sigma^2 - 4(\mu-q)^2}{\sigma^2 + 4(\mu-q)^2}. \end{cases}$$

Finally, if $q \ge \mu$, then

$$\Pi^{MVS}(q; s) = \begin{cases} p\left(\frac{(1-s)}{2} q + \frac{(1+s)}{2}\mu - \frac{\sigma}{2}\sqrt{1-s^2}\right) - cq, & \text{for } s \ge \frac{4(q-\mu)^2 - \sigma^2}{4(q-\mu)^2 + \sigma^2}, \\ p\mu - cq - \frac{p(1+s)\sigma^2}{8(q-\mu)}, & \text{for } \frac{2q-3\mu}{2q-\mu} \le s < \frac{4(q-\mu)^2 - \sigma^2}{4(q-\mu)^2 + \sigma^2}, \\ \frac{p}{2}\left(\mu + bq - \sqrt{(bq-\mu)^2 - (1-b)^2\mu^2 + \frac{1}{2}(1+s)\sigma^2 b}\right) - cq, & \text{for } s < \frac{2q-3\mu}{2q-\mu}. \end{cases}$$

The convexity of the functions in each of these ranges can be easily verified.

Figure 2.6 shows how the optimal bounds change as a function of the quantity $q$. Observe that the best-case bounds for the MV and MVS (with $s \ge 0$) are equal. While the MV policy recommends ordering about 85 units, the MVS policy recommends ordering around 70 units. Moreover, the MV policy is suboptimal under the MVS regime, with about 25% loss in the optimal worst-case profit.

# 3 Ambiguity-Averse, Risk-Averse Newsvendor

In Section 2, the newsvendor is assumed to be risk-neutral whose primary concern is in maximizing the expected profit. Intuitively, we expect knowledge of asymmetry to greatly affect the policies of a risk-averse newsvendor who heavily penalizes losses. A newsvendor who is risk-averse evaluates the quality of the ordering policy based on the risk he faces on the random payoff. The way risk is quantified depends on an individual's risk preferences.

A popular measure of risk used in finance is the Conditional Value-at-Risk (CVaR) (Rockafellar and Uryasev [22]). It is related to the Value-at-Risk (VaR) measure. By definition, with respect to a specific probability level $\alpha$, the $\alpha$-VaR is the lowest amount of loss $v$ such that, with probability $\alpha$, the loss will exceed that amount $v$. The $\alpha$-CVaR on the other hand is the conditional expectation of losses above that amount $v$ (see Artzner et al. [1]). The Conditional Value-at-Risk of a random loss $\tilde{x}$ can be written as

$$\text{CVaR}_\alpha(\tilde{x}) = \inf_{v \in \Re} \left\{ v - \frac{1}{1-\alpha} \mathrm{E}_f \left( \min\{0, v - \tilde{x}\} \right) \right\},$$

for some risk tolerance parameter $\alpha \in (0, 1)$. Values of $\alpha$ commonly considered are $0.90, 0.95$ and $0.99$. We can define the loss as the difference between some benchmark $M$ and the newsvendor profit, that is

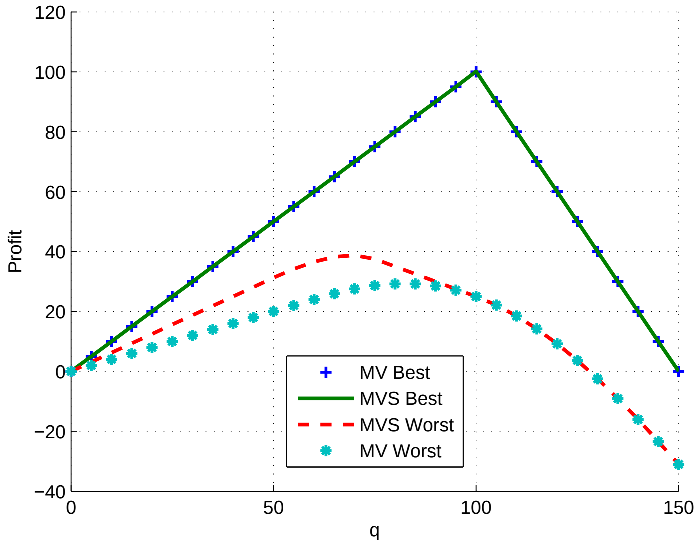


Figure 2.6: Bounds on the expected newsvendor profit with ambiguous normalized semivariance ($s \geq 0, p = 3, c = 2, \mu = 100, \sigma = 50$).

$\tilde{x} = M - p\min\{\tilde{d}, q\} + cq$. The benchmark $M$ defines what the newsvendor views as a loss. For instance, if $M = 0$, then any negative profits would be considered a loss. If $M = (p - c)\mu$, then loss would be the difference between the expected profit under the scenario where demand is known before ordering and the profit when ordering happens before observing the demand. In the newsvendor context, the CVaR risk measure has been analyzed in Gotoh [29] and Chen et.al. [30].

Suppose that the distribution is unknown, and only a few parameters such as mean or variance are known. We can reasonably expect a risk-averse newsvendor to exhibit ambiguity-aversion. That is, he can choose a policy that minimizes the ambiguous CVaR. The ambiguous CVaR is simply the CVaR with the expectation is taken with respect to the worst-case distribution. Mathematically, given an order quantity $q$, if the family of distributions is characterized by known mean and variance, then the ambiguous $\alpha$-CVaR is given by

$$\mathrm{AmCVaR}_\alpha^{MV}(q) \triangleq \inf_{v \in \Re} \left\{ v - \frac{1}{1-\alpha} \inf_{f \in \mathbb{F}} \mathrm{E}_f \left( \min \left\{ 0, v - M + p\tilde{d} - cq, v - M + (p-c)q \right\} \right) \right\},$$

where $\mathbb{F}$ is the set of all distributions with the given mean and variance. We similarly define $\mathrm{AmCVaR}_\alpha^{MVS}(\cdot)$ as the ambiguous $\alpha$-CVaR under the family of distributions with known mean, variance and semivariance. An ambiguity-averse newsvendor would choose an order quantity that minimizes the ambiguous CVaR of his loss. Mathematically, this is equivalent to

$$\min_{q \geq 0} \mathrm{AmCVaR}_\alpha(q).$$

Ambiguous risk measures have been studied in the context of portfolio management (Calafiore [5]; Natarajan et al. [16]). We extend this to the newsvendor model.

### 3.1 Mean-Variance (MV) Model

Suppose only the mean and variance of the demand is known. The following theorem gives a closed-form expression for the ambiguous CVaR of the newsvendor loss under a mean-variance framework.

**Theorem 3.1.** *Consider a newsvendor problem specified by a unit cost $c$ and a unit price $p$. Suppose the family of nonnegative demand distribution is characterized by known mean $\mu$ and standard deviation $\sigma$. For a given order quantity $q$, the ambiguous $\alpha$-CVaR of the newsvendor loss with respect to a target $M$ is*

$$AmCVaR_\alpha(q) = \begin{cases} M - (p-c)q + \frac{pq}{1-\alpha}\frac{\sigma^2}{\mu^2+\sigma^2}, & \text{for } 0 \le q < \frac{\mu^2+\sigma^2}{2\mu}, \\ M - (p-c)q + \frac{p}{2(1-\alpha)}\left(q - \mu + \sqrt{(q-\mu)^2+\sigma^2}\right), & \text{for } \frac{\mu^2+\sigma^2}{2\mu} \le q < \mu + \frac{\sigma(1-2\alpha)}{2\sqrt{\alpha(1-\alpha)}}, \\ M + cq - p\mu + p\sigma\sqrt{\frac{\alpha}{1-\alpha}}, & \text{for } q \ge \mu + \frac{\sigma(1-2\alpha)}{2\sqrt{\alpha(1-\alpha)}}, \end{cases}$$

*if $\alpha \le \frac{\mu^2}{\mu^2+\sigma^2}$. Otherwise, $AmCVaR_\alpha(q) = M + cq$ for all $q \ge 0$.*

*Proof.* See Appendix B.

The following theorem provides an optimal order quantity that minimizes the ambiguous CVaR risk under the mean-variance framework.

**Theorem 3.2.** *Consider a newsvendor problem specified by a unit cost $c$ and a unit price $p$. Suppose the family of nonnegative demand distributions is characterized by known mean $\mu$ and standard deviation $\sigma$. An ordering policy $q^*_{MV}$ that minimizes the worst-case $\alpha$-CVaR of the newsvendor loss with respect to a target $M$ is*

$$q^*_{MV} = \begin{cases} 0, & \text{if } \alpha \ge \frac{\mu^2}{\mu^2+\sigma^2} \text{ or } \beta < \frac{\sigma^2}{\mu^2+\sigma^2}, \\ \mu - \frac{\sigma(1-2\beta)}{2\sqrt{\beta(1-\beta)}}, & \text{if } \beta \ge \frac{\sigma^2}{\mu^2+\sigma^2}, \end{cases}$$

*where $\beta = (1-\alpha)\left(1 - \frac{c}{p}\right)$. The worst-case $\alpha$-CVaR attained by this policy is*

$$AmCVaR_\alpha^{MV}(q^*) = \begin{cases} M, & \text{if } \alpha \ge \frac{\mu^2}{\mu^2+\sigma^2} \text{ or } \beta < \frac{\sigma^2}{\mu^2+\sigma^2}, \\ M - (p-c)\left(\mu - \frac{\sigma}{2\sqrt{\beta(1-\beta)}}\right) + c\sigma\sqrt{\frac{\beta}{1-\beta}}, & \text{if } \beta \ge \frac{\sigma^2}{\mu^2+\sigma^2}. \end{cases}$$

*Proof.* See Appendix B.

First of all, observe that the risk-averse $MV$ policy looks exactly like the risk-neutral $MV$ policy, but with $\beta$ replacing $1 - \frac{c}{p}$. Intuitively, the newsvendor's behavior is not only governed by the cost and price, but also by his risk preference. Also observe that the newsvendor orders nothing if his risk tolerance is too low, or if the cost relative to the price is too large. Similar to the risk-neutral case, we also have a ordering rule-of-thumb: order more than the mean if $\beta > \frac{1}{2}$, otherwise, order less.

**Corollary 3.1.** *Let $q^*_{MV}$ be the optimal policy that minimizes the worst-case expected profit $AmCVaR^{MV}_{\alpha}(\cdot)$. Then we have the following:*

*(i) $q^*_{MV}/\mu$ is increasing (decreasing) in the coefficient of variation $\frac{\sigma}{\mu}$ if $\beta > \frac{1}{2}$ ($\beta < \frac{1}{2}$),*

*(ii) $q^*_{MV}$ is increasing in $\beta$ (i.e., it is decreasing in $\frac{c}{p}$ and decreasing in $\alpha$),*

*(iii) If $M = 0$ or $M = (p-c)\mu$, then $AmCVaR^{MV}_{\alpha}(q^*_{MV})/\mu$ is increasing in the coefficient of variation, $\frac{\sigma}{\mu}$,*

*(iv) $AmCVaR^{MV}_{\alpha}(q^*)$ is decreasing in $\alpha$ for $\alpha < \frac{1}{2}$, and increasing in $\alpha$ for $\alpha > \frac{1}{2}$.*

Note that $q^*_{MV}$ is monotone increasing or decreasing with respect to the coefficient of variation, depending on $\beta$. We can think of AmCVaR($q$) as the maximum guaranteed level of risk for a given quantity $q$. Thus, Corollary 3.1 states that the guaranteed upper bound on the CVaR if the MV policy is adopted is increasing as the distribution is more spread out.

### 3.2 Mean-Variance-Semivariance (MVS) Model

Now suppose that aside from the mean and variance, the distribution is also known to have a normalized semivariance $s$. Then the problem of minimizing the ambiguous $\alpha$-CVaR is equivalent to solving the following SOCP problem (see Appendix A):

$$
\begin{aligned}
\inf_{q,v,t,r,y_1,y_2,\tau_i^k} \quad & v - \tfrac{1}{1-\alpha}\left(t + \tfrac{1}{2}(1+s)\sigma^2 y_1 + \tfrac{1}{2}(1-s)\sigma^2 y_2\right) \\
\text{s.t.} \quad & -t - y_1 \geq \sqrt{(-t+y_1)^2 + \left(-r-\tau_1^1\right)^2}, \\
& v - M + p\mu - cq - t - y_1 \geq \sqrt{(v - M + p\mu - cq - t + y_1)^2 + \left(p - r - \tau_1^2\right)^2}, \\
& v - M + (p-c)q - t - y_1 \geq \sqrt{(v - M + (p-c)q - t + y_1)^2 + \left(-r - \tau_1^3\right)^2}, \\
& -t - y_2 + \tau_2^1 \geq \sqrt{\left(-t + y_2 - \tau_2^1\right)^2 + \left(r - \mu\tau_2^1\right)^2}, \\
& v - M + p\mu - cq - t - y_2 + \tau_2^2 \geq \sqrt{\left(v - M + p\mu - cq - t + y_2 - \tau_2^2\right)^2 + \left(-p + r - \mu\tau_2^2\right)^2}, \\
& v - M + (p-c)q - t - y_2 + \tau_2^3 \geq \sqrt{\left(v - M + (p-c)q - t + y_2 - \tau_2^3\right)^2 + \left(r - \mu\tau_2^3\right)^2}, \\
& q, \tau_i^k \geq 0, \qquad i = 1,2,\ k = 1,2,3.
\end{aligned}
$$

First let us observe how the MVS policies behave as a function of the model parameters. In the following plots, we plot the policies for three ambiguity-averse newsvendors with varying risk preferences, which are found by solving SOCP formulations. First, we argue that a lower risk parameter $\alpha$ corresponds to a higher optimal order quantity $q^*$. This is because the newsvendor with a higher risk parameter would prefer to be safe and order less, to reduce the magnitude of losses. Figure 3.1 illustrates how the optimal order quantity and AmCVaR change as the ratio $\frac{c}{p}$ increases. Observe that as the newsvendor becomes more risk-averse, the range of $\frac{c}{p}$ for which he will become conservative and order nothing increases. For instance, a newsvendor with $\alpha = 0.95$ will only order a positive amount if the cost is less than about 55% of the price. Now let us focus our attention to Figure 3.2, which shows how the MVS policies changes as the normalized semivariance increases. Observe that the risk-averse newsvendors will choose to order nothing if the distribution is negatively skewed. This is true, even if the risk parameter is fairly small ($\alpha = 0.75$). As soon as the distribution exhibits a large enough

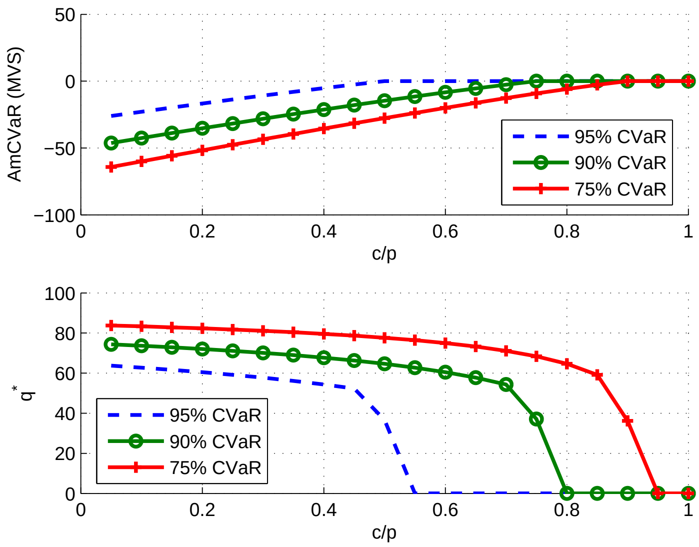


Figure 3.1: Sensitivity of the mean-variance-semivariance order quantity and the ambiguous CVaR against the unit cost to unit price ratio $c/p$ ($p = 1, \mu = 100, \sigma = 50, s = 0.8, M = 0$).

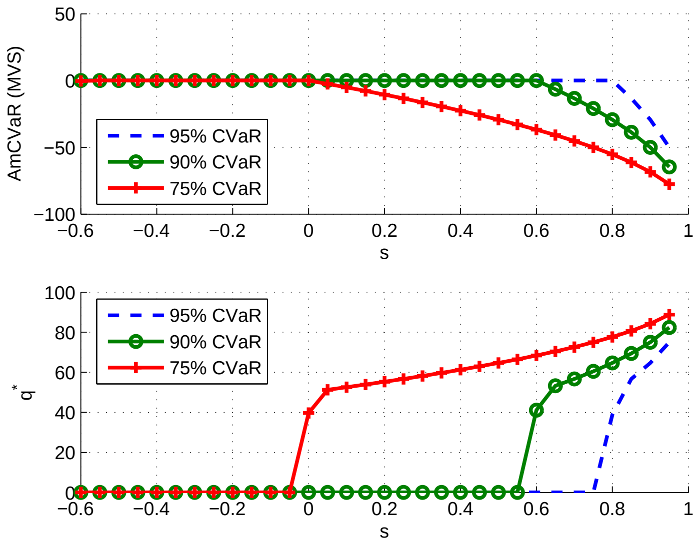


Figure 3.2: Sensitivity of the mean-variance-semivariance order quantity and the ambiguous CVaR against the normalized semivariance $s$ ($p = 2, c = 1, \mu = 100, \sigma = 50, M = 0$).

degree of positive skewness, then the risk-averse newsvendor will decide to order a positive amount. This optimal order quantity is monotonically increasing in the degree of asymmetry.

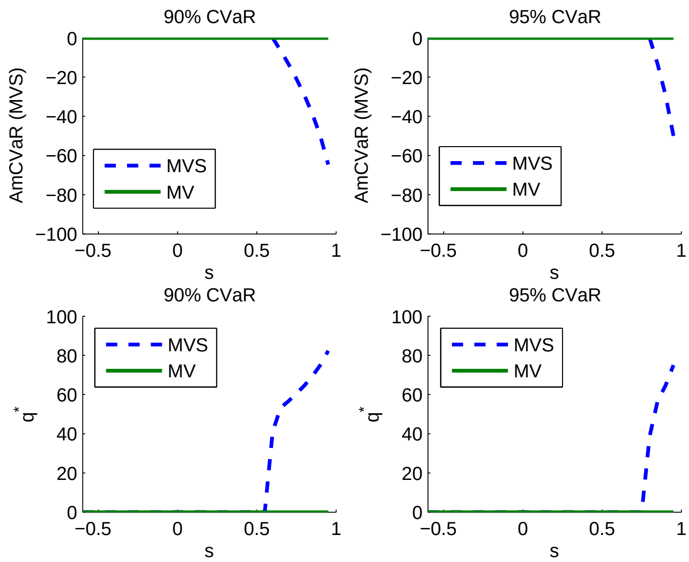


Figure 3.3: Comparison between mean-variance and mean-variance-semivariance risk-averse policies as the normalized semivariance varies ($p = 2, c = 1, \mu = 100, \sigma = 50, M = 0$).

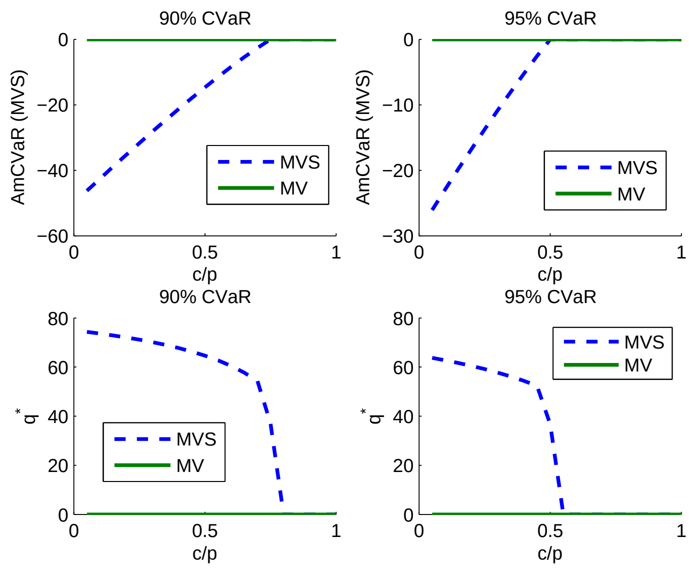


Figure 3.4: Comparison between mean-variance and mean-variance-semivariance risk-averse policies as $c/p$ varies ($s = 0.8, c = 1, \mu = 100, \sigma = 50, M = 0$).

## 3.3 Comparing MV and MVS policies

We have seen that the MVS risk-averse policies recommend positive order quantities if the distribution is known to have a normalized semivariance that is large enough. This is in contrast with the MV risk-averse policies that remain conservative even if the distribution is very positively skewed and the cost is small enough. Figure 3.3 illustrates how the MV and MVS risk-averse policies change as the normalized semivariance is increased. Since the MV policies are independent of $s$, they remain static, as we can observe in the plot. If the normalized is semivariance is negative (as it is for negatively skewed distributions), then both policies are conservative. However, as the semivariance increases, the MVS policies suggest ordering higher amounts. Moreover, the gap between the AmCVaR (evaluated under the MVS uncertainty set) of both policies becomes larger. For example, for $s = 0.8$, the MV policy achieves an AmCVaR that is about \$40 above the optimal MVS risk for $\alpha = 0.9$, and \$20 above the optimal MVS risk for $\alpha = 0.95$.

Figure 3.4 shows how the MV and MVS risk-averse policies vary as the ratio $\frac{c}{p}$ is increased. Observe that for the common risk parameters ($\alpha = 0.9, 0.95$), the MV risk-averse policies are conservative for all ranges of $\frac{c}{p}$. That is, even if the cost is very small, the MV policy will recommend ordering nothing. On the other hand, for the MVS policies, the optimal order quantity becomes positive for small enough cost to price ratios. Moreover, as this ratio decreases, the AmCVaR gap (under the MVS uncertainty set) increases. If the cost is 25% of the price, then the MV policy achieves an AmCVaR risk that is about \$30 above the optimal MVS risk for $\alpha = 0.9$ and about \$15 above the optimal MVS risk for $\alpha = 0.95$.

These two figures imply that knowledge of asymmetry has a huge effect on risk-averse policies. Without knowing the degree of asymmetry, then the newsvendor will choose a conservative policy, that can be suboptimal in some cases.

### 3.4 Uncertainty in the Semivariance Estimate

As with the risk-neutral case in Section 2.4, the same technique can be applied to a risk-averse newsvendor who aims to minimize the ambiguous CVaR and is unsure of the exact value of the semivariance. Suppose he believes that $s$ belongs in the range $[s_l, s_h]$. We can solve for the optimal order quantity through the SOCP formulation outlined in Appendix A.

Figure 3.5 plots the optimal order quantity and AmCVaR (evaluated under the MVS uncertainty set) of the MV and MVS policies as a function of the cost-to-price ratio. Observe that the MV policy remains conservative for all values of this ratio. However, the MVS policy recommends a positive order quantity for small enough costs. If the cost is about 5% of the price, then the MV policy has an AmCVaR that is \$6 above the optimal MVS AmCVaR.

## 4 Conclusion

We have introduced a decision making framework under which an ambiguity-averse newsvendor can make optimal decisions with knowledge of the asymmetry. In particular, this asymmetry is represented by a single parameter: the normalized semivariance. Moreover, this framework can be also applied to various extensions of the model, such as bounded support and multiple demand partitions. Since this measure of asymmetry is a second-order measure, the resulting models can be formulated as second-order cone programs. Moreover, in some special cases, the optimal order quantity can be found in closed-form. Knowledge of asymmetry can greatly reduce the uncertainty in the expected profit, which

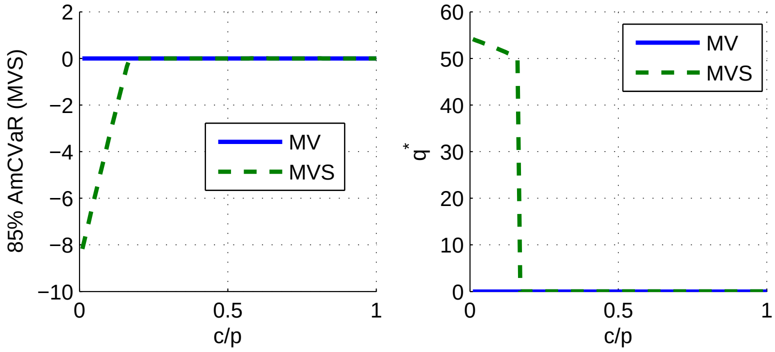


Figure 3.5: Comparison of mean-variance and mean-variance-semivariance policies under ambiguous normalized semivariance with respect to the unit cost to unit price ratio ($p = 1, \mu = 100, \sigma = 50, s \geq 0, M = (p - c)\mu$).

we saw by the reduction in the gap between the best- and worst-case bounds on the expected profit. We have also shown through several examples that the traditional mean-variance models can be suboptimal under the mean-variance-semivariance framework. Moreover, knowing the asymmetry can decrease the inherent conservatism in these mean-variance models. This has most impact in the risk-averse newsvendor models, where the mean-variance policies remain conservative even if the cost is a very small fraction of the price. We have also shown that we can relax the assumption that the semivariance is known exactly into an assumption of knowing a range of values for which the semivariance takes. Even in this relaxed model, we have shown that if the distribution is known to have positive normalized semivariance, the improvement of the model over the mean-variance models is still significant. This suggests that our model is useful especially for applications where the underlying distribution is known to be asymmetric.

## Appendix A General Partitioned Model

Consider a single random variable $\tilde{x}$ with a support set $A \subseteq \Re$. We can partition the support set in the following manner. Let $\{A_i\}_{i=1}^p$ be non-overlapping intervals whose union is $A$ (see Figure A.1). Therefore,

$$\tilde{x} = \sum_{i=1}^{p} h_i(\tilde{x}),$$

where

$$h_i(x) = \begin{cases} x, & \text{if } x \in A_i, \\ 0, & \text{otherwise.} \end{cases}$$

We can view each $h_i(\tilde{x})$ as a new partitioned random variable. If we have limited information on the distribution of $\tilde{x}$, say, the mean and variance, then this gives no light as to any asymmetry information. However, introducing the $p$ partitioned random variables gives us an idea of how distribution is divided among the support subintervals. Consider the expectation over a piecewise linear function of $\tilde{x}$

$$\mathrm{E}_f \left( \min_{k=1,\ldots,K} \{a_k \tilde{x} + b_k\} \right).$$

If $\tilde{x}$ represents a random payoff, we can interpret the piecewise linear function as some form of utility function. Using the partitioning, we can rewrite the expectation as

$$\mathrm{E}_f \left( \min_{k=1,\ldots,K} \left\{ a_k \sum_{i=1}^{p} h_i(\tilde{x}) + b_k \right\} \right).$$

If we have mean and variance information on each of the partitioned random variables, then a robust approach is to optimize the objective subject to the known information. Consider the problem of minimizing the objective, that is

$$\begin{array}{rcll} GP_m & = & \inf\limits_f & \mathrm{E}_f \left( \min\limits_{k=1,\ldots,K} \left\{ a_k \sum\limits_{i=1}^{p} h_i(\tilde{x}) + b_k \right\} \right) \\ & & \text{s.t.} & \mathrm{E}_f(h_i(\tilde{x})) = \mu_i, \quad \forall i = 1, \ldots, p \\ & & & \mathrm{E}_f(h_i(\tilde{x})^2) = \mu_i^2 + \sigma_i^2, \quad \forall i = 1, \ldots, p, \\ & & & \mathrm{E}_f(1) = 1, \\ & & & f(\tilde{x}) \geq 0, \quad \forall \tilde{x} \in A. \end{array}$$

We can think of the MVS model as a special case of the partitioning with $p = 2$. But for a model with more than two partitions, a valid question remains: how do you define reasonable partitions? We believe that partitions can naturally result from the model. For instance, they may come from quantiles of the empirical distribution. This choice of partitioning not only has theoretical value, but also leads to "fair" estimates of the moments given finite realizations of $\tilde{x}$. In other words, each partitioned random variable will have moment estimates using roughly the same number of samples. This enforces estimation errors to be spread out among the partitions.

By strong duality for distribution functions (Isii [11]; Bertsimas and Popescu [20]), if the moments of the problem lie strictly in the interior of the feasible moment cone, then the problem is equivalent to the dual formulation

$$\begin{array}{ll} \sup\limits_{t,r_i,y_i} & t + \sum\limits_{i=1}^{p} r_i\mu_i + \sum\limits_{i=1}^{p} y_i(\mu_i^2+\sigma_i^2) \\ \text{s.t.} & t + r_i x + y_i x^2 \le a_k x + b_k, \quad \forall x \in A_i,\ \forall i = 1,\ldots,p,\ \forall k = 1,\ldots,K. \end{array} \tag{A.1}$$

There are a total of $Kp$ constraints in the dual problem. In the previous section, we found a closed-form expression to the dual problem in the simple case of $K = 2$ and $p = 2$. Closed-form expressions may prove to be complicated for models with multiple partitions or piecewise functions with multiple linear pieces. However, we show that we can in fact write the dual problem into a second-order cone programming problem by invoking the widely used S-lemma (Pólik and Terlaky [19]), which we state here for completeness.

**Proposition A.1** (S-lemma)**.** *Consider two quadratic functions of* $\boldsymbol{z} \in \Re^N$, $q_i(\boldsymbol{z}) = \boldsymbol{z}'\boldsymbol{B}_i\boldsymbol{z} + 2\boldsymbol{b}_i'\boldsymbol{z} + c_i$, $i = 0, 1$, *with* $q_1(\bar{\boldsymbol{z}}) > 0$ *for some* $\bar{\boldsymbol{z}}$. *Then*

$$q_0(\boldsymbol{z}) \ge 0 \quad \forall \boldsymbol{z} \in \{\boldsymbol{z} : q_1(\boldsymbol{z}) \ge 0\}$$

*if and only if there exists* $\tau \ge 0$ *such that*

$$\begin{pmatrix} c_0 & \boldsymbol{b}_0' \\ \boldsymbol{b}_0 & \boldsymbol{B}_0 \end{pmatrix} - \tau \begin{pmatrix} c_1 & \boldsymbol{b}_1' \\ \boldsymbol{b}_1 & \boldsymbol{B}_1 \end{pmatrix} \in \mathcal{S}^+,$$

*where* $\mathcal{S}^+$ *denotes the set of all positive-semidefinite matrices of corresponding size.*

Each interval $A_i \subset \Re$ can either be bounded, unbounded above, or unbounded below. We show how to convert the constraints in (A.1) to a second-order constraints if the interval is bounded. The other two cases can be handled in a similar manner. Suppose $A_i = [\underline{x}_i, \overline{x}_i]$, where $\underline{x}_i < \overline{x}_i$ and $\underline{x}_i, \overline{x}_i$ are finite. Then the set of all $x$ that belong in $A_i$ can be written as

$$\{x \in \Re : (x - \underline{x}_i)(x - \overline{x}_i) \le 0\}.$$

By directly applying S-lemma, we can write the $K$ constraints

$$t + r_i x + y_i x^2 \le a_k x + b_k, \quad \forall x \in A_i,\ \forall k = 1,\ldots,K$$

as the positive semidefinite constraints

$$\begin{pmatrix} b_k - t & \dfrac{a_k - r_i}{2} \\ \dfrac{a_k - r_i}{2} & -y_i \end{pmatrix} - \tau_i^k \begin{pmatrix} -\underline{x}_i\overline{x}_i & \dfrac{\underline{x}_i + \overline{x}_i}{2} \\ \dfrac{\underline{x}_i + \overline{x}_i}{2} & -1 \end{pmatrix} \succeq 0, \quad \tau_i^k \ge 0, \quad \forall k = 1,\ldots,K.$$

Notice that the constraints only require $2 \times 2$ matrices to be positive semidefinite. In fact, we can rewrite each of the $K$ positive semidefiniteness constraints into the following simpler form:

$$\begin{aligned} & b_k - t + \underline{x}_i\overline{x}_i\tau_i^k \ge 0, \\ & \left(b_k - t + \underline{x}_i\overline{x}_i\tau_i^k\right)\left(-y_i + \tau_i^k\right) \ge \left(\frac{a_k - r_i - (\underline{x}_i + \overline{x}_i)\tau_i^k}{2}\right)^2, \end{aligned}$$

for $k = 1, \ldots, K$. The last inequality simply means that the determinant of the matrix must be nonnegative. Note that this constraint includes a hyperbolic term. We can further cast the constraints into the following second-order constraints (see Lobo et al. [14]):

$$b_k - t + \underline{x}_i \overline{x}_i \tau_i^k - y_i + \tau_i^k \geq \sqrt{\left(b_k - t + \underline{x}_i \overline{x}_i \tau_i^k + y_i - \tau_i^k\right)^2 + \left(a_k - r_i - (\underline{x}_i + \overline{x}_i)\tau_i^k\right)^2}, \quad k = 1, \ldots, K.$$

By a similar idea, if $A_i$ is unbounded above with $A_i = [\underline{x}_i, \infty)$, then the corresponding $K$ constraints in the dual form (A.1) can be written as:

$$b_k - t + \underline{x}_i \tau_i^k - y_i \geq \sqrt{\left(b_k - t + \underline{x}_i \tau_i^k + y_i\right)^2 + \left(a_k - r_i - \tau_i^k\right)^2}, \qquad k = 1, \ldots, K.$$

On the other hand, if $A_i = (\infty, \overline{x}_i]$ , then the corresponding constraints are:

$$b_k - t - \underline{x}_i \tau_i^k - y_i \geq \sqrt{\left(b_k - t - \underline{x}_i \tau_i^k + y_i\right)^2 + \left(a_k - r_i + \tau_i^k\right)^2}, \qquad k = 1, \ldots, K.$$

Finally, we can write the dual problem as an SOCP. Let $I_0$ be the set of all indices $i$ for which $A_i$ is bounded. Also denote $I_\infty$ (and $I_{-\infty}$) for indices in which $A_i$ is unbounded above (and below). Thus, minimization problem $GP_m$ can be solved through the following SOCP:

$$\begin{aligned}
\sup_{t, r_i, y_i, \tau_i^k} \quad & t + \sum_{i=1}^p r_i \mu_i + \sum_{i=1}^p y_i(\mu_i^2 + \sigma_i^2) \\
\text{s.t.} \quad & b_k - t + \underline{x}_i \overline{x}_i \tau_i^k - y_i + \tau_i^k \geq \sqrt{\left(b_k - t + \underline{x}_i \overline{x}_i \tau_i^k + y_i - \tau_i^k\right)^2 + \left(a_k - r_i - (\underline{x}_i + \overline{x}_i)\tau_i^k\right)^2}, \\
& \hspace{20em} \forall i \in I_0,\ \forall k = 1, \ldots, K, \\
& b_k - t + \underline{x}_i \tau_i^k - y_i \geq \sqrt{\left(b_k - t + \underline{x}_i \tau_i^k + y_i\right)^2 + \left(a_k - r_i - \tau_i^k\right)^2}, \\
& \hspace{20em} \forall i \in I_\infty,\ \forall k = 1, \ldots, K, \\
& b_k - t - \underline{x}_i \tau_i^k - y_i \geq \sqrt{\left(b_k - t - \underline{x}_i \tau_i^k + y_i\right)^2 + \left(a_k - r_i + \tau_i^k\right)^2}, \\
& \hspace{20em} \forall i \in I_{-\infty},\ \forall k = 1, \ldots, K, \\
& \tau_i^k \geq 0, \quad \forall i = 1, \ldots, K,\ \forall k = 1, \ldots, K.
\end{aligned}$$

The SOCP formulation can be solved efficiently using interior point methods (see Nesterov and Nemirovski [17]). Its optimal solution can be found using modern solvers such as CPLEX, MOSEK, SDPT3 or SeDuMi.

In fact, with some minor modifications, we can also find an equivalent SOCP formulation for the problem of maximizing the objective:

$$\begin{aligned}
GP_M \quad = \quad \sup_f \quad & \mathrm{E}_f\left(\min_{k=1,\ldots,K}\left\{a_k \sum_{i=1}^p h_i(\tilde{x}) + b_k\right\}\right) \\
\text{s.t.} \quad & \mathrm{E}_f(h_i(\tilde{x})) = \mu_i, \qquad \forall i = 1, \ldots, p \\
& \mathrm{E}_f(h_i(\tilde{x})^2) = \mu_i^2 + \sigma_i^2, \quad \forall i = 1, \ldots, p, \\
& \mathrm{E}_f(1) = 1, \\
& f(\tilde{x}) \geq 0, \quad \forall \tilde{x} \in A.
\end{aligned}$$

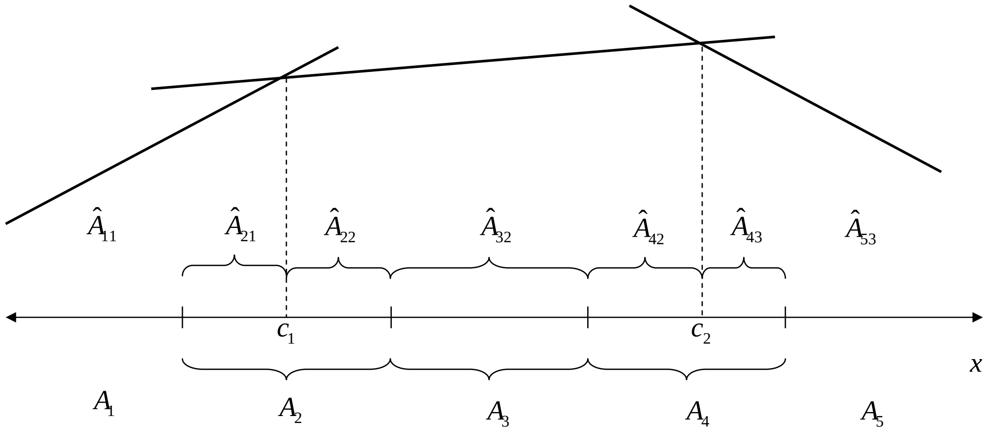


Figure A.1: Example of a partition for $A = \Re$ with $K = 3$ and $p = 5$.

We can assume without loss of generality that $a_1 > a_2 > \cdots > a_K$. Let $c_k$ be the point of intersection between the $k$th and $(k+1)$th linear piece, i.e. $a_k c_k + b_k = a_{k+1} c_k + b_{k+1}$. Let us define new subintervals $\hat{A}_{ik} = A_i \cap [c_{k-1}, c_k]$, for $i = 1, \ldots, p$ and $k = 1, \ldots, K$. We define $c_0 = -\infty$ and $c_K = \infty$. Figure A.1 illustrates how this new partition is constructed.

If the moments of the problem lie strictly in the interior of the feasible moment cone, then the problem is equivalent to

$$\begin{aligned} \inf_{t, r_i, y_i} \quad & t + \sum_{i=1}^{p} r_i \mu_i + \sum_{i=1}^{p} y_i (\mu_i^2 + \sigma_i^2) \\ \text{s.t.} \quad & t + r_i x + y_i x^2 \geq \min_{k=1,\ldots,K} \{a_k x + b_k\}, \quad \forall x \in A_i, \ \forall i = 1, \ldots, p, \end{aligned}$$

or, equivalently,

$$\begin{aligned} \inf_{t, r_i, y_i} \quad & t + \sum_{i=1}^{p} r_i \mu_i + \sum_{i=1}^{p} y_i (\mu_i^2 + \sigma_i^2) \\ \text{s.t.} \quad & t + r_i x + y_i x^2 \geq a_k x + b_k, \quad \forall x \in \hat{A}_{ik}, \ \forall i = 1, \ldots, p, \ \forall k = 1, \ldots, K. \end{aligned} \tag{A.2}$$

Note that if the interval $\hat{A}_{ik}$ is empty, then the corresponding constraint is ignored. We can therefore apply the same method of casting the constraints (A.2) into SOCP constraints as before. Suppose $\underline{x}_{ik}$ and $\overline{x}_{ik}$ are the lower and upper bounds of interval $\hat{A}_{ik}$ (if they exist). Let $\hat{I}_0$ consist of all $(i,k)$ for which $\hat{A}_{ik}$ is bounded, and $\hat{I}_\infty$ ($\hat{I}_{-\infty}$) consist of all $(i,k)$ for which $\hat{A}_{ik}$ is unbounded above (below). The maximization problem $GP_M$ can be solved through the following SOCP:

$$\begin{aligned} \inf_{t, r_i, y_i, \tau_i^k} \quad & t + \sum_{i=1}^{p} r_i \mu_i + \sum_{i=1}^{p} y_i (\mu_i^2 + \sigma_i^2) \\ \text{s.t.} \quad & -b_k + t + \underline{x}_{ik} \overline{x}_{ik} \tau_i^k + y_i + \tau_i^k \geq \sqrt{\left(-b_k + t + \underline{x}_{ik} \overline{x}_{ik} \tau_i^k - y_i - \tau_i^k\right)^2 + \left(-a_k + r_i - (\underline{x}_{ik} + \overline{x}_{ik}) \tau_i^k\right)^2}, \\ & \hspace{25em} \forall (i,k) \in \hat{I}_0, \\ & -b_k + t + \underline{x}_{ik} \tau_i^k + y_i \geq \sqrt{\left(-b_k + t + \underline{x}_{ik} \tau_i^k - y_i\right)^2 + \left(-a_k + r_i - \tau_i^k\right)^2}, \quad \forall (i,k) \in \hat{I}_\infty, \\ & -b_k + t - \underline{x}_{ik} \tau_i^k + y_i \geq \sqrt{\left(-b_k + t - \underline{x}_{ik} \tau_i^k - y_i\right)^2 + \left(-a_k + r_i + \tau_i^k\right)^2}, \quad \forall (i,k) \in \hat{I}_{-\infty}, \\ & \tau_i^k \geq 0, \quad \forall (i,k) \in \hat{I}_0 \cup \hat{I}_\infty \cup \hat{I}_{-\infty}. \end{aligned}$$

# Appendix B

## B.1 Proof of Proposition 2.1

For notational convenience, we define the nonnegative random variables $\tilde{d}_1 = (\tilde{d}-\mu)_+$ and $\tilde{d}_2 = (\mu-\tilde{d})_+$. From the definition of $s$, we have $\mathrm{E}(\tilde{d}_1^2) = \frac{1}{2}(1+s)\sigma^2$ and $\mathrm{E}(\tilde{d}_2^2) = \frac{1}{2}(1-s)\sigma^2$. Since $\tilde{d}$ is nonnegative, $\tilde{d}_2$ must not exceed $\mu$. This implies that $\mathrm{E}(\tilde{d}_2(\mu-\tilde{d}_2)) \geq 0$, or equivalently, $\mathrm{E}(\tilde{d}_2) \geq \mathrm{E}(\tilde{d}_2^2)/\mu = (1-s)\sigma^2/(2\mu)$. Also since $\mathrm{E}(\tilde{d}_2^2|\tilde{d}_2>0) \geq (\mathrm{E}(\tilde{d}_2|\tilde{d}_2>0))^2$, then we have

$$\Pr(\tilde{d}_2>0) \geq \frac{\Big(\mathrm{E}(\tilde{d}_2|\tilde{d}_2>0)\Pr(\tilde{d}_2>0) + 0\cdot\Pr(\tilde{d}_2=0)\Big)^2}{\mathrm{E}(\tilde{d}_2^2|\tilde{d}_2>0)\Pr(\tilde{d}_2>0) + 0^2\cdot\Pr(\tilde{d}_2=0)} = \frac{(\mathrm{E}(\tilde{d}_2))^2}{\mathrm{E}(\tilde{d}_2^2)} \geq \frac{(1-s)\sigma^2}{2\mu^2}. \tag{B.1}$$

By a similar argument, and since $\mathrm{E}(\tilde{d}_1) = \mathrm{E}(\tilde{d}_2)$, we also find that

$$\Pr(\tilde{d}_1>0) \geq \frac{\Big(\mathrm{E}(\tilde{d}_1|\tilde{d}_1>0)\Pr(\tilde{d}_1>0) + 0\cdot\Pr(\tilde{d}_1=0)\Big)^2}{\mathrm{E}(\tilde{d}_1^2|\tilde{d}_1>0)\Pr(\tilde{d}_1>0) + 0^2\cdot\Pr(\tilde{d}_1=0)} = \frac{(\mathrm{E}(\tilde{d}_1))^2}{\mathrm{E}(\tilde{d}_1^2)} \geq \frac{(1-s)^2}{(1+s)}\frac{\sigma^2}{2\mu^2}. \tag{B.2}$$

Finally, note that by our definition, $\Pr(\tilde{d}_1>0) = \Pr(\tilde{d}>\mu)$ and $\Pr(\tilde{d}_2>0) = \Pr(\tilde{d}<\mu)$. Thus, by inequalities (B.1)–(B.2), it follows that

$$1 \geq \Pr(\tilde{d}>\mu) + \Pr(\tilde{d}<\mu) \geq \frac{\sigma^2}{\mu^2}\frac{(1-s)}{(1+s)}, \tag{B.3}$$

which gives us the lower bound on $s$. Observe that if $s = \frac{\sigma^2-\mu^2}{\sigma^2+\mu^2}$, then inequalities (B.1)–(B.3) tight. This implies that

$$\begin{array}{ll} \Pr(\tilde{d}>\mu) = \frac{\mu^2}{\mu^2+\sigma^2}, & \mathrm{E}((\tilde{d}-\mu)_+) = \frac{\mu\sigma^2}{\mu^2+\sigma^2}, \\ \Pr(\tilde{d}<\mu) = \frac{\sigma^2}{\mu^2+\sigma^2}, & \mathrm{E}((\mu-\tilde{d})_+) = \frac{\mu\sigma^2}{\mu^2+\sigma^2}. \end{array}$$

Therefore, we have the following conditional moments:

$$\begin{array}{ll} \mathrm{E}(\tilde{d}_1|\tilde{d}>\mu) = \frac{\sigma^2}{\mu}, & \mathrm{E}(\tilde{d}_1^2|\tilde{d}>\mu) = \frac{\sigma^4}{\mu^2}, \\ \mathrm{E}(\tilde{d}_2|\tilde{d}<\mu) = \mu, & \mathrm{E}(\tilde{d}_2^2|\tilde{d}<\mu) = \mu^2, \end{array}$$

where the second moments follow from the expression for $s$. Thus, $\tilde{d}_1$ and $\tilde{d}_2$ can only have one positive mass above and below $\mu$ respectively. Thus, $\tilde{d}$ can only have strictly positive mass on 0 and $(\mu^2+\sigma^2)/\mu$, and a nonnegative mass at $\mu$. However, since $\Pr(\tilde{d}>\mu) + \Pr(\tilde{d}<\mu) = 1$, then the only possible distribution is a two-point distribution.

Now let us verify that we can construct a nonnegative distribution for any triplet $(\mu, \sigma, s)$ satisfying the conditions in Proposition 2.1. Consider the two-point distribution:

$$\tilde{d} = \begin{cases} \mu - \sigma\sqrt{\frac{1-s}{1+s}}, & \text{w.p. } \frac{1+s}{2}, \\ \mu + \sigma\sqrt{\frac{1+s}{1-s}}, & \text{w.p. } \frac{1-s}{2}. \end{cases}$$

These support points are nonnegative in the range (2.2). In fact, $s$ can be arbitrarily close but never equal to one since, by definition,

$$s = \frac{\mathrm{E}((\tilde{d}-\mu)_+^2) - \mathrm{E}((\mu-\tilde{d})_+^2)}{\sigma^2} = 1 - \frac{2}{\sigma^2}\mathrm{E}((\mu-\tilde{d})_+^2).$$

Since we assume that $\sigma > 0$, then there exists $\tilde{d} < \mu$ with nonzero probability. □

### B.2 Proof of Theorem 2.1

Consider the primal problem

$$\begin{aligned}
\inf_{f} \quad & \mathrm{E}_f\left(\min\{\tilde{d}, q\}\right) \\
\text{s.t.} \quad & \mathrm{E}_f\left((\tilde{d}-\mu)^+\right) - \mathrm{E}\left((\mu-\tilde{d})_+\right) = 0, \\
& \mathrm{E}_f\left((\tilde{d}-\mu)_+^2\right) = \tfrac{(1+s)}{2}\sigma^2, \\
& \mathrm{E}_f\left((\mu-\tilde{d})_+^2\right) = \tfrac{(1-s)}{2}\sigma^2, \\
& \mathrm{E}_f(1) = 1, \quad f(\tilde{d}) \geq 0, \quad \forall \tilde{d} \geq 0.
\end{aligned}$$

The constraints are equivalent to having $\tilde{d}$ be a nonnegative random variable with moments $\mathrm{E}(\tilde{d}) = \mu$, $\mathrm{var}(\tilde{d}) = \sigma^2$, and $\mathrm{sv}(\tilde{d}) = s\sigma^2$. Define $m_1 = \frac{(1+s)}{2}\sigma^2, m_2 = \frac{(1-s)}{2}\sigma^2$. Proposition 2.1 is equivalent to the following: $\frac{m_2}{m_1}(m_1+m_2) \leq \mu^2$ and $m_1, m_2 > 0$. Then we can write the dual as

$$\begin{aligned}
\sup_{t,r,y_1,y_2} \quad & t + m_1 y_1 + m_2 y_2 \\
\text{s.t.} \quad & t + r(x-\mu) + y_1(x-\mu)^2 \leq \min\{x, q\}, \quad \forall x \geq \mu, \\
& t - r(\mu - x) + y_2(\mu-x)^2 \leq \min\{x, q\}, \quad \forall 0 \leq x \leq \mu.
\end{aligned}$$

By a simple transformation, the dual problem can be rewritten as

$$\begin{aligned}
\sup_{t,r,y_1,y_2} \quad & t + m_1 y_1 + m_2 y_2 \\
\text{s.t.} \quad & t + rx + y_1 x^2 \leq \min\{x+\mu, q\}, \quad \forall x \geq 0, \\
& t - rx + y_2 x^2 \leq \min\{\mu - x, q\}, \quad \forall 0 \leq x \leq \mu.
\end{aligned}$$

Throughout this section, we will refer to this last formulation as the dual problem. Given a set of variables $t, r, y_1, y_2, \mu, q$, we define the following quadratic functions: $g_1(x) = t + rx + y_1 x^2$, $g_2(x) = t - rx + y_2 x^2$, and the following piecewise linear functions: $f_1(x) = \min\{x+\mu, q\}$, $f_2(x) = \min\{\mu - x, q\}$. Checking dual feasibility of a set of parameters $t, r, y_1, y_2$ is equivalent to checking if $g_1(x) \leq f_1(x)$ for $x \geq 0$, and $g_2(x) \leq f_2(x)$ for $0 \leq x \leq \mu$. Our strategy in finding a closed-form solution for finding an optimal solution to the primal problem is by constructing primal and dual feasible solutions which achieve the same objective cost. By weak duality, the two solutions are primal and dual optimal respectively.

**Case 1: $q \leq \mu$.** Under this case, we find that $f_1(x) = q$ for all $x \geq 0$. Figure B.1 illustrates a dual feasible solution. Table 1 shows the primal and dual optimal solutions under three subcases, which we discuss briefly.

**Case 1(a): $0 \leq q \leq \frac{\mu}{2}$.** By Proposition 2.1, the support points of the primal distribution are well-defined and the probabilities are valid. Moreover, the range of $\pi$ is nonempty. This range guarantees that the nonzero support points are at least $\mu$. Then we can verify that this is a primal feasible distribution. Note that $g_1(x) = f_1(x)$, for all $x \geq 0$, while $g_2(x)$ is concave and intersects $g_2(x)$ at points 0

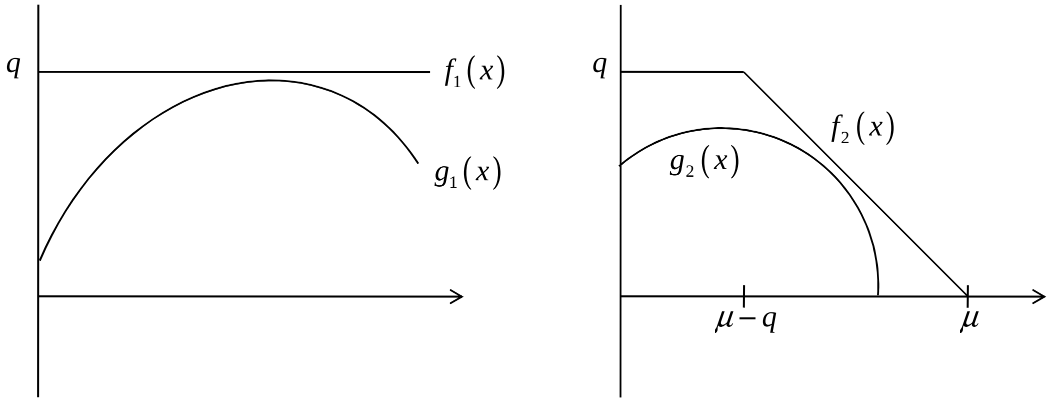


Figure B.1: Graphical illustration of functions satisfying feasibility conditions of the dual problem under the case when $q \leq \mu$.

and $\mu$. Moreover, $g_2(x) \leq q$ for all $x$. If $q \leq \frac{\mu}{2}$, then $g_2'(\mu) \geq -1$ and $g_2(x) \leq \mu - x$ for all $0 \leq x \leq \mu$. Thus, it is dual feasible.

**Case 1(b):** $\boldsymbol{\frac{\mu}{2} \leq q \leq \mu - \frac{m_2}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}}}$**.** Note that for the range of $q$ under this case, the support points are well-defined and the probabilities are valid. Moreover, the range of $\pi$ is nonempty and guarantees that the last two support points are at least $\mu$. We can then verify that the distribution is primal feasible. Note that $g_1(x) = f_1(x)$ for all $x \geq 0$, while $g_2(x) \leq f_2(x)$ for all $x$. Thus, the dual solution is always feasible.

**Case 1(c):** $\boldsymbol{\mu - \frac{m_2}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}} \leq q \leq \mu}$**.** Under Proposition 2.1, the support points are always nonnegative. It is easy to check that it is a primal feasible distribution. Moreover, $g_1(x)$ intersects $q$ and $g_2(x)$ intersects $\mu - x$ at exactly one nonnegative point each. Note that the dual solution is feasible if $g_1'(0) \geq 0$, or $q \geq \mu - \frac{m_2}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}}$.

**Case 2:** $\boldsymbol{q \geq \mu}$**.** Under this case, it is not difficult to verify that $f_2(x) = \mu - x$ for all $x \geq 0$. Figure B.2 shows the functions corresponding to a dual feasible solution. Table 2 shows the primal and dual optimal solutions under three subcases.

**Case 2(a):** $\boldsymbol{\mu \leq q \leq \mu + \frac{m_1}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}}}$**.** Note that the primal and dual optimal solutions are exactly the same as that of Case 1(c). We can verify that $g_1(x)$ intersects $q$ and $g_2(x)$ intersects $\mu - x$ at exactly one nonnegative point each. Dual feasibility is equivalent to the condition that $g_2'(0) \geq -1$, or $q \leq \mu + \frac{m_1}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}}$.

**Case 2(b):** $\boldsymbol{\mu + \frac{m_1}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}} \leq q \leq \mu + \frac{\mu m_1}{2 m_2}}$**.** Unlike the previous cases, it is not obvious that the

**Case 1(a):** $\boldsymbol{0 \leq q \leq \frac{\mu}{2}}$

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} 0, & \text{w.p. } \frac{m_2}{\mu^2}, \\ \frac{\mu}{\mu^2 - m_2}\left(\mu^2 - \sqrt{m_1(\mu^2 - m_2) - m_2^2}\sqrt{\frac{1-\pi}{\pi}}\right), & \text{w.p. } \pi\left(1 - \frac{m_2}{\mu^2}\right), \\ \frac{\mu}{\mu^2 - m_2}\left(\mu^2 + \sqrt{m_1(\mu^2 - m_2) - m_2^2}\sqrt{\frac{\pi}{1-\pi}}\right), & \text{w.p. } (1-\pi)\left(1 - \frac{m_2}{\mu^2}\right), \end{cases}$$

for any $\pi \in \left[1 - \frac{m_2^2}{m_1(\mu^2 - m_2)}, 1\right)$

Dual Optimal Solution: $t = q, r = 0, y_1 = 0, y_2 = -\frac{q}{\mu^2}$

Primal and Dual Optimal Cost: $q - \frac{m_2}{\mu^2}q$

**Case 1(b):** $\boldsymbol{\frac{\mu}{2} \leq q \leq \mu - \frac{m_2}{2}\sqrt{\frac{m_1 + m_2}{m_1 m_2}}}$

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} 2q - \mu, & \text{w.p. } \frac{m_2}{4(\mu-q)^2}, \\ \mu + \frac{2(\mu-q)}{4(\mu-q)^2 - m_2}\left(m_2 - \sqrt{4m_1(\mu-q)^2 - m_2(m_1+m_2)}\sqrt{\frac{1-\pi}{\pi}}\right), & \text{w.p. } \pi\left(1 - \frac{m_2}{4(\mu-q)^2}\right), \\ \mu + \frac{2(\mu-q)}{4(\mu-q)^2 - m_2}\left(m_2 + \sqrt{4m_1(\mu-q)^2 - m_2(m_1+m_2)}\sqrt{\frac{\pi}{1-\pi}}\right), & \text{w.p. } (1-\pi)\left(1 - \frac{m_2}{4(\mu-q)^2}\right), \end{cases}$$

for any $\pi \in \left[1 - \frac{m_2^2}{m_1(4(\mu-q)^2 - m_2)}, 1\right)$

Dual Optimal Solution: $t = q, r = 0, y_1 = 0, y_2 = \frac{-1}{4(\mu-q)}$

Primal and Dual Optimal Cost: $q - \frac{m_2}{4(\mu-q)}$

**Case 1(c):** $\boldsymbol{\mu - \frac{m_2}{2}\sqrt{\frac{m_1 + m_2}{m_1 m_2}} \leq q \leq \mu}$

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} \mu - m_2\sqrt{\frac{m_1+m_2}{m_1 m_2}}, & \text{w.p. } \frac{m_1}{m_1+m_2}, \\ \mu + m_1\sqrt{\frac{m_1+m_2}{m_1 m_2}}, & \text{w.p. } \frac{m_2}{m_2+m_2}, \end{cases}$$

Dual Optimal Solution:

$t = \frac{qm_2 + \mu m_1}{m_1 + m_2} - \frac{1}{2}\sqrt{\frac{m_1 m_2}{m_1 + m_2}}$,

$r = \frac{2}{m_1}\sqrt{\frac{m_1 m_2}{m_1 + m_2}}\left(\frac{m_1(q-\mu)}{m_1+m_2} + \frac{1}{2}\sqrt{\frac{m_1 m_2}{m_1+m_2}}\right)$,

$y_1 = \frac{-m_2}{m_1(m_1+m_2)}\left(\frac{m_1(q-\mu)}{m_1+m_2} + \frac{1}{2}\sqrt{\frac{m_1 m_2}{m_1+m_2}}\right)$,

$y_2 = \frac{m_1}{m_2(m_1+m_2)}\left(\frac{m_2(q-\mu)}{m_1+m_2} - \frac{1}{2}\sqrt{\frac{m_1 m_2}{m_1+m_2}}\right)$,

Primal and Dual Optimal Cost: $\frac{qm_2 + \mu m_1}{m_1 + m_2} - \sqrt{\frac{m_1 m_2}{m_1 + m_2}}$

Table 1: Primal and dual optimal solutions when $q \leq \mu$

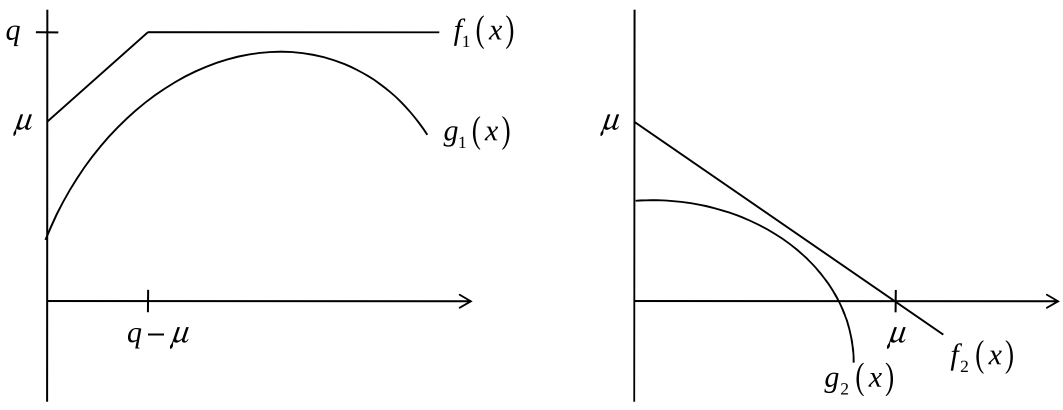


Figure B.2: Graphical illustration of functions satisfying feasibility conditions of the dual problem under the case when $q \geq \mu$.

range of $\pi$ is nonempty. To prove that the range is valid, consider the following random variable

$$\tilde{x} = \begin{cases} \frac{2(q-\mu)}{4(q-\mu)^2-m_1}\left(m_1 - \sqrt{4m_2(q-\mu)^2 - m_1(m_1+m_2)}\sqrt{\frac{1-\hat{\pi}}{\hat{\pi}}}\right), & \text{w.p. } \hat{\pi}, \\ \frac{2(q-\mu)}{4(q-\mu)^2+m_1}\left(m_1 - \sqrt{4m_2(q-\mu)^2 - m_1(m_1+m_2)}\sqrt{\frac{\hat{\pi}}{1-\hat{\pi}}}\right), & \text{w.p. } 1-\hat{\pi}, \end{cases}$$

for any $\hat{\pi} \in (0,1)$. Note that under the range of $q$, $\mathrm{E}(\tilde{x}^2) \leq \mu\mathrm{E}(\tilde{x})$, or equivalently, $\mathrm{var}(\tilde{x}) \leq \mathrm{E}(\tilde{x})\left(\mu - \mathrm{E}(\tilde{x})\right)$. Taking the square of both sides and adding $\mathrm{var}(\tilde{x})\left(\mu - \mathrm{E}(\tilde{x})\right)^2$, we find that

$$\mathrm{var}(\tilde{x})\left(\mathrm{var}(\tilde{x}) + (\mu - \mathrm{E}(\tilde{x}))^2\right) \leq (\mu - \mathrm{E}(\tilde{x}))^2\left(\mathrm{var}(\tilde{x}) + \mathrm{E}(\tilde{x})^2\right).$$

Substituting the values of the mean and variance of $\tilde{x}$, we find that this is equivalent to saying that the range of $\pi$ is nonempty. The range of values for $\pi$ ensures that the first two support points of $\tilde{d}$ are nonnegative and no more than $\mu$. We can then verify that the distribution is primal feasible. Note that $g_1(x) \leq f_1(x)$ for all $x$, while $g_2(x) = f_2(x)$ for all $x \geq 0$. Thus the solution is dual feasible.

**Case 2(c):** $\boldsymbol{q \geq \mu + \frac{\mu m_1}{2m_2}}$**.** Define $a = \frac{m_1}{2(q-\mu)^2} - \frac{m_2}{\mu(q-\mu)}$, and $b = 1 - \frac{m_2}{\mu^2}$. Note that $a \leq 0$ for the range of $q$ in Case 2(c). By Proposition 2.1, $b \in (0,1)$ and $2a + b \geq 0$ for all $q$. All the probabilities take values within the range of 0 to 1 by Proposition 2.1. It can be verified that the distribution is primal feasible. Note that $g_1(x) \leq f_1(x)$ for all $x$ and they intersect at exactly two points. Moreover, $g_2(x)$ is convex and intersects $f_2(x)$ at some negative value and at $\mu$. Thus, the dual solution is feasible.

If we multiply the primal optimal cost by $p$ and subtract $cq$ in all cases, then we get $\Pi^{MVS}(q)$. Combining all the cases and letting $m_1 = (1+s)\sigma^2/2$ and $m_2 = (1-s)\sigma^2/2$, we get the closed-form expression. □

**Case 2(a):** $\boldsymbol{\mu \le q \le \mu + \frac{m_1}{2}\sqrt{\frac{m_1+m_2}{m_1 m_2}}}$

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} \mu - m_2\sqrt{\frac{m_1+m_2}{m_1m_2}}, & \text{w.p. } \frac{m_1}{m_1+m_2}, \\ \mu + m_1\sqrt{\frac{m_1+m_2}{m_1m_2}}, & \text{w.p. } \frac{m_2}{m_2+m_2}, \end{cases}$$

Dual Optimal Solution:

$$t = \frac{qm_2+\mu m_1}{m_1+m_2} - \frac{1}{2}\sqrt{\frac{m_1m_2}{m_1+m_2}},$$
$$r = \frac{2}{m_1}\sqrt{\frac{m_1m_2}{m_1+m_2}}\left(\frac{m_1(q-\mu)}{m_1+m_2} + \frac{1}{2}\sqrt{\frac{m_1m_2}{m_1+m_2}}\right),$$
$$y_1 = \frac{-m_2}{m_1(m_1+m_2)}\left(\frac{m_1(q-\mu)}{m_1+m_2} + \frac{1}{2}\sqrt{\frac{m_1m_2}{m_1+m_2}}\right),$$
$$y_2 = \frac{m_1}{m_2(m_1+m_2)}\left(\frac{m_2(q-\mu)}{m_1+m_2} - \frac{1}{2}\sqrt{\frac{m_1m_2}{m_1+m_2}}\right),$$

Primal and Dual Optimal Cost: $\frac{qm_2+\mu m_1}{m_1+m_2} - \sqrt{\frac{m_1m_2}{m_1+m_2}}$

---

**Case 2(b):** $\boldsymbol{\mu + \frac{m_1}{2}\sqrt{\frac{m_1+m_2}{m_1m_2}} \le q \le \mu + \frac{\mu m_1}{2m_2}}$

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} \mu - \frac{2(q-\mu)}{4(q-\mu)^2-m_1}\left(m_1 + \sqrt{4m_2(q-\mu)^2 - m_1(m_1+m_2)}\sqrt{\frac{\pi}{1-\pi}}\right), & \text{w.p. } (1-\pi)\left(1 - \frac{m_1}{4(q-\mu)^2}\right), \\ \mu - \frac{2(q-\mu)}{4(q-\mu)^2-m_1}\left(m_1 - \sqrt{4m_2(q-\mu)^2 - m_1(m_1+m_2)}\sqrt{\frac{1-\pi}{\pi}}\right), & \text{w.p. } \pi\left(1 - \frac{m_1}{4(q-\mu)^2}\right), \\ 2q-\mu, & \text{w.p. } \frac{m_1}{4(\mu-q)^2}, \end{cases}$$

for any $\pi \in \left[1 - \frac{m_2^2}{m_1(4(\mu-q)^2-m_2)}, 1\right)$

Dual Optimal Solution: $t = \mu, r = 1, y_1 = \frac{-1}{4(q-\mu)}, y_2 = 0$

Primal and Dual Optimal Cost: $\mu - \frac{m_1}{4(q-\mu)}$

---

**Case 2(c):** $\boldsymbol{q \ge \mu + \frac{\mu m_1}{2m_2}}$**.** Define $a = \frac{m_1}{2(q-\mu)^2} - \frac{m_2}{\mu(q-\mu)}$, and $b = 1 - \frac{m_2}{\mu^2}$.

Primal Optimal Distribution:

$$\tilde{d} = \begin{cases} 0, & \text{w.p. } 1-b, \\ \mu + (q-\mu)\left(1 - \sqrt{\frac{2a+b}{b}}\right), & \text{w.p. } \frac{1}{2}\left(b + \left(b - \frac{(1-b)\mu}{(q-\mu)}\right)\sqrt{\frac{b}{2a+b}}\right), \\ \mu + (q-\mu)\left(1 + \sqrt{\frac{2a+b}{b}}\right), & \text{w.p. } \frac{1}{2}\left(b - \left(b - \frac{(1-b)\mu}{(q-\mu)}\right)\sqrt{\frac{b}{2a+b}}\right). \end{cases}$$

Dual Optimal Solution:

$$t = \frac{1}{2}\left(q + \mu - \frac{(q-\mu)(a+b)}{\sqrt{b(2a+b)}}\right),$$
$$r = \frac{1}{2}\left(1 + \sqrt{\frac{b}{2a+b}}\right),$$
$$y_1 = \frac{-1}{4(q-\mu)}\sqrt{\frac{b}{2a+b}},$$
$$y_2 = \frac{q-\mu}{2\mu^2}\left(\frac{(a+b)}{\sqrt{b(2a+b)}} - 1\right) + \frac{1}{2\mu}\left(\sqrt{\frac{b}{2a+b}} - 1\right),$$

Primal and Dual Optimal Cost: $\frac{1}{2}\left(\mu + bq - (q-\mu)\sqrt{b(2a+b)}\right)$

Table 2: Primal and dual optimal solutions when $q \ge \mu$

### B.3 Proof of Theorem 2.2

The worst-case profit $MVS(q)$ is concave and continuous in $q$. Note that unless the profit is maximized by a range of $q$, the optimal quantity can only occur at 0 or in Regions (ii), (iv) or (v) (as defined in Theorem 2.1). Let $q^*$ be an order quantity that maximizes $MVS(q)$. Define

$$\begin{aligned} q^*_{(ii)} &= \mu - \frac{\sigma}{2}\sqrt{\frac{(1-s)p}{2(p-c)}}, \\ q^*_{(iv)} &= \mu + \frac{\sigma}{2}\sqrt{\frac{(1+s)p}{2c}}, \\ q^*_{(v)} &= \frac{\mu}{b} + \frac{(pb-2c)}{2b}\sqrt{\frac{(1+s)\sigma^2 b - 2(1-b)^2\mu^2}{2c(pb-c)}}, \end{aligned}$$

which are the unconstrained maximizers of the aggregate functions in Regions (ii), (iv) and (v), respectively. Suppose $\frac{c}{p} > 1 - \frac{(1-s)\sigma^2}{2\mu^2}$. Then $MVS(q)$ is strictly decreasing in Region (i). This implies that the function is strictly decreasing for $q \geq 0$. Thus, $q^* = 0$. Note that $q^*_{(ii)}$ lies in Region (ii) if $\frac{1}{2}(1-s) \leq \frac{c}{p} \leq 1 - \frac{(1-s)\sigma^2}{2\mu^2}$. Under this case, $q^* = q^*_{(ii)}$. Similarly, $q^*_{(iv)}$ lies in Region (iv) if $\frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2} \leq \frac{c}{p} \leq \frac{1}{2}(1-s)$. If this is true, then $q^* = q^*_{(iv)}$. Finally, if $\frac{c}{p} < \frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}$, then $MVS(q)$ is increasing in Regions (i) to (iv). Therefore, the maximum is attained in Region (v). Thus, $q^* = q^*_{(v)}$. □

### B.4 Proof of Theorem 3.1

Define the following function:

$$g(v,q) = \inf_{f\sim(\mu,\sigma^2,\Re^+)} \mathrm{E}_f\left(\min\left\{v - M + (p-c)q, v - M + p\tilde{d} - cq, 0\right\}\right).$$

Note that for a fixed $v$, $q$, the minimum of the three piece linear function can be achieved by at most two pieces. Thus, we can apply Scarf's closed-form expression in intervals.

**Case 1: $v \leq M - (p-c)q$.** We have that

$$g(v,q) = v - M - cq + \inf_{f\sim(\mu,\sigma^2,\Re^+)} p\mathrm{E}_f\left(\min\left\{\tilde{d}, q\right\}\right)$$

which, if $q \leq \frac{\mu^2+\sigma^2}{2\mu}$, is equal to

$$v - M - cq + pq\frac{\mu^2}{\mu^2+\sigma^2},$$

or otherwise, if $q \geq \frac{\mu^2+\sigma^2}{2\mu}$, is equal to

$$v - M - cq + \frac{p(\mu+q)}{2} - \frac{p}{2}\sqrt{(q-\mu)^2+\sigma^2}.$$

**Case 2: $v \geq M - (p-c)q$.** We have that

$$g(v,q) = v - M - cq + \inf_{f\sim(\mu,\sigma^2,\Re^+)} p\mathrm{E}_f\left(\min\left\{\tilde{d}, \frac{-v+M+cq}{p}\right\}\right).$$

**Case 2(a): $M - (p-c)q \leq v < M + cq - p\frac{\mu^2+\sigma^2}{2\mu}$.** This corresponds to the nondegenerate case of the Scarf closed-form expression.

$$g(v,q) = \frac{1}{2}\left(v - M - cq + p\mu - \sqrt{(v - M - cq + p\mu)^2 + p^2\sigma^2}\right)$$

**Case 2(b):** $\boldsymbol{M + cq - p\frac{\mu^2+\sigma^2}{2\mu} \leq v < M + cq}$. This corresponds to the degenerate case.

$$g(v,q) = (v - M - cq)\frac{\sigma^2}{\mu^2+\sigma^2}$$

**Case 2(c):** $\boldsymbol{v \geq M + cq}$. It is easy to check that due to the nonnegative support, under this case, $g(v,q) = 0$.

Note that the interval of Case 2(a) can be empty. In particular, it is empty if $q \leq \frac{\mu^2+\sigma^2}{2\mu}$. Thus, summarizing Cases 1 and 2, we find that if $q \leq \frac{\mu^2+\sigma^2}{2\mu}$,

$$v - \frac{1}{1-\alpha}g(v,q) = \begin{cases} v - \frac{1}{1-\alpha}\left(v - M - cq + pq\frac{\mu^2}{\mu^2+\sigma^2}\right), & \text{for } v < M - (p-c)q, \\ v - \frac{1}{1-\alpha}(v - M - cq)\frac{\sigma^2}{\mu^2+\sigma^2}, & \text{for } M - (p-c)q \leq v < M + cq, \\ v, & \text{for } v \geq M + cq, \end{cases}$$

whereas, if $q \geq \frac{\mu^2+\sigma^2}{2\mu}$,

$$v - \tfrac{1}{1-\alpha}g(v,q) = \begin{cases} v - \frac{1}{1-\alpha}\left(v - M - cq + \frac{p(\mu+q)}{2} - \frac{p}{2}\sqrt{(q-\mu)^2+\sigma^2}\right), & \text{for } v < M - (p-c)q, \\ v - \frac{1}{2(1-\alpha)}\left(v - M - cq + p\mu - \sqrt{(v-M-cq+p\mu)^2 + p^2\sigma^2}\right), & \text{for } M - (p-c)q \leq v < M + cq - p\frac{\mu^2+\sigma^2}{2\mu}, \\ v - \frac{1}{1-\alpha}(v - M - cq)\frac{\sigma^2}{\mu^2+\sigma^2}, & \text{for } M + cq - p\frac{\mu^2+\sigma^2}{2\mu} \leq v < M + cq, \\ v, & \text{for } v \geq M + cq. \end{cases}$$

Now, note that

$$\text{AmCVaR}_\alpha(q) = \inf_{v\in\Re}\left\{v - \frac{1}{1-\alpha}g(v,q)\right\}.$$

The objective function is convex in $q$. If $q \leq \frac{\mu^2+\sigma^2}{2\mu}$, then the minimizer is easily computed since the objective consists of only linear functions. In particular, if $\alpha \leq \frac{\mu^2}{\mu^2+\sigma^2}$, then $v^* = M-(p-c)q$. Otherwise, $v^* = M + cq$.

Now let us focus on the case where $q \geq \frac{\mu^2+\sigma^2}{2\mu}$. Note that the objective is a piecewise convex function, where all the pieces are linear functions of $v$, except for the second piece. The unconstrained minimizer of the second piece is

$$v_2^* = M + cq - p\mu - \frac{p\sigma(1-2\alpha)}{2\sqrt{\alpha(1-\alpha)}}.$$

If the third linear piece is decreasing in $v$, that is $\alpha \geq \frac{\mu^2}{\mu^2+\sigma^2}$, then the global minimizer is $v^* = M + cq$. Moreover, we can check that if $\alpha \leq \frac{\mu^2}{\mu^2+\sigma^2}$, then $v_2^* \leq M + cq - p\frac{\mu^2+\sigma^2}{2\mu}$. Thus, under this case, the global minimizer is $v^* = v_2^*$ if $v_2^* \geq M-(p-c)q$, or equivalently, $q \geq \mu + \frac{\sigma(1-2\alpha)}{2\sqrt{\alpha(1-\alpha)}}$. Otherwise, $v^* = M-(p-c)q$. Substituting the values for $v^*$, we get the closed-form expression for $\text{AmCVaR}_\alpha(q)$. □

## B.5 Proof of Theorem 3.2

Note that if $\alpha \geq \frac{\mu^2}{\mu^2+\sigma^2}$, then $q^* = 0$, and $\text{AmCVaR}_\alpha(q^*) = M$. Now let us focus on the case when $\alpha \leq \frac{\mu^2}{\mu^2+\sigma^2}$. Under this case, $\text{AmCVaR}_\alpha(q)$ is a three-piece function of $q$. All the pieces are linear, except for the second one, whose unconstrained minimizer is

$$q_2^* = \mu - \frac{\sigma(1-2\beta)}{2\sqrt{\beta(1-\beta)}},$$

where we define $\beta = \left(1 - \frac{c}{p}\right)(1-\alpha)$. Note that if $\beta < \frac{\sigma^2}{\mu^2+\sigma^2}$, then $\mathrm{AmCVaR}_\alpha(q)$ is increasing for $q \geq 0$. Thus, $q^* = 0$ and $\mathrm{AmCVaR}_\alpha(q^*) = M$. Now let us focus on the case where $\beta \geq \frac{\sigma^2}{\mu^2+\sigma^2}$. Note that $q_2^* \geq \frac{\mu^2+\sigma^2}{2\mu}$ under this case. If in addition, $\frac{2\beta-1}{\sqrt{\beta(1-\beta)}} < \frac{1-2\alpha}{\sqrt{\alpha(1-\alpha)}}$, then $q^* = q_2^*$. In fact, we can show that this is always true. Note that $0 < \beta \leq 1-\alpha$, since $0 < c \leq p$. Moreover, $\alpha(1-\alpha) < \beta(1-\beta)$ if and only if $\alpha < \frac{\lambda}{1+\lambda}$, where $\lambda = 1 - \frac{c}{p}$. Consider the case when $0 \leq 2\beta - 1 < 1 - 2\alpha$. This implies that $\alpha \leq 1 - \frac{1}{2\lambda} \leq \frac{\lambda}{1+\lambda}$. Thus, $\frac{2\beta-1}{\sqrt{\beta(1-\beta)}} < \frac{1-2\alpha}{\sqrt{\alpha(1-\alpha)}}$. This condition is obviously satisfied if $2\beta - 1 < 0 \leq 1 - 2\alpha$. We can repeat the same exercise for the case when $2\beta - 1 \leq 1 - 2\alpha \leq 0$, and find that $\alpha(1-\alpha) > \beta(1-\beta)$. Thus, $\frac{2\beta-1}{\sqrt{\beta(1-\beta)}} < \frac{1-2\alpha}{\sqrt{\alpha(1-\alpha)}}$. Thus,

$$q^* = \begin{cases} 0, & \text{if } \alpha \geq \frac{\mu^2}{\mu^2+\sigma^2} \text{ or } \beta < \frac{\sigma^2}{\mu^2+\sigma^2}, \\ \mu - \frac{\sigma(1-2\beta)}{2\sqrt{\beta(1-\beta)}}, & \text{if } \beta \geq \frac{\sigma^2}{\mu^2+\sigma^2}, \end{cases}$$

where $\beta = \left(1 - \frac{c}{p}\right)(1-\alpha)$. Substituting $q^*$ into the original expression, we get the optimal ambiguous $\alpha$-CVaR. □

## References


[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] P. Berck and J. M. Hihn. Using the semivariance to estimate safety-first rules. *American Journal of Agricultural Economics*, 64(2):298–300, 1982.

[3] S. A. Bond. Asymmetry and downside risk in foreign exchange markets. University of Cambridge. Working Paper, 2000.

[4] S. A. Bond and S. E. Satchell. Statistical properties of the sample semi-variance. *Applied Mathematical Finance*, 9:219–239, 2002.

[5] G. C. Calafiore. Ambiguous risk measures and optimal robust portfolios. *SIAM Journal on Optimization*, 18(3):853–877, 2007.

[6] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, pages 643–669, 1961.

[7] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*, volume 27 of *de Gruyter Studies in Mathematics*, section 2.5, pages 86–99. Walter de Gruyter, Germany, second edition, 2004.

[8] G. Gallego and I. Moon. The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society*, 44(8):825–834, 1993.

[9] I. Gilboa and D. Schmeidler. Maximin Expected Utility Theory with Non-Unique. *Journal of Mathematical Economics*, 18:141–153, 1989.

[10] S. He, J. Zhang, and S. Zhang. Bounding probability of small deviations: A fourth moment approach. Working Paper, 2007.

[11] K. Isii. On the sharpness of Chebyshev-type inequalities. *Annals of the Institute of Statistical Mathematics*, 12:185–197, 1963.

[12] K. Jansen, J. Haezendonck, and M. J. Goovaerts. Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance: Mathematics and Economics*, 5:315–334, 1986.

[13] A. W. Lo. Semi-parametric upper bounds for option prices and expected payoffs. *Journal of Financial Economics*, 19(2):373–387, 1987.

[14] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Its Applications*, 284:193–228, 1998.

[15] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments.* Wiley, New York, 1959.

[16] K. Natarajan, M. Sim, and J. Uichanco. Tractable robust expected utility and risk models for portfolio optimization. To appear in Mathematical Finance, 2008.

[17] Y. Nesterov and A. Nemirovski. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathemathics.* SIAM, Philadelphia, PA, 1994.

[18] G. Perakis and G. Roels. Regret in the newsvendor model with partial information. *Operations Research*, 56(1):188–203, 2008.

[19] I. Pólik and T. Terlaky. A survey of the S-lemma. *SIAM Review*, 49(3):371–418, 2007.

[20] I. Popescu and D. Bertsimas. On the relation between option and stock prices: a convex optimization approach. *Operations Research*, 50(2):358–374, 2002.

[21] A. Ridder, E. Van der Laan, and M. Salomon. How larger demand variability may lead to lower costs in the newsvendor problem. *Operations Research*, 46(6):934–936, 1998.

[22] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[23] L. J. Savage. The theory of statistical decisions. *Journal of American Statistical Association*, 46:55–67, 1951.

[24] H. Scarf. A min-max solution to an inventory problem. In K. J. Arrow, S. Karlin, and H. Scarf, editors, *Studies in the Mathematical Theory of Inventory and Production*, chapter 12, pages 201–209. Stanford University Press, 1958.

[25] A. De Schepper and B. Heijnen. Distribution-free option pricing. *Insurance: Mathematics and Economics*, 40:179–199, 2007.

[26] M. Song, D. Klabjan, and D. Simchi-Levi. Robust stochastic lot-sizing by means of histograms. 2007.

[27] F. A. Sortino and H. J. Forsey. On the use and misuse of downside risk. *Journal of Portfolio Management*, 22(2):35–42, 1996.

[28] P. T. von Hippel. Mean, median, and skew: Correcting a textbook rule. *Journal of Statistics Education*, 13(2), 2005.

[29] Jun ya Gotoh and Yuichi Takanoa. Newsvendor solutions via conditional value-at-risk minimization. *European Journal of Operational Research*, 179(1):80–96, 2007.

[30] Zhe George Zhang Youhua Chen, Minghui Xu. Technical note: A risk-averse newsvendor model under the cvar criterion. *To Appear in Operations Research*.

[31] L. F. Zuluaga, J. Peña, and D. Du. Extensions of Lo's semiparametric bound for European call options. *European Journal of Operational Research*, 198(2):557–570, 2008.