Supplier Choice: Market Selection under Uncertainty

by

Zohar Maia-Aliya Strinka

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Industrial and Operations Engineering) in the University of Michigan 2016

Doctoral Committee:

Professor Edwin Romeijn, Co-Chair
Professor Jon Lee, Co-Chair
Associate Professor Damian R. Beil
Professor Xiuli Chao
Assistant Professor Siqian Shen

©Zohar Maia-Aliya Strinka

2016

For Riker and Saria.

ACKNOWLEDGMENTS

This thesis collects much of my work during my time at the University of Michigan. It would not be what it is today were it not for the advice, guidance, and support of many people.

First I want to thank my advisor Prof. Edwin Romeijn for his guidance and support throughout my time in graduate school. I started at Michigan with a lot to learn and his consistent encouragement and mentoring has helped me to succeed here. This thesis would not be what it is today without his help, nor would I have learned how to balance the big picture with the mathematical details. I also want to thank my co-advisor Prof. Jon Lee for his career advice and helpful research discussions in addition to agreeing to serve as a co-chair on my committee when the need arose.

Thanks as well to my committee members. Prof. Damian Beil for sharing his perspective on research and publishing, and for helping me get comfortable with game theory. His mentoring has helped me discover my personal style in research. Prof. Siqian Shen had many helpful comments and careful reading. Prof. Xiuli Chao also shared helpful insights and engaging lectures.

I also want to thank the faculty and staff in the Industrial and Operations Engineering department. In particular Prof. Marina Epelman for her help as I started to learn mathematical rigor and Mint for conversation and technical support. To my undergrad mentor Bruce Schmeiser, I appreciate the support and encouragement to pursue a PhD.

In addition to the above, many friends helped me through “rubber duck proofs” and other discussions. Thanks to Gregg Schell for five great years as office mates, Troy Long for game nights, Arleigh Waring for having a helpful dissertation, Victor Wu for help with figures, and Kathryn Schumacher for conference adventures. Thanks as well (in no particular order) to Johannes, Sara, Robert, Vernnaliz, Rama, Kayse, Rose, Brandon, Jake, Katie T. and Katie R., Jax, and Ursula.

I also want to thank my friend Tamar Chestnut for her unwavering support. To Chris Miller for discussing work, or not. To Lisa and Tony Putman for their guidance. To John and Mary Strinka for their support as well as Alex, Greg, Katya, Kristina, and Molly. To my aunt Clove for many conversations and support. To my brother Tai and my grandparents Rita and Gerald for their encouragement. Most of all, I would like to thank my husband Stephen for everything he does and has done, including waiting five, and then one more years before moving to mountains.

I also want to thank the institutions which provided me funding to complete this thesis: The Rackham Merit Fellowship and the National Defense Science and Engineering Graduate Fellowship.

Finally, thank you to the American Dialect Society for choosing singular “they” as the word of the year for 2015.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Supplier Choice: Market Selection under Uncertainty

by

Zohar Maia-Aliya Strinka

Co-Chairs: Edwin Romeijn and Jon Lee

Suppliers and Manufacturers generally have some say in which subset of all possible demand they will meet. In some cases that choice is implicit through pricing decisions and feature selection. Other times it is made explicitly by choosing only specific regions to stock a product in. This thesis includes models using both approaches and incorporates random demands. We present several methods for choosing a subset of all candidate customers given uncertain demands.

In this thesis we consider four models of demand selection. The first two research problems consider market selection, which has been studied in the literature. The Selective Newsvendor Problem (SNP) looks at a decision maker choosing a subset of candidate markets to serve, and then receiving revenues and paying newsvendor-type costs based on the selected collection. In this thesis we consider a generalization with normally distributed demands which includes a multi-period problem as a special case and develop both exact and

heuristic algorithms to solve it. When demands are not normally distributed, the problem is considerably more complex and is in general $\mathcal{NP}$-hard. We develop an approximation algorithm using sample average approximation and a rounding approach to efficiently solve the problem. In addition to the work on market selection, we propose two other models for demand selection. We study auctions as a tool for a supplier with a fixed capacity to allocate the limited supply to retailers with newsvendor-type costs. Finally, we present a model for a supplier who must ensure demand is met in all markets, but has the option to work with subsidiary suppliers to meet that demand.

CHAPTER 1

Introduction

1.1 Supplier choice problems

When a supplier is designing a new product, there are a number of decisions they need to make before it reaches customers. Those decisions include which features the product will have, what price to sell it for, how many options should be available, and a host of other choices. One common perspective is that decision makers have specific products which they then meet all demand for. In this thesis we take the more comprehensive view that suppliers make both implicit and explicit customer selection decisions which ultimately results in a subset of all possible demand being met.

Since there is no way for a single product to meet all needs for every person, designing something inherently involves deciding what your target market is. Further, suppliers will have different priorities which help them to define that target market. When a new product is being developed there are also parallel marketing decisions including if it will be sold in all regions, how limited supply should be allocated, and risk tolerance for having a limited supply. Solving all of these issues is tricky, especially when determining which simplifying assumptions to make to ensure tractability. There are several research areas which touch on this underlying question of how to choose a subset of all possible demand to meet. These include at a minimum marketing, pricing, revenue management, auctions, and market selection.

Both the literature in marketing and Operations Management (OM) have studied problems related to those in this thesis. The marketing approach is beyond the scope of this thesis, but see Krishnan and Ulrich [1] for a review of the product development decision literature which discusses the differences in approach between OM and marketing. Within OM, revenue management is the field which considers how to best balance price limited capacity, particularly when demand for that capacity arrives over time. Chiang et al. [2] provides an overview of the wide field of revenue management including pricing, auctions, and other techniques. Within revenue management, Yano and Gilbert [3] provide a review of the pricing literature as it relates to production and procurement decisions. Petruzzi and Dada [4] also look at pricing and procurement specifically for the newsvendor problem.

Auctions are sometimes used as a tool within revenue management to maximize profits. However, auctions and other market mechanisms have been gaining popularity within supply chain optimization as a means not only to increase revenues, but also to allow information revelation which facilitates coordination between independent agents. Krishna [5] provides an overview of auction theory including multi-unit auctions which are most relevant to supply chain. In the auction literature there are a variety of auction mechanisms for selling multiple units including the optimal auction presented in Maskin and Riley [6] and the efficient auction developed by Ausubel in [7]. Within supply chain, auctions have mostly been used for the reverse of the situation presented in this thesis. Specifically, a decision maker choosing which of many similar suppliers to order from. For example, Chen [8] presents optimal procurement strategies by combining contracts and auctions while Duenyas et al. [9] use a modified descending-price auction to accomplish the same goal.

While auctions and pricing can be used to implicitly select a subset of demand to meet, in some cases suppliers instead explicitly make that selection. Rhim et al. [10] study a market selection problem with deterministic demands which also includes competition. Geunes et al. [11] take an explicit market selection approach and present both deterministic and stochastic versions of the problem. Van den Heuvel et al. [12] provide complexity results for general market selection problems as well as special cases. The authors also show that for the profit maximization version of the general problem, not even a constant-factor approximation algorithm can solve the problem efficiently. Geunes et al. [13] do present a constant-factor approximation algorithm for the cost-minimization formulation of a deterministic market selection problem.

In addition to the work on deterministic market selection, there has been a stream of literature considering stochastic demands. Carr and Lovejoy [14] present the inverse newsvendor problem for choosing the demand distribution from an opportunity set given a capacity constraint. Taaffe et al. [15] present the selective newsvendor problem (SNP) where the supplier chooses a binary market selection and makes associated procurement decisions. Other papers have added to the work on the SNP including Chahar and Taaffe [16] which includes risk as an additional objective and Taaffe and Chahar [17] which considers all-or-nothing demands. Waring [18] also studies the SNP and considers risk-aversion measures including Conditional Value at Risk (CVaR) and Value at Risk (VaR).

The stochastic market selections discussed above build on much of the research on the newsvendor problem as a means of successfully incorporating uncertainty into supply chain problems. The newsvendor problem has been extensively studied as a model in retail for goods which experience a discount at the end of a selling period. Khouja [19] is a review paper which covers many of the extensions which have been considered for the single period newsvendor problem. The author's summary includes extensions to multiple locations and pricing as is studied in this thesis as well as a variety of other extensions including random yields, multi-echelon systems, and related objectives. Since this review paper, there has continued to be an active stream of literature using the newsvendor model to address important questions.

Qin et al. [20] provide a more recent review of the newsvendor problem. The authors build on [19] and present more recent contributions which include addressing marketing effort, elements of competition and contracts, and risk. Kraiselburd et al. [21] consider market effort and competition in a setting with newsvendor-type costs and find the value in vendor-managed inventory to be dependent on substitution rates and the effects of manufacturer effort. Gotoh and Takano [22] show that Conditional Value at Risk for a single-period newsvendor can be formulated as a Linear Program while Ahmed et al. [23] consider coherent risk measures as a whole in inventory problems. Weng [24] considers how manufacturers and retailers can use contracts to coordinate inventory decisions in a newsvendor setting and Chen and Xiao [25] consider outsourcing as a response to disruption risk and uncertain capacity.

In summary, this thesis builds on many of the most important issues which have been identified in the literature relating to procurement and inventory management. In Yano and Gilbert [3] the authors comment that they believe in the future

> the most significant contributions will come not from continued examination of recognized trade-offs, but from expansive inquiry that breaks down traditional boundaries and identifies new issues.

This thesis takes an optimization-grounded view of suppliers choosing which subset of all possible demand to meet using newsvendor-type retailers as their customers. Using this perspective we identify explicit market selection, auctions, and outsourcing as tools the supplier can use to decide who to serve.

1.2 Chapter summaries

In this thesis, I extend the existing market selection literature for the selective newsvendor problem. In addition, I consider using an auction to implicitly choose a subset of customer demand. Finally, I present an outsourcing problem which demonstrates one more way suppliers have choice about which subset of demand they will serve. The remainder of this section will present a brief overview of the thesis.

1.2.1 Chapter 2: The SNP with normally distributed demands

In this chapter we study a class of selective newsvendor problems, where a decision maker has a set of raw materials each of which can be customized shortly before satisfying demand. The goal is then to select which subset of customizations maximizes expected profit. We show that certain multi-period and multi-product versions of the SNP fall within our problem class. Under the assumption that the demands are independent and normally, but not necessarily identically, distributed we show that some problem instances from our class can be solved efficiently using an attractive sorting property that was also established in the literature for some related problems. For our general model we use the KKT conditions to develop an exact algorithm that is efficient in the number of raw materials. In addition, we develop a class of heuristic algorithms. In a numerical study we compare the performance of the algorithms, and the heuristics are shown to have excellent performance and running times as compared to available solvers.

1.2.2 Chapter 3: Approximation algorithms

While most of the existing literature on the SNP assumes that demands are normally distributed. In this chapter we consider more general demand distributions and use an approximation algorithm to solve the resulting problem efficiently. We study a class of problems with both binary selection decisions and continuous variables (e.g., procurement quantity) which result in stochastic rewards and costs. The rewards are received based on the decision maker's selection and the costs depend both on all decision variables and realizations of the stochastic variables. We consider a family of risk-based objective functions that contains the more traditional risk-neutral expected-value objective as a special case. We use a combination of rounding and sample average approximation to produce solutions which are guaranteed to be close to the optimal solution with high probability. We also provide an empirical comparison of the performance of the algorithms on a set of randomly generated test problems. We find that for our supply chain example, high-quality solutions can be found with small computational effort.

1.2.3 Chapter 4: Allocating goods via an efficient auction

In this chapter we study a supplier who has a fixed and limited inventory to sell to customers. We consider customers who are themselves retailers facing uncertain demand with newsvendor-type costs. Because demand is uncertain and the retailers face newsvendor-type costs, they order gradually decreasing quantities from the supplier for higher per-unit prices. This trade-off between price charged and order quantity motivates the supplier to use an auction to allocate the limited goods. We consider a supplier who wants to ensure that those with the highest value for an object receive it. This is called an efficient allocation and is most commonly used for government decision making or other times when welfare maximization is important. In addition, since the retailer costs are driven by customer demand, an efficient allocation could be a reflection of the underlying end-customer demands. We use the work of Ausubel [7] and assume retailers face newsvendor-type costs and only know their own private information as well as existing bids. We present auction mechanisms for discrete and continuous auctions and also introduce an "ϵ-efficient" auction mechanism.

1.2.4 Chapter 5: Outsourcing to a subsidiary

While Chapters 2 and 3 study the selective newsvendor problem, in some settings the supplier may need to ensure that all market demand is met. However, even in those cases the supplier may have the option to outsource markets to subsidiaries if they so choose. In this chapter, we consider such a setting of a single supplier and a number of subsidiaries who are all able to produce the desired good. The supplier then assigns each market to either themselves or one of the subsidiaries. The supplier tries to maximize their own expected profit while ensuring that all subsidiaries have positive expected profit through transfer payments if necessary. By solving both this problem and the equivalent SNP, we are able to study the resulting market assignment decisions and gain insight into the market structure of the product. In this chapter we also include numerical results which can be compared to the work in Chapter 3.

CHAPTER 2

The Selective Newsvendor Problem with Normally Distributed Demands

This chapter discusses a project I completed jointly with H. Edwin Romeijn and Jingchen Wu which appeared in the April 2013 issue of Omega [26].

2.1 Introduction

In this chapter we study a class of selective newsvendor problems (SNPs) that generalizes the classical newsvendor problem by incorporating a degree of flexibility regarding the shape of the demand distribution faced by an inventory manager into the decision making progress. In particular, our generic model considers a set of raw materials that can be customized immediately prior to satisfying demand. The raw materials could be physically different items, but also simply a single item in different periods. The processes by which customization can take place are identical for each of the raw materials; e.g., we can think of coloring, packaging, etc., or preparation for satisfying demand in a particular market or segment. The selection flexibility lies in the ability to invest in a collection of customization methods or options. This model is applicable in several important practical settings. For example, it could be used as a prescriptive model for a manufacturer who has the opportunity to make an optimal selection, but also as a tool for gaining managerial insights for a manufacturer who is considering a modification of their selection. In addition, the type of selection models that we will consider in this chapter also often appear as a subproblem for solving assignment, facility location, or other more comprehensive models (see, e.g., Freling et al. [27], Huang et al. [28], Shen et al. [29], Shu et al. [30], Taaffe et al. [31]).

As mentioned above, our problem is based on the single-period newsvendor problem (see Porteus [32] for a general overview of stochastic inventory models). Eppen [33] considered a generalization of this problem to multiple locations, which allows for a reduction in the expected costs associated with variability in demand by considering all locations together and planning for aggregate demand. This observation has more recently led to research on market selection problems where the manufacturer has a choice of a set of markets (e.g., locations) that may be served. Taaffe et al. [15] first introduced a Selective Newsvendor Problem (SNP), there defined as a market selection problem with independent and normally distributed demands for each market. The authors demonstrate that such a market selection problem can be solved efficiently using a sorting algorithm that ranks the markets according to the ratio of net expected revenue to demand variance. Taaffe et al. [34] studied the case of all-or-nothing demand distributions, and Taaffe and Chahar [17] and Chahar and Taaffe [16] included risk as an additional objective. In related work, deterministic market selection problems with Economic Order Quantity (EOQ) (Geunes et al. [35]) and Economic Lot Sizing (ELS) costs (Van den Heuvel et al. [12]) were studied, while Geunes et al. [11] reviewed demand selection and assignment problems. Chen and Zhang [36] and Huang and Sošić [37] study allocations of profits for a newsvendor game. Testing if an allocation of profits is in the core of the game is closely related to selection problems.

In the context of this chapter, the basic SNP introduced and studied by Taaffe et al. [15] can be viewed as a customization selection problem corresponding to only a single raw material. In this chapter we (i) discuss a wider range of applications of this general problem class; and (ii) generalize that model to account for multiple raw materials. Taaffe et al. [15] showed that the case of a single raw material can be solved efficiently using a sorting-based algorithm. We develop an exact solution approach for the more general and computationally much more challenging class of selective newsvendor problems with multiple raw materials. Finally, we propose a class of heuristics inspired by the exact solution approaches and show, through extensive computational tests, that particular implementations of these are both effective and efficient. In Chapter 3 we go in a different direction and instead consider non-normally distributed demands. We then propose several algorithms which approximately and efficiently solve the SNP with nonnegative demands with high probability.

The remainder of the chapter is organized as follows. Section 2.2 describes the model and some applications. In Section 2.3 we develop exact and heuristic solution approaches, while Section 2.4 provides a variety of computational results. Finally, in Section 2.5 we summarize our results and conclude.

2.2 Problem formulation

2.2.1 Notation and general model

Let $A = \{1, \ldots, a\}$ be a set of *raw materials* that may be *customized* shortly before satisfying demand in one of several different ways, indexed by the set $N = \{1, \ldots, n\}$. The problem that we will study in this chapter is to determine a subset of customizations and a set of order quantities for the raw materials that maximize expected profit. To this end, define the binary decision variables $z_i = 1$ when customization i is selected and $z_i = 0$ otherwise ($i \in N$), as well as the continuous decision variables Q_j ($j \in A$), denoting the raw material order quantities. For convenience, let $Q = (Q_j, j \in A)^\top$ and $z = (z_i, i \in N)^\top$.

Let the random variable $\mathbf{D}_{ij}$ denote the demand for raw material $j \in A$ customized according to $i \in N$, and let $\mathbf{D}_j = (\mathbf{D}_{ij}, i \in N)^\top$ ($j \in A$) denote the corresponding demand vectors. Furthermore, let:

- F_i = fixed charge associated with selecting customization method $i \in N$;

- $\bar{r}_{ij}$ = unit revenue for raw material $j \in A$ customized according to $i \in N$;
- c_j = unit cost of raw material $j \in A$, and $c = (c_j, j \in A)^\top$;
- v_j = unit salvage value for raw material $j \in A$;
- e_j = unit expediting cost for raw material $j \in A$.

For convenience, we let $\mu_{ij} = E[\mathbf{D}_{ij}]$ (for $j \in A, i \in N$) and define $\bar{\mathcal{R}}_i = \sum_{j \in A} \bar{r}_{ij}\mu_{ij} - F_i$ (for $i \in N$). To ensure that the problem is meaningful, we assume that $e_j > c_j > v_j$ for all $j \in A$. The expected profit as a function of the decision variables can then be expressed as:

$$\bar{\mathcal{R}}^\top z - c^\top Q + \sum_{j \in A} v_j E\left[\left(Q_j - \mathbf{D}_j^\top z\right)^+\right] - \sum_{j \in A} e_j E\left[\left(\mathbf{D}_j^\top z - Q_j\right)^+\right].$$

It is easy to see that, given values for the selection variables, this problem decomposes into a traditional newsvendor problem for each raw material. This means that the optimal order quantity for raw material $j \in A$ is the $\rho_j \equiv \left(\frac{e_j - c_j}{e_j - v_j}\right)$-fractile of the distribution of demand for that raw material.

In the remainder of this chapter, we will follow Taaffe et al. [15] and assume that $\mathbf{D}_{ij} \sim$ Normal$(\mu_{ij}, \sigma_{ij}^2)$ (for $(i, j) \in N \times A$). In addition, we assume that, for each $j \in A$, the elements of $\mathbf{D}_j$ are independent; however, the vectors $\mathbf{D}_j$ may be dependent. Clearly, both the normality and independence assumptions may be violated in practice. In fact, this assumption is relaxed in Chapter 3 where approximation algorithms for such problems are developed. However, in this chapter we will limit ourselves to the special case since under these assumptions the optimization problem takes on an interesting form. We then develop exact and heuristic approaches for a more general class of problems that may be of independent interest.

Taaffe et al. [15] show that normality and independence of the demand vector (for fixed $j \in A$) can be used to further simplify the expected profit function. In particular, letting $s_{ij} = \sigma_{ij}^2$ ($i \in N$, $j \in A$), $s_j = (s_{ij}, i \in N)^\top$, $\mathcal{R}_i = \sum_{j \in A}(\bar{r}_{ij} - c_j)\mu_{ij} - F_i$, and

$\mathcal{R} = (\mathcal{R}_i, i \in N)^\top$ we obtain the optimization problem

$$\max_{z \in \{0,1\}^N} \ \mathcal{R}^\top z - \sum_{j \in A} f_j(s_j^\top z) \tag{P}$$

where, for all $j \in A$, $f_j(x) = K_j\sqrt{x}$ with $K_j = (c_j - v_j)\Phi^{-1}(\rho_j) + (e_j - v_j)L\left(\Phi^{-1}(\rho_j)\right)$ a nonnegative constant, where Φ denotes the c.d.f. of the standard normal distribution and L denotes the associated loss function. This model generalizes the basic selective newsvendor problem (SNP) as introduced by Taaffe et al. [15]. As noted earlier in this chapter, their problem is a special case of (P) where $a = 1$ and N is interpreted as a collection of markets that may or may not be entered by the supplier. Despite the fact that (P) is still a convex maximization problem for general values of a and therefore there exists an optimal extreme point (i.e., binary) solution to its continuous relaxation, this generalization makes the mathematical programming problem (P) considerably more challenging to solve, since a sorting approach can no longer be applied in general. Although in our application the functions f_j have the form given above, all of our results in fact apply more generally to the case where these functions are concave and nondecreasing. Moreover, without loss of generality we will assume that $\mathcal{R}_i > 0$ for all $i \in N$ (since it is easy to see that an optimal solution exists for which $z_i = 0$ for all $i \in N$ for which $\mathcal{R}_i \le 0$).

2.2.2 Examples

In this section we will discuss several multi-period and multi-item selective newsvendor problems that can be formulated as special cases of (P).

Multi-item selective newsvendor problems Of course, the generic description of our problem class can be viewed as a multi-item selective newsvendor problem where, by definition, we have to select a common set of customizations for the raw materials. However, note that if we can select a different set of customizations for each raw material then the problem decomposes by raw material. In other words, for each raw material we obtain an optimization problem that selects the set of customizations for that raw material. (We will refer to this case as C1.)

Multi-period selective newsvendor problems A generalization of the basic SNP described above to a multi-period setting considers the demands for a single product in a set of M different markets over a horizon of time periods indexed by the set T, denoted by the random variables D_{it} ($i \in M$, $t \in T$). We assume that all selection and ordering decisions have to be made at the start of the horizon. In this setting, we must decide (i) whether or not the set of selected markets can change between periods and (ii) whether or not inventory or backlogging is allowed between periods. In the remainder of this section, we will show how this problem reduces to one of the form (P) under a number of different simplifying assumptions.

With respect to the market selection, we will consider models that do not allow the set of selected markets to change between periods as well as models that do allow for costless changes in the set of selected markets. With respect to the inventory carryover and backlogging, we assume that these are either costless or not allowed.

T1. If inventory carryover and backlogging as well as changes in the market selection are allowed and costless, then demand can be pooled across all periods and all markets. In this case, we obtain an instance of (P) with $A = \{1\}$, $N = M \times T$, and $\mathbf{D}_{(i,t),1} = D_{it}$ ($i \in M$, $t \in T$), provided all demand variables are independent.

T2. If inventory carryover and backlogging are allowed and costless but the market selection is set across all periods, then the market demands can be pooled across all periods. In this case, we have $A = \{1\}$, $N = M$ and $\mathbf{D}_{i1} = \sum_{t \in T} D_{it}$ ($i \in M$) provided the aggregate market demands are independent.

T3. If inventory carryover and backlogging are not allowed but changes in the market selection are costless, then the problem decomposes by period. In this case, for each

$t \in T$, we solve a problem with $A = \{t\}$, $N = M$, and $\mathbf{D}_{it} = D_{it}$ $(i \in M)$ provided the market demands within a given period are independent.

T4. If neither carryover and backlogging nor changes in the market selection are allowed we have $A = T$, $N = M$, and $\mathbf{D}_{it} = D_{it}$ $(i \in M,\ t \in T)$ provided the market demands within a given period are independent.

2.3 Solution approaches

Several of the examples described in the previous section are of the same form as the basic SNP (i.e., have $a = 1$). This means that we can solve these problems using the exact algorithm provided in Taaffe et al. [15]. In Section 2.3.1 we will, for completeness' sake as well as insights into the implications for the examples, briefly review this solution approach.

Other interesting examples have $a > 1$ though, which yields a much harder optimization problem. In the remainder of this section, we will therefore develop an exact algorithm that runs in polynomial time in n but exponential time in a. Motivated by this analysis as well as the exact algorithm for the case $a = 1$ we also propose a class of heuristics for solving the problem.

2.3.1 Single raw material ($a = 1$)

When $a = 1$ and eliminating the corresponding index, the selection problem in profit maximization form is:

$$\max_{z \in \{0,1\}^n} \ \mathcal{R}^\top z - f(s^\top z).$$

Taaffe et al. [15] show that the optimal solution to this problem is either $z = \mathbf{0}$ (i.e., a vector all of whose elements are equal to zero) or of the form

$$z_i = \begin{cases} 1 & \text{if } i \preceq \ell \\ 0 & \text{if } i \succ \ell \end{cases}$$

for some $\ell \in N$, provided that the elements of N are sorted in such a way that

$$\frac{\mathcal{R}_i}{s_i} \geq \frac{\mathcal{R}_{i'}}{s_{i'}} \quad \Longleftrightarrow \quad i \preceq i'$$

for all $i, i' \in N$. Sorting the elements of n takes $O(n \log n)$ time, and the optimization problem can be solved by choosing the best among $n + 1$ candidate solutions.

For the examples with $a = 1$ discussed in the previous section (i.e., cases C1 and T1–T3) we then obtain the following sortings:

C1. For each raw material $j \in A$, we order the customizations $i \in N$ in non-increasing order of $\mathcal{R}_i/\sigma_{ij}^2$.

T1. We consider all possible (market, period)-pairs $(i, t) \in N \times T$ and sort these in non-increasing order of $\mathcal{R}_{it}/\sigma_{it}^2$ (where $\mathcal{R}_{it}$ is the net expected revenue for market i in period t).

T2. We aggregate the demands for each market $i \in N$ over the planning horizon and sort the markets in non-increasing order of $\mathcal{R}_i/\sigma_i^2$ (where $\mathcal{R}_i$ is the net expected revenue for market i over the entire planning horizon and σ_i^2 is the variance of the aggregate demand in market $i \in N$).

T3. For each period $t \in T$, we order the markets $i \in N$ in decreasing order of $\mathcal{R}_{it}/\sigma_{it}^2$.

2.3.2 General case

When there is more than one raw material ($a > 1$), (P) is no longer equivalent to the basic SNP studied in Taaffe et al. [15]. We therefore propose to solve this problem (or, in fact, a generalization thereof) by deriving the KKT conditions and using the resulting structure to find candidate solutions.

Consider the following linear relaxation of (P):

$$\max_{z \in [0,1]^n} \ \mathcal{R}^\top z - \sum_{j \in A} f_j \left(s_j^\top z \right). \tag{R}$$

Since the functions f_j ($j \in A$) are concave the objective function of (R) is convex, and this problem has an extreme point (i.e., binary) optimal solution. We will therefore focus on solving this continuous optimization problem. Under mild differentiability conditions (see, e.g., Bazaraa et al. [38]), the optimal solution to this problem can be found among the KKT solutions. Since the square root function, which is used in all of our applications, is not differentiable at 0 these conditions are violated at $z = \mathbf{0}$. However, this simply means that, in addition to the KKT solutions, we also need to consider $z = \mathbf{0}$ as a potential solution to the problem. Therefore, in the following analysis we will for convenience assume that the functions f_j are differentiable everywhere and that the KKT conditions are necessary for optimality.

The KKT conditions for this problem can be written as:

$$\begin{aligned} \mathcal{R}_i - \sum_{j \in A} f_j' \left(s_j^\top z \right) s_{ij} - \mu_i &= 0 && i \in N && (2.1) \\ \mu_i^- z_i &= 0 && i \in N \\ \mu_i^+ (1 - z_i) &= 0 && i \in N \end{aligned}$$

where $\mu_i^+ = \max\{\mu_i, 0\}$ and $\mu_i^- = \max\{-\mu_i, 0\}$. We will use these conditions to develop a collection of candidate solutions that is guaranteed to contain a (binary) optimal solution to (P), provided that $n \geq a$ (we will deal with the case $n < a$ later). Before we proceed with the analysis, let us introduce some notation that will be useful. First, let $\mathbb{S} = [s_{ij}]_{i \in N, j \in A}$. Furthermore, for $K \subseteq N$ let $\mathbb{S}_K = [s_{ij}]_{i \in K, j \in A}$ and $\mathcal{R}^K = (\mathcal{R}_i, i \in K)^\top$. In addition, let

$$\mathcal{K} = \{K \subseteq N : |K| = a \text{ and } \mathbb{S}_K \text{ is invertible}\}.$$

Finally, we use the 0-norm $\|\cdot\|_0$ to denote the number of nonzero elements in a vector.

Theorem 2.3.1. *If $n \geq a$, an optimal binary solution to (R) can be found by solving a problem of the form (R) for each $K \in \mathcal{K}$, where the problem corresponding to $K \in \mathcal{K}$ has $n - \|\mathbb{S}_{K^c}\mathbb{S}_K^{-1}\mathcal{R}^K - \mathcal{R}^{K^c}\|_0$ decision variables.*

Proof. The following class of linear programs (parameterized by values V_j , $j \in A$) is closely related to problem (R):

$$\text{maximize} \sum_{i \in N} \mathcal{R}_i z_i$$

subject to (LP)

$$\sum_{i \in N} s_{ij} z_i = V_j \qquad j \in A$$

$$0 \leq z_i \leq 1 \qquad i \in N.$$

In particular, an optimal solution to this problem is optimal to (R) if the values V_j are chosen equal to $f_j(s_j^\top z^*)$, where z^* is itself optimal solution to (R). (In fact, in many cases z^* will be the unique optimal solution to (LP).) We will use this linear program to derive structural properties of an optimal (binary) solution to (R). To this end, we formulate the dual of (LP):

$$\text{minimize} \sum_{j \in A} V_j \lambda_j + \sum_{i \in N} \mu_i^+$$

subject to

$$\begin{aligned}
\sum_{j\in A} s_{ij}\lambda_j + \mu_i^+ &\geq \mathcal{R}_i && i \in N \\
\mu_i^+ &\geq 0 && i \in N \\
\lambda_j \quad &\text{free} && j \in A.
\end{aligned}$$

The complementary slackness conditions for this pair of problems include:

$$\begin{aligned}
z_i\left(\sum_{j\in A} s_{ij}\lambda_j + \mu_i^+ - \mathcal{R}_i\right) &= 0 && i \in N \\
\mu_i^+(1 - z_i) &= 0 && i \in N.
\end{aligned}$$

Now note that, regardless of the values of V_j ($j \in A$), any basic feasible solution to (LP) can be characterized by exactly A basic variables, indexed by a set $K \in \mathcal{K}$. For a basic solution given by a set K to be optimal to (LP) it must, by duality theory for linear programming, satisfy the complementary slackness conditions with a dual solution that satisfies $\hat{\lambda}^K = \mathbb{S}_K^{-1}\mathcal{R}^K$. Using the complementary slackness conditions this, in turn, defines a partial solution to (LP):

$$z_i^K = \begin{cases} 0 & \text{if } \mathcal{R}_i < \sum_{j\in A} s_{ij}\hat{\lambda}_j^K \\ 1 & \text{if } \mathcal{R}_i > \sum_{j\in A} s_{ij}\hat{\lambda}_j^K. \end{cases}$$

It is easy to see that if our goal were to find *some* optimal solution to (P) we may restrict ourselves to solutions of this form and complete the solution by solving the following problem:

$$\max_{z_i\in\{0,1\},\, i\in K} \sum_{i\in K} \mathcal{R}_i z_i - \sum_{j\in A} f_j\left(\sum_{i\in N\setminus K} s_{ij} z_i^K + \sum_{i\in K} s_{ij} z_i\right). \tag{P^K}$$

However, note that we are not just interested in finding *any* optimal solution to (P), but in finding a *binary* optimal solution. Of course, if the optimal solution to (P) is unique there is no distinction. However, if (P) has multiple optimal solutions we have to make sure that we do not exclude all binary optimal solutions by restricting ourselves to solution of the form (2.3.2). Now note that what we will find is all optimal solutions to (P) that are *extreme point* optimal solutions of a corresponding (LP). Since no binary vector z can be a non-extreme point of the feasible region of (LP), we ensure that identifying all candidate optimal solutions corresponding to sets $K \in \mathcal{K}$ will contain a binary optimal solution to (P). $\square$

We will now assume that the following regularity condition holds, which will ensure that an optimal binary solution to (R) can be found in polynomial time in the number of customizations n.

Assumption 2.3.2. *Let* $\Phi = \left[\frac{s_{ij}}{\mathcal{R}_i}\right]_{i \in N, j \in A}$ *and* $\mathbf{1}$ *a vector all of whose elements are equal to 1. Then the unique solution to the system*

$$\Phi^\top x = \mathbf{0}$$

$$\mathbf{1}^\top x = 0$$

is $x = \mathbf{0}$.

The regularity condition captured by Assumption 2.3.2 is mild since, as we will show in the next lemma, it boils down to none of the extreme points of (LP) being degenerate.

Lemma 2.3.3. *If* $n \geq a$ *then Assumption 2.3.2 is equivalent to no subsystem of* $\mathbb{S}\lambda = \mathcal{R}$ strictly *containing a system of the form* $\mathbb{S}_K \lambda^K = \mathcal{R}^K$ *for any* $K \in \mathcal{K}$ *having a solution.*

Proof. First note that $\mathcal{K}$ is empty if $n < a$, so we will only address the case $n \geq a$.

"$\Rightarrow$" We will prove this by contradiction. Let $K \in \mathcal{K}$ and $\ell \in N \backslash K$ such that $s_{\ell \cdot}^\top \lambda^K = \mathcal{R}_\ell$

with $s_{\ell\cdot} = [s_{\ell j}]_{j \in A}$ and, as above, $\lambda^K = \mathbb{S}_K^{-1}\mathcal{R}^K$. Since $\mathcal{R} > 0$ this means that

$$\Phi_K \lambda^K = \mathbf{1}$$
$$\phi_{\ell\cdot}^\top \lambda^K = 1$$

where $\Phi_K = \left[\frac{s_{ij}}{\mathcal{R}_i}\right]_{i \in K, j \in A}$ and $\phi_{\ell\cdot} = \left[\frac{s_{\ell j}}{\mathcal{R}_\ell}\right]_{j \in A}$. Note that Φ_K is invertible, so that $\lambda^K = \Phi_K^{-1}\mathbf{1}$. Now let

$$\bar{x}_i = \begin{cases} \left((\Phi_K^\top)^{-1}\phi_{\ell\cdot}\right)_i & \text{if } i \in K \\ -1 & \text{if } i = \ell \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\Phi^\top \bar{x} = \Phi_K^\top \bar{x}^K + \phi_{\ell\cdot}\bar{x}_\ell = \Phi_K^\top(\Phi_K^\top)^{-1}\phi_{\ell\cdot} - \phi_{\ell\cdot} = 0$$

and

$$\mathbf{1}^\top \bar{x} = \mathbf{1}^\top(\Phi_K^\top)^{-1}\phi_{\ell\cdot} - 1 = \phi_{\ell\cdot}^\top \lambda^K - 1 = 0.$$

This is in contradiction with Assumption 2.3.2 so that we may conclude that no subsystem of $\mathbb{S}\lambda = \mathcal{R}$ *strictly* containing a system of the form $\mathbb{S}_K\lambda^K = \mathcal{R}^K$ for any $K \in \mathcal{K}$ has a solution.

"$\Leftarrow$" Again we will prove this by contradiction. Suppose that $\bar{x} \neq \mathbf{0}$ is a solution to the system in Assumption 2.3.2. This means that at least one element of $\bar{x}$ is nonzero; denote the index of such an element by $\ell \in N$. This implies that the system

$$\sum_{i \in N \setminus \{\ell\}} \frac{s_{ij}}{\mathcal{R}_i} \cdot x_i' = -\frac{s_{\ell j}}{\mathcal{R}_\ell} \qquad j \in A$$

has a solution (in particular, $x_i' = \bar{x}_i/\bar{x}_\ell$ for $i \in N\backslash\{\ell\}$ is a solution). This, in turn, implies that there exists a set $K \subseteq N\backslash\{\ell\}$ with $|K| = a$ for which the matrix $\left[\frac{s_{ij}}{\mathcal{R}_i}\right]_{i\in K, j\in A}$ is invertible. We have therefore identified a subsystem of $\mathbb{S}\lambda = \mathcal{R}$ *strictly* containing the system $\mathbb{S}_K\lambda^K = \mathcal{R}^K$ which has a solution. This proves the desired result.

□

Theorem 2.3.4. *Under Assumption 2.3.2, an optimal solution to (P) can be found by examining $O(n^a)$ solutions, each of which can be characterized in $O(a^3)$ time and evaluated in $O(na)$ time.*

Proof. Assumption 2.3.2 and Lemma 2.3.3 imply that, for all $K \in \mathcal{K}$, the problem (P^K) has a decision variables. Simply enumerating 2^a solutions yields that we can find an optimal binary solution to (R) by examining no more than $(2n)^a = O(n^a)$ solutions. For each $K \in \mathcal{K}$, the most time-consuming operation is inverting the matrix $\mathbb{S}_K$ which can be done in $O(a^3)$ time, and evaluating the objective function value of a solution can be done in $O(na)$ time. When $1 < n < a$, we can enumerate all $2^n \leq 2^a \leq n^a$ solutions to (P). Finally, when $n = 1$ there are only 2 solutions to consider. This yields the desired result. □

Exact algorithm for solving (P)

Step 0. Let $\pi^* = -\sum_{j\in A} f_j(0)$, $z^* = \mathbf{0}$, and choose $K \in \mathcal{K}$ arbitrarily.

Step 1. Let $\lambda^K = \mathbb{S}_K^{-1}\mathcal{R}^K$ and determine the partial solution z^K according to Eq. (2.3.2).

Step 2. Complete this solution by solving problem (P^K) and let

$$\pi^K = \mathcal{R}^\top z^K - \sum_{j\in A} f_j(s_j^\top z^K).$$

Step 3. If $\pi^* < \pi^K$ then let $z^* = z^K$ and $\pi^* = \pi^K$.

Step 4. Let $\mathcal{K} = \mathcal{K}\backslash\{K\}$. If $\mathcal{K} \neq \emptyset$, return to Step 1.

Upon termination of this algorithm, z^* is the optimal solution to (P) and π^* is the corresponding optimal value.

2.3.3 A class of heuristics

When a is large, the exact algorithm derived earlier in this section may no longer be computationally efficient. We therefore propose a class of heuristic algorithms that is inspired by both the KKT conditions as well as the sorting algorithm that provides the optimal solution in case $a = 1$. In particular, recall from Eq. (2.3.2) that a quantity of the form

$$\mathcal{R}_i - \sum_{j \in A} s_{ij} \lambda_j$$

can be used to, at least partially, determine the values of the selection variables. In our class of heuristics we will use quantities of this form to measure the attractiveness of selecting the customizations $i \in N$, where the vector λ parameterizes the heuristic. Then, given an instance of the heuristic, we will consider the customizations $i \in N$ sorted in non-increasing order of attractiveness, and select the first ℓ from this sorted list (for some $\ell \in \{1, \ldots, n\}$. Note that not selecting any customization is always feasible, so we will also consider this obvious solution to the problem as a candidate. Of course, there are many ways in which the vector λ can be chosen. Inspired by KKT condition (2.1) and given the fact that an optimal solution to (P) is also a KKT solution to (R), we will choose parameters of the form

$$\lambda_j = f_j'(s_j^\top z) \qquad\qquad j \in A$$

for some $z \in \{0, 1\}^n$. In Section 2.4 we will consider several possibilities for choosing one or more values of z when applying the heuristic to a particular problem instance of (P). All of these will take the form of selecting z from some probability distribution on $\{0, 1\}^n$.

More formally, the heuristic can be described as follows:

Heuristic for solving (P)

Step 0. Let $\hat{\pi} = -\sum_{j \in A} f_j(0)$ and $\hat{z} = \mathbf{0}$.

Step 1. Select a vector $z \in \{0,1\}^n$ according to some probability distribution.

Step 2. Let

$$\begin{aligned} \lambda_j &= f_j'(s_j^\top z) && j \in A \\ \mu_i &= \mathcal{R}_i - \sum_{j \in A} s_{ij} \lambda_j && i \in N. \end{aligned}$$

Step 3. Define the following ordering of the elements of N: $\mu_i \geq \mu_{i'} \Longleftrightarrow i \preceq i'$ for all $i, i' \in N$.

Step 4. Let $L \subseteq N$ and determine partial solutions given by

$$z_i^{K,\ell} = \begin{cases} 1 & \text{if } i \preceq \ell \\ 0 & \text{if } i \succ \ell \end{cases}$$

with corresponding objective function values $\tilde{\pi}^{K,\ell} = \mathcal{R}^\top z^{K,\ell} - \sum_{j \in A} f_j(s_j^\top z^{K,\ell})$, for all $\ell \in L$.

Step 5. Let $\ell^* = \arg\max_{\ell \in L} \pi^{K,\ell}$, $z^K = z^{K,\ell^*}$, and $\pi^K = \pi^{K,\ell^*}$.

Step 6. If $\hat{\pi} < \pi^K$ then let $\hat{z} = z^K$, $\hat{\pi} = \pi^K$, and (if additional candidate solutions are desired) set $z = \hat{z}$ and return to Step 2.

Upon termination of the heuristic, $\hat{z}$ is the best solution found and $\hat{\pi}$ is the corresponding objective function value.

Note that Step 4 in the heuristic is motivated by a combination of the exact algorithm for $a = 1$ discussed in Section 2.3.1 and the exact algorithm for the general case discussed in Section 2.3.2. In Section 2.4 we will show how different choices in Steps 4 and 6 impact the efficiency and efficacy of the heuristic, and also explore multistart implementations of the heuristic.

2.4 Computational results

In Sections 2.3.2 and 2.3.3 we developed an exact algorithm as well as a class of heuristic algorithms for solving (P) when $a > 1$. In this section, we will study and evaluate the performance of the exact algorithm as well as several variants of the class of heuristics. All tests were performed on a PC with an Intel Xeon Quad Core 3.2 GHz processor with 8 GB RAM.

2.4.1 Test problem instances

We created problem instances by randomly generating problem data using uniform distributions for the problem parameters as given in Table 2.1. However, it turns out that using this data the distribution of the number of customizations selected in the optimal solution is skewed towards larger numbers, while on the other hand a sizable number of problem instances have optimal solution **0**. Since the performance of the exact algorithm is insensitive to the nature of the optimal solution and our heuristics always consider the solution **0**, we decided to

(i) eliminate all problem instances with optimal solution **0**;

(ii) use an acceptance/rejection method to ensure that we have an equal number of problem instances with $1, 2, \ldots, n$ customizations selected in the optimal solution.

For larger problem instances applying (ii) exactly turned out to be impractical. However, as we show in the remainder of this section, one of our heuristics performs very well.

Therefore, for larger instances we use the solution found by this heuristic in the acceptance/rejection method in (ii).

K_j	μ_{ij}	σ_{ij}	$r_{ij} - c_j$	F_i
$U(0,6)$	$U(0, \frac{100}{n})$	$U(0,5)$	$U(0, \frac{8}{a})$	0

Table 2.1: Test problem parameters

2.4.2 Performance evaluation

2.4.2.1 Exact algorithm vs. complete enumeration

In Section 2.3.2 we developed an exact algorithm that, under a mild regularity condition, runs in polynomial time in the number of customizations but in exponential time in the number of raw materials. It is also easy to see that complete enumeration of all solutions would run in exponential time in the number of customizations but in polynomial time in the number of raw materials. In this section, we will compare the performance of both on a set of test instances, with the goal of providing some insight into the problem dimensions that make the exact algorithm attractive as compared to complete enumeration. In particular, in Figure 2.1 we show, using a log-scale, the running time for both algorithms as a function of both n and a. As expected, the threshold for the number of customizations at which the exact algorithm starts to outperform complete enumeration increases in the number of raw materials. However, for the class of instances that we used, this threshold is reasonably small, highlighting the value of the exact algorithm for a large class of problem instances. Using similar data for $a = 4, 6, 8$ we showed empirically that the running time of our exact algorithm and complete enumeration satisfy the following relationships very closely:

$$t^{\text{exact}}(n, a) = 2^{-17.71} \times \binom{n}{a} a^4$$

$$t^{\text{enum}}(n, a) = 2^{-13.66} \times 2^n$$

which can help determine the actual values of n and a for which the exact algorithm outperforms complete enumeration.

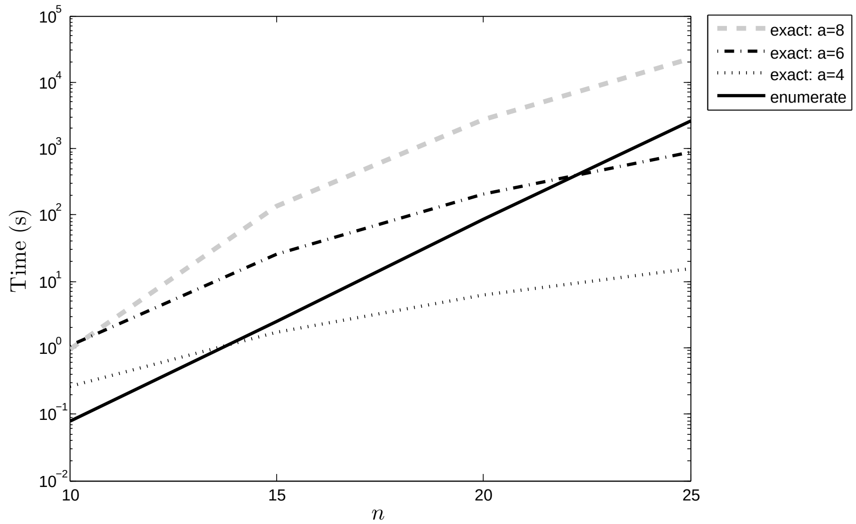


Figure 2.1: Run time as a function of n and a for exact algorithm and enumeration.

In addition to the relative performance, it is important to observe from Figure 2.1 that both algorithms quickly become impractical for large values of n, while such larger values could easily occur in practice. For example, a typical Walmart distribution center serves about $n = 75$–100 stores, and we might consider a small set of products that must be sold together such as an electronic good and related accessories for a total of $a = 2$–10 products. Both the exact algorithm and enumeration take approximately 200 seconds when n is as little as 22 and $a = 6$. This motivates the use of heuristics for many practical-sized problems.

2.4.2.2 Heuristic performance

In Section 2.3.3 we proposed a class of heuristics for solving (P). We will start by comparing the efficiency and efficacy of several variants to publicly available solvers in MATLAB. Since the SNP is a convex maximization problem any local maximum will occur at an extreme (i.e., in our case binary) point of the feasible region we chose to use general nonlinear optimization solvers. In particular, we used both BONMIN and LGO, which do not guarantee optimality and hence qualify as heuristics. LGO has several solver options, and we provide results for both the pure local search setting (single run from a given starting point, denoted by LGO-1) and the global multistart random search and local search setting (denoted by LGO-∞). The variants of our heuristic that we consider differ with respect to (i) the choice of the set L, and (ii) the (maximum) number of additional candidate solutions that are generated via Step 6.

(i) We consider two choices of L, both of which depend on the current solution. The first choice (indicated by *KKT*) is motivated by the exact algorithm from Section 2.3.2 (or the KKT conditions), and sets $L = \{\ell\}$ where ℓ is the largest element of N with respect to the ordering defined in Step 3 for which $\mu_i > 0$. The second choice (indicated by *ranking*) is motivated by the sorting algorithm that provides the optimal solution when $a = 1$ and sets $L = N$.

(ii) We consider two choices in Step 6. The first choice (indicated by *1*) simply considers a single (inner) iteration of the algorithm, while the second choice (indicated by ∞) does not set an upper bound on the number of iterations (and hence terminates the algorithm if no improvement is found).

Tables 2.2 and 2.3 summarize the performance of the heuristics on a set of smaller problem instances. For each problem dimension, we generated either 1000 or 10,000 instances as outlined in Section 2.4.1 and computed the optimal solution using our exact algorithm. As can be seen from the tables, our custom heuristics perform very well overall. From a randomly generated starting point, the KKT-based variants find the optimal solution to about 45–75% of the instances in negligible time. In somewhat more (but still negligible) time, the ranking based heuristics find the optimal solution to almost all (98–100%) of the instances. The ratio of the time required by BONMIN to that required by the ranking-∞ heuristic is substantial, while the former is much less effective. LGO-1 took somewhat

		time (10^{-3} sec)							optimum found (%)						
		KKT		ranking		BONMIN	LGO		KKT		ranking		BONMIN	LGO	
n	a	1	∞	1	∞		1	∞	1	∞	1	∞		1	∞
10	4	0.04	0.08	0.20	0.46	14.22	17.55	1034	57	75	99	100	78	77	94
	6	0.04	0.08	0.23	0.51	14.09	17.69	1033	54	73	99	100	78	75	94
	8	0.04	0.08	0.25	0.56	14.11	17.61	1033	53	73	99	100	77	75	94
	10	0.05	0.10	0.24	0.53	13.57	18.83	1103	54	73	99	100	78	76	96
	15	0.05	0.09	0.29	0.65	13.71	18.99	1107	56	74	99	100	79	76	96
	20	0.06	0.10	0.35	0.77	13.81	19.21	1108	52	72	100	100	79	74	96
	25	0.06	0.10	0.41	0.90	14.08	19.61	1117	52	73	100	100	78	76	95
15	4	0.04	0.08	0.28	0.62	17.98	30.25	2513	46	69	99	100	72	73	91
	6	0.04	0.09	0.32	0.70	17.59	30.24	2526	44	68	99	100	73	73	91
	8	0.04	0.09	0.35	0.76	17.57	30.22	2519	44	69	99	100	73	73	91
	10	0.04	0.09	0.33	0.71	16.95	32.38	2692	46	70	99	100	73	73	91
	15	0.05	0.09	0.42	0.89	17.28	32.22	2711	46	71	99	100	76	76	93
	20	0.06	0.10	0.50	1.07	17.48	32.18	2708	48	71	100	100	75	73	93
	25	0.07	0.11	0.61	1.28	18.07	33.30	2821	45	71	100	100	74	74	97

Table 2.2: Algorithms average run times (10,000 instances for $a \leq 8$, 1,000 instances for $a \geq 10$).

		error (%) in profit, all							error (%) in profit, non-optimal						
		KKT		ranking		BONMIN	LGO		KKT		ranking		BONMIN	LGO	
n	a	1	∞	1	∞		1	∞	1	∞	1	∞		1	∞
10	4	8.56	5.11	0.05	0.01	4.60	4.05	1.34	20.12	20.35	5.33	28.31	20.45	17.90	23.03
	6	14.29	8.98	0.08	0.03	7.77	7.54	2.30	31.12	33.40	9.43	64.24	34.61	30.16	40.44
	8	17.98	11.53	0.08	0.01	10.14	10.08	3.21	38.42	42.20	11.04	100	44.50	40.07	54.43
	10	18.24	11.50	0.11	0.10	10.28	9.80	2.71	39.40	42.74	17.72	100	45.68	41.00	61.52
	15	17.82	11.43	0.18	—	10.02	10.62	2.36	40.50	44.13	29.17	—	46.82	43.36	54.87
	20	19.94	11.85	0.06	—	9.28	10.47	1.83	41.72	42.48	21.37	—	43.37	39.50	46.90
	25	17.80	9.95	0.09	—	8.36	8.84	2.12	37.46	36.57	30.51	—	38.19	36.07	43.19
15	4	12.89	8.24	0.08	0.01	7.55	6.56	3.26	23.79	26.41	6.25	46.83	27.27	24.52	37.96
	6	20.24	12.98	0.07	0.03	11.74	10.49	5.16	35.91	40.75	6.92	60.04	43.19	38.23	59.53
	8	22.11	14.10	0.08	0.04	12.76	11.32	5.74	39.43	44.78	10.93	100	46.86	41.64	67.00
	10	21.02	13.11	0.07	—	12.46	11.51	5.53	38.57	43.56	13.04	—	46.50	42.15	64.31
	15	21.65	12.52	0.00	—	10.89	9.48	4.50	40.09	42.88	0.57	—	44.99	39.16	64.34
	20	18.86	11.64	0.00	—	10.83	9.87	4.53	36.47	40.42	2.18	—	42.81	36.67	62.91
	25	18.88	10.26	0.00	—	8.50	7.72	1.38	34.58	35.38	0.41	—	32.68	29.35	41.76

Table 2.3: Heuristic relative error (among all and only non-optimally solved instances).

more time than BONMIN, with no clear difference in quality of solutions. Lastly, LGO-∞ took significantly more time than any other heuristic, and still found the optimal solution significantly less often than the ranking heuristics (91–97%). In Table 2.3 we list the error for all heuristics, both across all instances and only across the instances in which the heuristic was not able to find the optimal solution. The latter errors are sometimes substantial, but for the most successful heuristic the frequency of such errors is extremely small.

In Figure 2.2 and Table 2.4 we analyze a more comprehensive set of results for $n = 15$ and $a = 6$. In Figure 2.2 we show the empirical pdf of the relative error in the solution

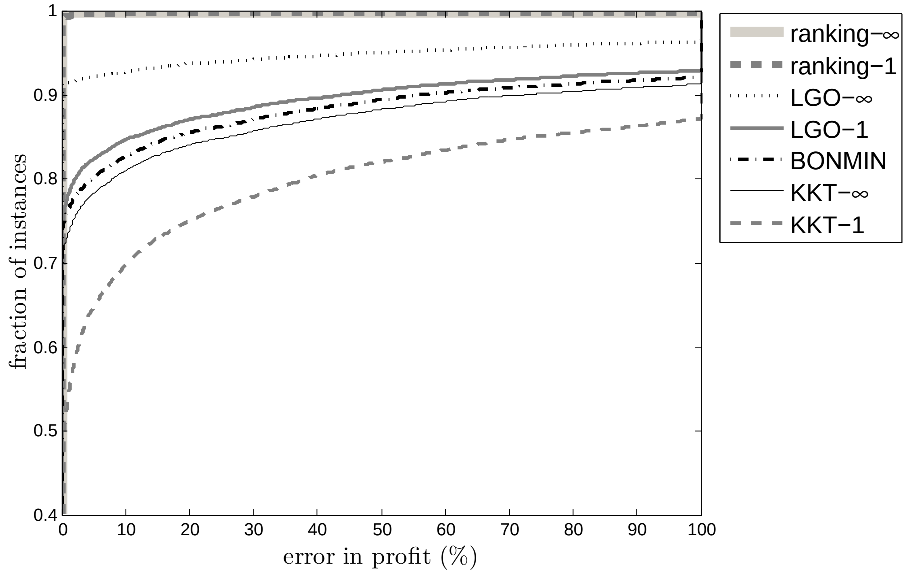


Figure 2.2: Comparison of performance of heuristics and BONMIN.

Class #	KKT 1	KKT ∞	ranking 1	ranking ∞	BONMIN	LGO 1	LGO ∞
	% of optimal solutions found						
1	0.30	4.20	99.85	99.85	4.95	13.49	23.99
2	0.91	11.78	99.70	99.85	14.05	29.00	60.57
3	3.75	25.94	99.25	100	27.74	43.18	87.41
4	15.14	40.48	99.10	99.85	42.88	60.57	98.20
5	29.84	58.92	99.40	100	62.97	70.16	99.55
6	48.43	71.51	99.10	100	74.06	72.26	100
7	62.82	79.76	99.10	100	83.66	76.91	100
8	69.12	85.91	98.65	100	92.80	80.81	100
9	71.51	88.16	98.50	99.85	97.30	83.21	100
10	66.87	91.15	98.05	99.85	98.65	86.36	100
11	59.22	90.55	98.80	100	97.90	90.40	100
12	46.78	87.41	96.85	100	98.20	90.70	100
13	51.12	91.90	98.95	100	97.90	94.30	100
14	58.92	96.25	99.70	100	99.10	98.05	100
15	69.42	97.75	100	100	99.70	98.80	100

Table 2.4: Success rate for $n = 15, a = 6$ as a function of $\sum_{i=1}^{n} z_i$.

# starts	time (10^{-3} sec)						optimum found (%)					
	KKT		ranking		Solvers		KKT		ranking		Solvers	
	1	∞	1	∞	BONMIN	LGO-1	1	∞	1	∞	BONMIN	LGO-1
1	0.04	0.09	0.32	0.70	17.28	30.03	43.80	68.44	98.90	99.96	72.93	72.59
2	0.09	0.17	0.65	1.43	34.54	60.09	55.48	72.54	99.63	99.98	73.60	80.67
5	0.22	0.43	1.61	3.62	86.24	150.01	67.64	77.85	99.90	99.99	74.39	89.57
10	0.43	0.87	3.22	7.26	172.28	299.27	74.68	81.48	99.94	100	74.94	93.69
20	0.87	1.73	6.44	14.54	343.62	597.00	79.91	84.66	99.96	100	75.54	96.32
30	1.30	2.59	9.66	21.83	514.97	894.14	82.46	86.56	99.97	100	75.99	97.39

Table 2.5: Algorithm running time and success rate for $n = 15$, $a = 6$.

# starts	error (%) in profit, all						error (%) in profit, non-optimal					
	KKT		ranking		Solvers		KKT		ranking		Solvers	
	1	∞	1	∞	BONMIN	LGO-1	1	∞	1	∞	BONMIN	LGO-1
1	20.35	12.90	0.09	0.04	11.83	10.55	36.22	40.88	7.77	100	43.71	38.48
2	16.97	11.95	0.04	0.02	11.66	7.97	38.11	43.51	11.10	100	44.17	41.22
5	13.88	10.42	0.02	0.01	11.46	5.14	42.88	47.02	23.03	100	44.76	49.25
10	11.95	9.32	0.01	—	11.28	3.49	47.19	50.33	20.36	—	45.02	55.38
20	10.31	8.11	0.01	—	11.13	2.21	51.33	52.86	24.28	—	45.50	60.10
30	9.30	7.33	0.01	—	11.00	1.69	52.99	54.52	32.36	—	45.80	64.73

Table 2.6: Average relative error per instance for $n = 15$ and $a = 6$.

given by each algorithm. Table 2.4 shows how the various algorithms perform as a function of the number of markets selected in the optimal solution. We see that both the KKT heuristics as well as all of the publicly available solvers perform significantly worse when few markets are selected in the optimal solution, while the ranking heuristics perform very well regardless of the nature of the optimal solution.

In Tables 2.5 and 2.6 the performance of the algorithms is compared for $n = 15$ and $a = 6$ when multiple randomly generated initial vectors are used. In these tables we do not consider the LGO-∞ algorithm since it includes a global search phase as part of the algorithm. We used a nested structure to generate these instances. The results show that using multiple random starting vectors can be an effective way to improve the performance of the heuristics. In addition, we observe that using LGO-1 with multiple random starting solutions is more effective than the global search option for our problem.

Table 2.7 compares the performance of the ranking-∞ algorithm and the multistart BONMIN and 1-LGO algorithms on large problem instances (since the acceptance/rejection method described in remark (ii) in Section 4.1 we slightly modified the model for generating problem instances by choosing $F_i = F \in U(0,a)$ for $i = 1, \ldots, n$.). In this case, we use the solution of the 50-start ranking-∞ algorithm to approximate the optimal solution and solve a total of 100 instances for each problem dimension (in no case did the publicly available solvers identify a solution better than the 50-start ranking-∞ algorithm). As can be seen, the time required to get good solutions with the ranking-∞ algorithm continues to perform very well. In addition, the solution times for the commercial solvers begin to be quite significant. Note also that the solution time required by both our heuristics and the publicly available solvers appear to be relatively insensitive to the value of a. Therefore, the results in Table 2.7 are expected to hold for a wide range of a values.

			time (s)			approximate optimum found (%)		
n	a	# of starts	ranking-∞	BONMIN	LGO-1	ranking-∞	BONMIN	LGO-1
25	4	1	0.001	0.026	0.071	100	78	79
		25	0.020	0.649	1.760	100	82	96
		50	0.039	1.294	3.537	100	82	97
	8	1	0.001	0.025	0.073	100	85	83
		25	0.025	0.629	1.769	100	87	99
		50	0.051	1.258	3.522	100	88	99
50	4	1	0.001	0.059	0.239	100	83	80
		25	0.036	1.542	5.987	100	84	96
		50	0.072	3.171	11.957	100	84	98
	8	1	0.002	0.068	0.241	100	95	87
		25	0.047	1.676	6.014	100	95	100
		50	0.095	3.397	12.067	100	95	100
100	4	1	0.003	0.419	0.891	100	94	93
		25	0.072	9.126	22.261	100	95	100
		50	0.145	18.515	44.421	100	95	100
	8	1	0.004	0.294	0.916	100	98	93
		25	0.093	6.392	22.915	100	98	100
		50	0.186	12.870	45.736	100	98	100
150	4	1	0.004	0.893	1.976	100	97	91
		25	0.114	22.127	49.685	100	98	100
		50	0.227	42.651	99.423	100	98	100
	8	1	0.006	0.851	2.066	100	96	91
		25	0.145	24.348	50.925	100	98	100
		50	0.289	49.154	101.617	100	98	100
200	4	1	0.006	1.682	3.612	100	96	94
		25	0.157	45.268	90.015	100	99	100
		50	0.314	92.276	179.830	100	99	100
	8	1	0.008	1.863	3.669	100	96	95
		25	0.200	44.826	90.534	100	98	100
		50	0.399	88.193	181.232	100	98	100

Table 2.7: Results for large problem instances.

Finally, Figure 2.3 compares the computation times of the heuristics, the exact algorithm, and complete enumeration as a function of n for $a = 6$. It clearly follows that our heuristics significantly outperform the publicly available solvers as well as exact approaches.

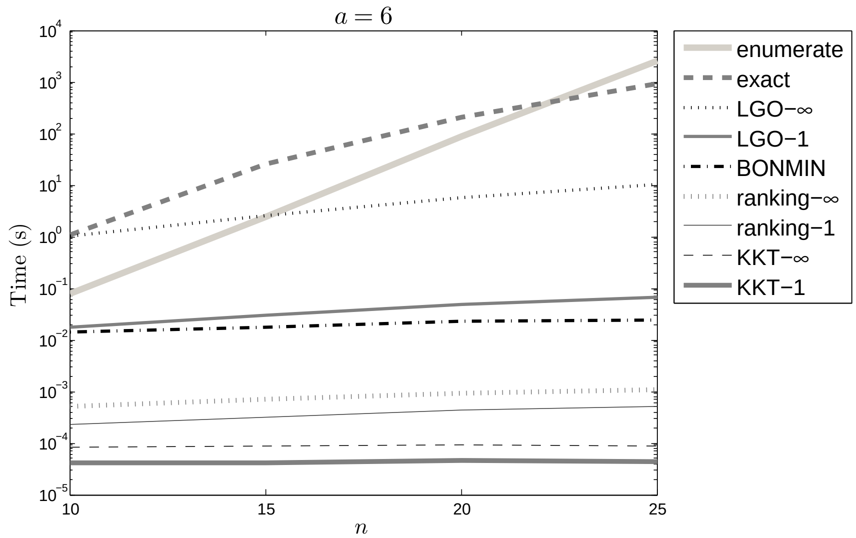


Figure 2.3: Running time as a function of m for $a = 6$.

2.5 Concluding remarks

In this chapter we have developed tools to solve a class of selective newsvendor problems with independent and normally distributed demands. These results can be used to solve certain multi-product, multi-period, or other selection problems. We have shown that some problems in our class can be solved efficiently and exactly using the sorting algorithm by Taaffe et al. [15]. In addition, we have developed an exact algorithm which is efficient in the number of items as well as a class of heuristics. We compared the effectiveness of, in particular, our heuristics with publicly available solvers, demonstrating that a particular variant of our heuristic represents a significant improvement over using other solvers, both in terms of computation time and solution quality. One of the key limitations of our current research is that we can only handle independent and normally distributed demands, and this issue will be addressed in Chapter 3. In addition, we hope that some future work may be able to address the limitation that in the multi-period interpretation of our model we do not allow for nonzero and finite inventory carryover and backlogging costs or changes to the market selection over time.

CHAPTER 3

Approximation Algorithms for Stochastic Selection Problems

3.1 Introduction

In Chapter 2 we considered a generalization of the original selective newsvendor problem (see Taaffe et al. [15]) to a class which included the multi-period market selection problem as long as the demands in each market were independent and normally distributed. In this chapter we stick with the single-period setting, but allow the demand vector to come from a nonnegative joint distribution, including dependence between market demands. The model presented in this chapter is in fact much more general than the SNP. Our class of 2-stage stochastic selection problems includes both a (binary) selection vector as well as other associated (continuous) decisions. The objective function includes rewards based on the subset chosen and costs based on all decisions. Both the rewards and the costs may be stochastic, and so we consider a risk-based objective function. We assume the cost function is convex, and also that the objective function has certain scalability properties. This class of problems is in general hard to solve in part because of the integer selection choices as well as the stochasticity. Because of these challenges, we combine a rounding approach and sample average approximation to efficiently find high-quality solutions with high probability.

In our problem, we study general stochastic distributions where even computing the objective value for a decision vector may be computationally intensive. Even if this is not the case, our general class of problems is $\mathcal{NP}$-hard: Chen and Zhang [36] show that testing membership of the core for the newsvendor game is $\mathcal{NP}$-hard, and this problem can be shown to be equivalent to solving an instance of the single-period, which is a special case of our model, to optimality. Because of the difficulties inherent in our class of problems, we seek to develop efficiently implementable approximation algorithms. In the spirit of [13] we consider algorithms that are based on the idea of rounding the solution to the continuous relaxation of the problem. However, even the continuous relaxation is often difficult to solve, leading to a final approximation algorithm that uses a combination of sample average approximation (SAA) and rounding. Using these techniques, we are able to develop several classes of approximation algorithms with explicit (albeit probabilistic) performance guarantees.

To use the approach we present here, the optimization problem must have a few key features:

- A reward-cost selection structure such that (binary) selected options provide a reward and the decision vector results in stochastic costs.
- Scalability of the objective function so that a continuous decision vector can be rounded to a feasible binary selection. The other decisions may need to be carefully chosen in order to ensure the overall cost is "not too bad" by comparison to the cost of the continuous decision.
- Complete recourse for the second-stage problem.

These features allow our rounding algorithm combined with SAA to behave well and produce decision vectors that, with high probability, result in high-quality solutions to the optimization problem. However, some of these assumptions could be relaxed if we were not concerned with providing a performance guarantee.

Our model has a few key features which distinguish it from most of the literature on similar problems. In contrast with earlier related models, our setting allows for state-based dependence among the random variables. We also assumes a rather general cost structure which encompasses a variety of applications in supply chain and other areas. In addition, many supply chain and resource allocation optimization problems under uncertainty use an expected value based objective function. However, we can accommodate a very general class of risk-based objectives (where an expected value objective is a special case).

The major contributions of this chapter are primarily methodological as we develop approximation algorithms which combine rounding and SAA to provide explicit performance guarantees. We also provide a framework which identifies which features of the objective function allow our approximation algorithms to be effective. Our key results include the development of three classes of approximation algorithms, each with three variants. Specifically, we develop algorithms which solve the selection problem either by simply solving the linear relaxation and rounding (possibly in a strategic way), or by directly solving the integer problem using SAA. Using these algorithms we can then choose the other decision variables at the same time as the selection, using a second instance of the SAA problem, or by using true information. We provide performance guarantees for each algorithm, and also empirically demonstrate their performance for a supply chain example.

Past work considers several approaches to modeling and solving deterministic market selection problems. Geunes et al. [35] and Geunes et al. [11] consider a market selection version of the economic order quantity problem and capacitated extensions thereof. Van den Heuvel et al. [12] study a market selection problem with economic lot-sizing costs. They show that the problem is $\mathcal{NP}$-hard and, in addition, that a profit-maximization formulation of the problem cannot efficiently be approximated to within a constant factor. They also study polynomially solvable special cases and propose heuristics. Geunes et al. [13] consider a class of deterministic selection problems that contains the market selection problem with lot-sizing costs as a special case, and derive conditions under which a polynomial-time constant-factor approximation algorithm exists for a cost-minimization version of the problem. An alternative generalization of deterministic market selection problems is presented in [39]. In that work, the authors consider a multi-echelon market selection problem with inventory held at both the distribution center (DC) level, as well as at the retailers. They then develop an efficient solution approach for their profit-maximizing problem.

Market selection problems with stochastic demand have mostly focused on settings with newsvendor-type costs, independent and normally distributed market demands, and linear revenue functions. Our model builds on previous work in the area of market selection in general, and the abovementioned SNP in particular. Taaffe et al. [15] introduced the SNP, and other variants and extensions were studied by Carr and Lovejoy [14] and Taaffe et al. [34], and in Chapter 2 of this thesis. Alptekinoğlu and Tang [40] also considered normally distributed demands and studied how demand in a collection of markets could be met using multiple channels. Finally, Lin and Ng [41] consider a more general demand distribution from a bounded interval, and then consider a minimax regret objective.

The remainder of this chapter is organized as follows. Section 3.2 provides our problem framework and key assumptions, as well as examples of cost functions and risk measures we consider. In Section 3.3 we present our core approximation result using a rounding scheme for a class of problems, and show that two supply-chain examples fall into that problem class. We then introduce using sample average approximation in Section 3.4, and provide bounds on the number of samples needed. In Section 3.5 we provide our computational experiments, and demonstrate excellent performance of our approximation algorithms for our problem. Finally, in Section 3.6 we conclude.

3.2 General framework and preliminaries

3.2.1 Problem formulation and assumptions

Consider a stochastic selection problem with m choices. Here we repeat the use of z as a selection decision vector where $z \in \{0,1\}^m$ and $z_i = 1$ ($z_i = 0$) means that choice i is (not) selected ($i = 1, \ldots, m$). Similarly, we generalize the procurement decisions Q from Chapter 2 to a continuous η-dimensional decision vector $y \in \mathcal{Y} \subseteq \mathbb{R}^\eta$ which includes any continuous decisions related to the selection. Then denote the real-valued total (random) cost associated with decisions (z, y) by $\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D})$, where $\mathbf{R}$ and $\mathbf{F}$ are the random reward and fixed costs similar to the deterministic $\mathcal{R}$ and F in Chapter 2 and $\mathbf{D}$ remains the random demands (or, more generally, requirements). The random vector $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ is assumed to have a joint probability distribution on $\mathbb{R}^{3m}$, and the random cost variable $\Gamma(\cdot)$ is viewed as an element of a linear space $\mathcal{X}$ of measurable functions, defined on an appropriate sample space. We assume that $\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D})$ is defined for continuous "selection" vectors $z \in [0,1]^m$, although we are ultimately only interested in binary selection vectors. Finally, we let ρ denote a *risk measure* that assigns a real value to a random cost variable, and consider the following class of selection problems:

$$\underset{z \in \{0,1\}^m, y \in \mathcal{Y}}{\text{minimize}} \;\; \Psi(z, y) \equiv \rho\left[\Gamma\left(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}\right)\right]. \tag{P}$$

Denote the continuous relaxation of (P) (obtained by relaxing the binary constraints to $z \in [0,1]^m$) by (R). We next provide our main assumptions on the problem inputs. Note, however, that some of our results hold under weaker assumptions, and we will specify that where appropriate.

First, we assume that the risk measure ρ is *coherent* (see, e.g., [42]):

Assumption 3.2.1. *Coherence of ρ*

(i) Convexity*: If $X_1, X_2 \in \mathcal{X}$ and $\alpha \in [0,1]$ then $\rho[\alpha X_1 + (1-\alpha)X_2] \leq \alpha\rho[X_1] + (1-$*

$\alpha)\rho[X_2]$.

(ii) Monotonicity*: If $X_1, X_2 \in \mathcal{X}$ and $X_2 \succeq X_1$ then $\rho[X_2] \geq \rho[X_1]$.*

(iii) Translation Equivariance*: If $X \in \mathcal{X}$ and $a \in \mathbb{R}$ then $\rho[X + a] = \rho[X] + a$.*

(iv) Positive Homogeneity*: If $X \in \mathcal{X}$ and $\lambda > 0$ then $\rho[\lambda X] = \lambda \rho[X]$.*

Here $X_2 \succeq X_1$ means that a realization of X_1 is no larger than the realization of X_2 corresponding to the same outcome of the underlying random experiment; i.e., in our case for the same realization r, f, d of rewards, fixed selection costs, and demands. Next, we impose some regularity assumptions on the random rewards, costs, and demands:

Assumption 3.2.2. *Regularity of distributions*

(i) $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ has a nonnegative support.

(ii) The variance-covariance matrix of $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ has finite elements.

Finally, we make some convexity and attainment assumptions:

Assumption 3.2.3. *Convexity and attainment*

(i) $\mathcal{Y}$ is a convex cone.

(ii) For all $r, f, d \in \mathbb{R}_+^m$, $\Gamma(\cdot, \cdot; r, f, d)$ is convex on $[0, 1]^m \times \mathcal{Y}$.

(iii) The optimal solution value to problem (R) is finite and it is attained.

For convenience, we define the function

$$\Psi^*(z) = \min_{y \in \mathcal{Y}} \Psi(z, y) \qquad \text{for } z \in [0, 1]^m$$

as well as the scalars

$$\Psi^* = \min_{z \in \{0,1\}^m} \Psi^*(z) = \min_{z \in \{0,1\}^m, y \in \mathcal{Y}} \Psi(z, y)$$

$$\underline{\Psi} = \min_{z \in [0,1]^m} \Psi^*(z) = \min_{z \in [0,1]^m, y \in \mathcal{Y}} \Psi(z, y)$$

(where Assumption 3.2.3 is sufficient to ensure that all minima are indeed attained).

3.2.2 Examples: Cost functions

In this section we will describe two illustrative examples of cost functions Γ. We will use these examples throughout the chapter to show how the general results can be applied. The first example generalizes the basic SNP, where the supplier faces the aggregate demands from the selected markets, while the second is a two-stage problem where market order quantities are determined early but transshipments can be employed after realization of demands to redistribute inventories.

3.2.2.1 Example 1: Aggregate newsvendor costs

Suppose that a decision maker faces aggregate newsvendor costs given a collection of candidate market demands to serve. The function Γ then consists of two parts; the first part represents lost revenues in the markets that are not selected while the second part represents procurement and newsvendor costs. Note that the problem can be formulated in profit maximization form as we did in Chapter 2. However, motivated by the negative result that the optimal profit in the deterministic market selection problem with economic lot-sizing costs cannot be efficiently approximated (see [12]) we choose a cost minimization formulation (see also [13]) that, as we will see, does lend itself to the development of approximation algorithms.

Recall that c, v, and e are the unit procurement cost, salvage value, and expediting cost, and let y be the aggregate order quantity (with $y \in \mathcal{Y} = \mathbb{R}_+$ so that $\eta = 1$). Then define

$$\Gamma\left(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}\right) = \Gamma_1\left(z; \mathbf{R}\right) + \Gamma_2\left(z, y; \mathbf{F}, \mathbf{D}\right)$$

where $\Gamma_1(z;\mathbf{R}) = \mathbf{R}^\top(\mathbf{1}-z)$ and $\Gamma_2(z,y;\mathbf{F},\mathbf{D}) = \mathbf{F}^\top z + cy + e\left(\mathbf{D}^\top z - y\right)^+ - v\left(y-\mathbf{D}^\top z\right)^+$. We assume that $0 < v < c < e$, which ensures convexity of Γ, existence of an optimal solution to (P), and monotonicity of Γ_2 in z (which will prove useful later on).

3.2.2.2 Example 2: Market newsvendor costs with transshipment recourse

Now suppose that a decision maker faces newsvendor costs in each (selected) market. An order quantity needs to be determined for each selected market before demand realizes, but transshipment between the markets can be used to reallocate goods. The function Γ again consists of two parts; the first part represents lost revenues in the markets that are not selected while the second part represents procurement, newsvendor, and transshipment costs.

As before we let c_i, v_i, and e_i be unit procurement cost, salvage value, and expediting cost in market i and s_{ij} the unit transportation cost between market i and j after demand is realized ($i,j = 1,\ldots,m$). In the first stage of the problem we select the markets to serve as well as the order quantity for each of the markets. We denote the latter by $y = (y_1,\ldots,y_m)^\top$ (so that $\eta = m$) and let $\mathcal{Y} = \mathbb{R}^m_+$. We then define

$$\Gamma(z,y;\mathbf{R},\mathbf{F},\mathbf{D}) = \Gamma_1(z;\mathbf{R}) + \Gamma_2(z,y;\mathbf{F},\mathbf{D})$$

where $\Gamma_1(z;\mathbf{R}) = \mathbf{R}^\top(\mathbf{1}-z)$ and $\Gamma_2(z,y;\mathbf{F},\mathbf{D}) = \mathbf{F}^\top z + c^\top y + g(z,y;\mathbf{D})$, with $c = (c_1,\ldots,c_m)^\top$ and $g(z,y;d)$ denoting the optimal solution value to the following optimization problem:

$$\min \sum_{i=1}^{m}\sum_{j=1}^{m} s_{ij}x_{ij} + \sum_{i=1}^{m} e_i\left(d_i z_i - x_{i0}\right)^+ - \sum_{i=1}^{m} v_i\left(x_{i0} - d_i z_i\right)^+$$

subject to

$$x_{i0} + \sum_{j=1}^{m} x_{ij} = y_i + \sum_{j=1}^{m} x_{ji} \qquad\qquad i = 0, 1, \ldots, m$$

$$x_{ij} \geq 0 \qquad\qquad i, j = 0, 1, , \ldots, m.$$

Here x_{ij} $(i, j = 1, 2, \ldots, m)$ are the (recourse) transshipment quantities between the markets and x_{i0} $(i = 1, \ldots, m)$ are the quantities used to satisfy market demands. These decisions are made in the second stage after demands have been realized. We assume that $0 < v_i < \min_{j=1,\ldots,m}(c_j + s_{ji}) < e_i$ (for $i = 1, \ldots, m$). This ensures convexity, existence of an optimal solution to (P), and monotonicity of Γ_2 in z. Moreover, it is easy to see that g is homogeneous in (z, y), i.e., for any positive constant $\lambda > 0$ we have $g(\lambda z, \lambda y; d) = \lambda g(z, y; d)$ for all $d > 0$. All of these properties will prove useful later on. Finally, note that this model could allow for transshipment through an unselected market. The reasonable condition that $s_{ij} > c_j - c_i$ (for $i, j = 1, \ldots, m$) prevents this from happening.

3.2.3 Examples: Risk measures

3.2.3.1 A class of optimization-based coherent risk measures

A coherent risk measure that has gained a significant amount of attention is *Conditional Value-at-Risk* (CVaR) (see, e.g., [43, 44, 45]). This measure is given by:

$$\rho[X] = \inf_{\theta \in \Theta} E\left[\theta + \frac{1}{1-\alpha}(X - \theta)^+\right]$$

for some $\alpha \in [0, 1)$ (e.g., $\alpha = 0.05$) and $\Theta = \mathbb{R}$. Later in this chapter we will consider risk measures of the following general form:

$$\rho[X] = \inf_{\theta \in \Theta} E\left[G(X; \theta)\right]$$

where G is some real-valued function parameterized by θ. This choice is motivated by a desire to, on the one hand, consider a wide class of risk measures that includes CVaR while, on the other hand, allowing for an efficiently implementable approximation algorithm. The following lemma characterizes conditions on G under which ρ is a coherent risk measure:

Lemma 3.2.4. *Suppose that Θ is a linear space and $G : \mathbb{R} \times \Theta \to \mathbb{R}$ is*

(i) convex in $(x; \theta)$;

(ii) monotonely nondecreasing in x;

(iii) translation invariant, i.e., $G(x + a; \theta + a\mathbf{1}) = G(x; \theta) + a$ for all $a \in \mathbb{R}$, where $\mathbf{1}$ is the vector of all ones; and

(iv) positively homogeneous, i.e., $G(\lambda x; \lambda\theta) = \lambda G(x; \theta)$ for all $\lambda > 0$.

*Then the risk measure ρ is coherent, i.e., it satisfies Assumption 3.2.1 (*coherence*).*

Proof. See Appendix A.1. □

This definition includes CVaR as a special case, but is a more general class of (still coherent) risk measures.

3.2.3.2 A class of utility-function–based risk measures

A second class of risk measures that we would like to mention explicitly is the certainty equivalent of a random cost with respect to a utility function $u : \mathbb{R}_+ \to \mathbb{R}$:

$$\rho[X] = u^{-1} \left(E \left[u(X) \right] \right).$$

The following lemma characterizes conditions on u under which ρ satisfies at least three of the four conditions to be a coherent risk measure:

Lemma 3.2.5. *Suppose that u is four times continuously differentiable and* *(i)* $u(x), u'(x), u''(x) > 0$ *for all* $x > 0$

(ii) *either* $u'(x)/u''(x)$ *is concave in* x *or* $u'(x)u'''(x)/u''(x)^2$ *is increasing in* x *for* $x > 0$*;*

(iii) u *is positively homogeneous of order* $k \geq 1$*, i.e.,* $u(\lambda x) = \lambda^k u(x)$ *for all* $\lambda > 0$ *and* $x > 0$.

Then the risk measure ρ *satisfies parts (i)–(ii) and (iv) of Assumption 3.2.1 (coherence).*

Proof. [46] show that conditions (i) and (ii) in the lemma imply that ρ is convex, so that the part (i) of Assumption 3.2.1 (*coherence*) is satisfied. Also, parts (ii) and (iv) of Assumption 3.2.1 (*coherence*) follow immediately from the fact that u is increasing (since $u' > 0$) and condition (iii) in the lemma. $\square$

Although this class of risk measures is not coherent, most of the results derived in this chapter will still apply. Note also that if we would simply let $\rho[X] = E\left[u(X)\right]$ then, under the conditions in Lemma 3.2.5, parts (i)–(ii) of Assumption 3.2.1 (*coherence*) are still satisfied, as well as a relaxation of part (iv) of that assumption to positive homogeneity of order k; this is also sufficient for most of the results derived in this chapter to apply, with minor modifications.

3.2.4 Convexity of (R)

The following lemma provides conditions on $\mathcal{Y}$, Γ, and ρ that guarantee that (R) is a convex optimization problem.

Lemma 3.2.6. *Suppose that parts (i)–(ii) of Assumption 3.2.1 (coherence) and Assumption 3.2.3 (convexity and attainment) hold. Then the function* Ψ *is convex on* $[0,1]^m \times \mathcal{Y}$.

Proof. Let $z, z' \in [0,1]^m$, $y, y' \in \mathcal{Y}$, and $\lambda \in [0,1]$. Then

$$\Psi\left(\lambda z + (1-\lambda)z', \lambda y + (1-\lambda)y'\right) = \rho\left[\Gamma\left(\lambda z + (1-\lambda)z', \lambda y + (1-\lambda)y'; \mathbf{R}, \mathbf{F}, \mathbf{D}\right)\right]$$

$$\leq \rho\left[\lambda\Gamma(z,y;\mathbf{R},\mathbf{F},\mathbf{D}) + (1-\lambda)\Gamma(z',y';\mathbf{R},\mathbf{F},\mathbf{D})\right]$$

(by convexity of Γ and monotonicity of ρ)

$$\leq \lambda\rho\left[\Gamma(z,y;\mathbf{R},\mathbf{F},\mathbf{D})\right] + (1-\lambda)\rho\left[\Gamma(z',y';\mathbf{R},\mathbf{F},\mathbf{D})\right]$$

(by convexity of ρ)

$$= \lambda\Psi(z,y) + (1-\lambda)\Psi(z',y')$$

which yields the desired result. □

We see from this lemma that $\mathbf{R}$ is a convex optimization problem as long as a modest set of assumptions hold. Note that convexity of (R) does not necessarily mean that the problem is efficiently solvable. Computing the objective function will often be very computationally intensive or impossible because of randomness and the risk measure, even if the demand distributions are explicitly known. In Section 3.3 we will first propose approximation algorithms that do rely on the optimal solution to (R). In Section 3.4 we will then modify the approaches to allow for solving (R) approximately.

3.3 Core approximation approach

In this section we present our core approximation result. The approximation algorithm that we will develop is based on (i) solving the continuous optimization problem (R), and (ii) rounding the solution to obtain a feasible solution to (P). In particular, we use the following rounding operator that generalizes ordinary rounding to the nearest integer:

$$[z_i]_\beta = \begin{cases} 1 & \text{if } z_i \geq 1 - \beta \\ 0 & \text{otherwise} \end{cases}$$

where $z \in [0,1]^m$ and $[z]_\beta = ([z_1]_\beta, \ldots, [z_m]_\beta)^\top$ for $\beta \in [0,1]$. For convenience, let $b = 1/\min(\beta, 1-\beta)$. Now assume that the function Γ satisfies the following

Assumption 3.3.1. *Scalability:*

For all $z \in [0,1]^m$ and $y \in \mathcal{Y}$,

$$\Gamma\left([z]_\beta, \frac{1}{1-\beta} y; \mathbf{R}, \mathbf{F}, \mathbf{D}\right) \preceq b\Gamma\left(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}\right).$$

This assumption is key to the results in this chapter. In particular, if Γ satisfies Assumption 3.3.1, a rounded decision vector will not be "too much" more costly than the original continuous vector. Intuitively, this scaling is present in many kinds of problems where there is a linear benefit to selections (revenue in a market selection is one example, but we could also consider tax breaks for capital investments), and has linear cost when failing to satisfy demands (or in general terms, requirements).

3.3.1 Approximation results

Our first result is that, under Assumption 3.3.1 (*scalability*), any feasible solution to (R) can be rounded and scaled to a feasible solution to (P) so that the cost of the latter is within a factor b of the cost of the former. This is formalized in the following lemma.

Lemma 3.3.2. *Suppose that ρ satisfies parts (i)–(ii) and (iv) of Assumption 3.2.1 (coherence) and Γ satisfies Assumption 3.3.1 (scalability). Let $(\hat{z}, \hat{y})$ be a feasible solution to (R). Then*

$$\Psi\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}\right) \leq b\Psi(\hat{z}, \hat{y})$$

i.e., $([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y})$ *is a feasible solution to (P) with cost no more than* b *times that of* $(\hat{z}, \hat{y})$.

Proof. Since $\mathcal{Y}$ is a cone we know that $\frac{1}{1-\beta}\hat{y} \in \mathcal{Y}$ so that $\Psi\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}\right)$ is well-defined. Then Assumption 3.3.1 (*scalability*) implies that

$$\Psi\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}\right) = \rho\left[\Gamma\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}; \mathbf{R}, \mathbf{F}, \mathbf{D}\right)\right]$$

$$\leq \rho\left[b\Gamma\left(\hat{z}, \hat{y}; \mathbf{R}, \mathbf{F}, \mathbf{D}\right)\right]$$

(by the monotonicity property of ρ)

$$= b\rho\left[\Gamma\left(\hat{z}, \hat{y}; \mathbf{R}, \mathbf{F}, \mathbf{D}\right)\right]$$

(by the positive homogeneity property of ρ)

$$= b\Psi(\hat{z}, \hat{y})$$

which is the desired result. □

The previous lemma demonstrates the implementation of scaling and how to choose an appropriate y for the rounded selection decision. As the following theorem formalizes, this result can be further used to obtain a feasible solution to (P) with cost no more than b times that of the overall *optimal* solution if we let $(\bar{z}, \bar{y})$ be an optimal solution to (R).

Theorem 3.3.3. *Suppose that* ρ *satisfies parts (i)–(ii) and (iv) of Assumption 3.2.1 (coherence) and* Γ *satisfies Assumption 3.3.1 (scalability). Let* $(\bar{z}, \bar{y})$ *be an optimal solution to (R). Then*

$$\Psi^*([\bar{z}]_\beta) \leq \Psi\left([\bar{z}]_\beta, \frac{1}{1-\beta}\bar{y}\right) \leq b\Psi^*.$$

Proof. The first inequality follows easily since $\frac{1}{1-\beta}\bar{y} \in \mathcal{Y}$, but it is not necessarily optimal given the selection vector $[\bar{z}]_\beta$. The second inequality follows from the result of Lemma

3.3.2:

$$\Psi\left([\bar{z}]_\beta, \frac{1}{1-\beta}\bar{y}\right) \leq b\Psi(\bar{z}, \bar{y}) = b\underline{\Psi} \leq b\Psi^*.$$

$\square$

This theorem in fact provides two feasible solutions to (P), each of which has cost no more than b times the optimal cost: $([\bar{z}]_\beta, \frac{1}{1-\beta}\bar{y})$ and $([\bar{z}]_\beta, \bar{y}^*)$, where $\bar{y}^*$ is the optimal value of y corresponding to the selection vector $[\bar{z}]_\beta$. In other words, Theorem 3.3.3 can be viewed as implicitly providing two b-approximation algorithms for (P). Choosing $\beta = \frac{1}{2}$ yields the smallest value of $b = 2$ so that we have 2-approximation algorithms for (P). In case Γ separates into a Γ_1-term that represents deterministic rewards equal to $\bar{r}^\top(\mathbf{1} - z)$ for some constant vector $\bar{r}$ and a Γ_2-term that only depends on fixed costs and demands and satisfies assumption (ii) in the following theorem (which is somewhat stronger than Assumption 3.3.1), we can improve this approximation result. Corollary 3.3.5 provides the improved result.

Theorem 3.3.4. *Suppose that*

(i) $\Gamma_1(z; \mathbf{R}) = \bar{r}^\top(\mathbf{1} - z)$ *for all* $z \in [0,1]^m$*;*

(ii) Γ_2 *satisfies*

$$\Gamma_2\left([z]_\beta, \frac{1}{1-\beta}y; \mathbf{F}, \mathbf{D}\right) \preceq \frac{1}{1-\beta}\Gamma_2\left(z, y; \mathbf{F}, \mathbf{D}\right);$$

(iii) ρ *satisfies Assumption 3.2.1 (coherence).*

Let $(\hat{z}, \hat{y})$ *be a feasible solution to (R). If* $\boldsymbol{\beta}$ *is a random variable uniformly distributed on* $[0, \delta]$ *with* $0 < \delta \leq 1$ *then*

$$E_{\boldsymbol{\beta}}\left[\Psi\left([\hat{z}]_{\boldsymbol{\beta}}, \frac{1}{1-\boldsymbol{\beta}}\hat{y}\right)\right] \leq \frac{1}{\delta}\max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\}\Psi(\hat{z}, \hat{y})$$

i.e., the expected solution value of the random solution $\left([\hat{z}]_{\boldsymbol{\beta}}, \frac{1}{1-\boldsymbol{\beta}}\hat{y}\right)$ *is no larger than* $\frac{1}{\delta}\max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\}$ *times that of* $(\hat{z}, \hat{y})$.

Proof. See Appendix A.2. $\square$

Now $\min_{0<\delta\leq 1}\frac{1}{\delta}\max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\} = \left(1-e^{-1}\right)^{-1}$ and $1-e^{-1}$ is the corresponding optimal value of δ. The following corollary then says that we can obtain a feasible solution to (P) with cost no more than $\left(1-e^{-1}\right)^{-1}$ times that of the overall *optimal* solution if we let $(\bar{z}, \bar{y})$ be an optimal solution to (R). This improves the result of Theorem 3.3.3.

Corollary 3.3.5. *Let the assumptions of Theorem 3.3.4 be satisfied, let* $(\bar{z}, \bar{y})$ *be an optimal solution to (R), and let*

$$\bar{Z} = \left\{[\bar{z}]_\beta : \beta \in [0, 1)\right\}.$$

Then when $\boldsymbol{\beta}$ *is uniformly distributed on* $[0, 1-e^{-1}]$ *we have that*

$$\min_{z\in\bar{Z}}\Psi^*(z) \leq E_{\boldsymbol{\beta}}\left[\Psi\left([\bar{z}]_{\boldsymbol{\beta}}, \frac{1}{1-\boldsymbol{\beta}}\bar{y}\right)\right] \leq \frac{1}{1-e^{-1}}\Psi(\bar{z}, \bar{y}) = \frac{1}{1-e^{-1}}\underline{\Psi} \leq \frac{1}{1-e^{-1}}\Psi^*.$$

Proof. The first inequality follows by the definition of $\bar{Z}$ and the fact that there always exists a realization of a random variable with value no more than the expected value of the random variable. The second inequality follows from the result of Theorem 3.3.4, and the remaining steps are straightforward. $\square$

Note that when $\Gamma_1(z; \mathbf{R}) = \mathbf{R}^\top(\mathbf{1} - z)$ and $\rho = E$, Assumption (i) in Theorem 3.3.4 is immediately satisfied with $\bar{r} = E[\mathbf{R}]$.

Remark 3.3.6. *Corollary 3.3.5 enumerates each rounding cutoff* β *that leads to a distinct selection decision and chooses the best one. This is an improvement over Theorem 3.3.3 which, while defined for any* β*, is optimized by choosing* $\beta = 1/2$ *and results in a 2-approximation.*

3.3.2 Examples revisited

In this section we show that the assumptions we have made on Γ in Section 3.3.1 are satisfied for the two examples we introduced in Section 3.2.2. Some properties associated with the rounding operator will be very useful. In particular it is easy to show that, for all $z_i \in [0,1]$ and $\beta \in (0,1)$,

$$[z_i]_\beta \leq \frac{1}{1-\beta} z_i \tag{3.1}$$

and

$$1 - [z_i]_\beta \leq \frac{1}{\beta}(1 - z_i). \tag{3.2}$$

The following lemmas show that, for Examples 1 and 2 in Sections 3.2.2.1 and 3.2.2.2, Γ satisfies Assumption 3.3.1 so that the approximation results apply.

Lemma 3.3.7. *Suppose Γ is as in Example 1 (Section 3.2.2.1). Then Assumption 3.3.1 is satisfied.*

Proof. Recall that $\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}) = \Gamma_1(z; \mathbf{R}) + \Gamma_2(z, y; \mathbf{F}, \mathbf{D})$ with Γ_1 and Γ_2 as in Section 3.2.2.1. We will therefore study the two components of Γ separately. First,

$$\Gamma_1([z]_\beta; \mathbf{R}) = \mathbf{R}^\top (\mathbf{1} - [z]_\beta) \preceq \frac{1}{\beta} \mathbf{R}^\top (\mathbf{1} - z) = \frac{1}{\beta} \Gamma_1(z; \mathbf{R})$$

where the inequality $\preceq$ follows from inequality (**??**) and the nonnegativity of $\mathbf{R}$. Next,

$$\begin{aligned}
&\Gamma_2\left([z]_\beta, \frac{1}{1-\beta} y; \mathbf{F}, \mathbf{D}\right) \\
&= \mathbf{F}^\top [z]_\beta + \frac{1}{1-\beta} c y + e\left(\mathbf{D}^\top [z]_\beta - \frac{1}{1-\beta} y\right)^+ - v\left(\frac{1}{1-\beta} y - \mathbf{D}^\top [z]_\beta\right)^+ \\
&\preceq \frac{1}{1-\beta} \mathbf{F}^\top z + \frac{1}{1-\beta} c y + e\left(\frac{1}{1-\beta} \mathbf{D}^\top z - \frac{1}{1-\beta} y\right)^+ - v\left(\frac{1}{1-\beta} y - \frac{1}{1-\beta} \mathbf{D}^\top z\right)^+
\end{aligned}$$

(by inequality (3.1), and the facts that $\mathbf{F}$ has nonnegative support and v, e are nonnegative)

$$
\begin{aligned}
&= \frac{1}{1-\beta}\mathbf{F}^\top z + \frac{1}{1-\beta}cy + \frac{1}{1-\beta}e\left(\mathbf{D}^\top z - y\right)^+ - \frac{1}{1-\beta}v\left(y - \mathbf{D}^\top z\right)^+ \\
&= \frac{1}{1-\beta}\Gamma_2\left(z, y; \mathbf{F}, \mathbf{D}\right).
\end{aligned}
$$

Combining these two inequalities yields the desired result. □

Lemma 3.3.8. *Suppose Γ is as in Example 2 (Section 3.2.2.2). Then Assumption 3.3.1 is satisfied.*

Proof. Since Γ_1 is of the same form as in Example 1, it immediately follows that $\Gamma_1\left([z]_\beta; \mathbf{R}\right) \preceq \frac{1}{\beta}\Gamma_1\left(z; \mathbf{R}\right)$. Next,

$$
\begin{aligned}
\Gamma_2\left([z]_\beta, \frac{1}{1-\beta}y; \mathbf{F}, \mathbf{D}\right) &= \mathbf{F}^\top [z]_\beta + \sum_{i=1}^m c_i \frac{1}{1-\beta}y_i + g\left([z]_\beta, \frac{1}{1-\beta}y; \mathbf{D}\right) \\
&\preceq \frac{1}{1-\beta}\mathbf{F}^\top z + \frac{1}{1-\beta}\sum_{i=1}^m c_i y_i + g\left(\frac{1}{1-\beta}z, \frac{1}{1-\beta}y; \mathbf{D}\right)
\end{aligned}
$$

(by inequality (3.1), nonnegativity of $\mathbf{F}$, and monotonicity of g in z)

$$
= \frac{1}{1-\beta}\mathbf{F}^\top z + \frac{1}{1-\beta}\sum_{i=1}^m c_i y_i + \frac{1}{1-\beta}g\left(z, y; \mathbf{D}\right)
$$

(since g is homogeneous in (z, y))

$$
= \frac{1}{1-\beta}\Gamma_2\left(z, y; \mathbf{F}, \mathbf{D}\right).
$$

Combining the two inequalities yields the desired result. □

3.3.3 Towards efficiently implementable approximation algorithms

These results may, in some situations, yield an efficiently implementable b-approximation algorithm for (P). For example, if the support of $(\mathbf{R}, \mathbf{D})$ is finite, Γ is a newsvendor-type cost function, and ρ is a convex combination of expectation and Conditional Value-at-Risk, the relaxation (R) reduces to a linear program. However, in general (R) will not be a linear program or the cardinality of the support of $(\mathbf{R}, \mathbf{D})$ makes (R) intractable. We will therefore resort to the use of sample average approximation (SAA), a technique often used for solving stochastic programming problems (see, e.g., [47, 48, 49]). We will show that this approach can be employed to develop efficiently implementable algorithms that, with high probability, can find a feasible solution to (P) bounded by an affine function of the optimal cost for a large class of selection problems with risk measure ρ as in Section 3.2.3.2 and convex Ψ as in Section 3.2.4.

3.4 Sampling-based approximation approach

In the previous section we presented a collection of algorithms for solving (P). In this section, we introduce using sample average approximation as a method to approximately solve (P) when even the relaxation (R) is intractible. SAA is particularly effective if the second stage problem has complete recourse (i.e., for every set of first-stage decisions, it is possible to find a feasible solution to the second stage problem), as is the case when retailers face newsvendor-type costs with a reasonable shortage cost e_i.

3.4.1 Approximation algorithms

As mentioned in Section 3.3.3, the approximation approach developed thus far is practical if the support of $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ is finite (and of manageably small cardinality) and the resulting convex optimization problem (R) is tractable. However, in general the support of $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ may be continuous and/or the cardinality of its (finite) support too large to handle efficiently. We will therefore consider discrete approximations $(\mathbf{R}_{\mathcal{N}}, \mathbf{F}_{\mathcal{N}}, \mathbf{D}_{\mathcal{N}})$ of the random vector $(\mathbf{R}, \mathbf{F}, \mathbf{D})$ with finite support $\mathcal{N}$. We suppose that the discrete approximation is sampled in an i.i.d. fashion from the underlying distribution vectors (i.e., each sample is independent of other samples, but the individual sample may have underlying dependencies). Then define the corresponding approximate risk function:

$$\Psi_{\mathcal{N}}(z, y) \equiv \rho \left[\Gamma \left(z, y; \mathbf{R}_{\mathcal{N}}, \mathbf{F}_{\mathcal{N}}, \mathbf{D}_{\mathcal{N}} \right) \right],$$

let ($\mathrm{P}_{\mathcal{N}}$) denote the approximation of (P) where Ψ is replaced by $\Psi_{\mathcal{N}}$, and let ($\mathrm{R}_{\mathcal{N}}$) denote its continuous relaxation. We choose the discrete approximation $(\mathbf{R}_{\mathcal{N}}, \mathbf{F}_{\mathcal{N}}, \mathbf{D}_{\mathcal{N}})$ to consist of scenarios of equal probability, which results in an SAA problem.

In the remainder we will assume that problems of the form ($\mathrm{R}_{\mathcal{N}}$) can be solved efficiently. In particular, we will restrict ourselves to the class of expectation-based risk measures introduced in Section 3.2.3.1. In that case, (P) can be reformulated as

$$\underset{z \in \{0,1\}^m, y \in \mathcal{Y}, \theta \in \Theta}{\text{minimize}} \quad E \left[G \left(\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}); \theta \right) \right].$$

Under the assumptions in Lemma 3.2.4 and Assumption 3.2.3, the continuous relaxation (R) of (P) is a convex programming problem. For notational convenience we will, in the remainder of this chapter, simply merge the two decision vectors y and θ into a single vector. With a slight abuse of notation we will still refer to this new and (typically) larger vector as $y \in \mathcal{Y}$, so that (P) simply reads

$$\underset{z \in \{0,1\}^m, y \in \mathcal{Y}}{\text{minimize}} \quad E \left[G \left(\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}) \right) \right]$$

and the corresponding SAA of (R):

$$\underset{z \in [0,1]^m, y \in \mathcal{Y}}{\text{minimize}} \quad \frac{1}{|\mathcal{N}|} \sum_{(r,f,d) \in \mathcal{N}} G \left(\Gamma(z, y; r, f, d) \right). \tag{$\mathrm{R}_{\mathcal{N}}$}$$

In many cases, ($R_{\mathcal{N}}$) may be reformulated into a linear program.

The remarks following Theorem 3.3.3 now suggest the following class of LP and rounding-based approximation algorithms:

Rounding-based Approximation Algorithm (RAA)

Step 0. Select a sample $\mathcal{N}_1$ and solve ($R_{\mathcal{N}_1}$) to optimality, yielding a solution $\left(\bar{z}^{(1)}, \bar{y}^{(1)}\right)$.

Step 1. Set z equal to $z^{(1)} \equiv [\bar{z}^{(1)}]_\beta$.

Step 2. Set y equal to $y^{(1)} \equiv \frac{1}{1-\beta}\bar{y}^{(1)}$.

Note that Step 3 could be replaced by one of the following:

Step 3′. Select a sample $\mathcal{N}_2$ and set y equal to $\hat{y}^{(1)} \equiv \arg\min_{y \in \mathcal{Y}} \Psi_{\mathcal{N}_2}\left(z^{(1)}, y\right)$.

or

Step 3″. Set y equal to $y^{*(1)} \equiv \arg\min_{y \in \mathcal{Y}} \Psi\left(z^{(1)}, y\right)$.

We will refer to these modified algorithms as $\widehat{\mathrm{RAA}}$ and RAA^*, respectively, where, as a default, we will choose $\beta = \frac{1}{2}$, which optimizes the performance guarantee.

Next, Corollary 3.3.5, suggests the following extensions of these algorithms that essentially perform a search over different values of β. Due to this search, there is an additional complication involved in comparing the quality of the different candidate solutions.

Optimal Rounding-based Approximation Algorithm (ORAA)

Step 0. Select a sample $\mathcal{N}_1$ and solve ($R_{\mathcal{N}_1}$) to optimality, yielding a solution $\left(\bar{z}^{(2)}, \bar{y}^{(2)}\right)$.

Step 1. Select a sample $\mathcal{N}_2$ and set β equal to $\bar{\beta} \equiv \arg\min_{\beta \in [0,1)} \Psi_{\mathcal{N}_2}\left([\bar{z}^{(2)}]_\beta, \frac{1}{1-\beta}\bar{y}^{(2)}\right)$.
Set z equal to $z^{(2)} \equiv [\bar{z}^{(2)}]_{\bar{\beta}}$.
Set y equal to $y^{(2)} \equiv \frac{1}{1-\bar{\beta}}\bar{y}^{(2)}$.

Note that Step 2 could be replaced by one of the following (where $\bar{Z}^{(2)} = \{[\bar{z}^{(2)}]_\beta : \beta \in [0,1)\})$):

Step 2′. Select a sample $\mathcal{N}_2$ and set (z,y) equal to $\left(z^{(2)},\hat{y}^{(2)}\right) \equiv \arg\min_{z\in\bar{Z}^{(2)},y\in\mathcal{Y}} \Psi_{\mathcal{N}_2}(z,y)$.

or

Step 2″. Set (z,y) equal to $\left(z^{(2)},y^{*(2)}\right) \equiv \arg\min_{z\in\bar{Z}^{(2)},y\in\mathcal{Y}} \Psi(z,y)$.

We will refer to these modified algorithms as $\widehat{\mathrm{ORAA}}$ and ORAA^*, respectively.

Finally, as an alternative to linear programming and rounding, we could apply SAA directly to (P), yielding another class of approximation algorithms:

IP-based Approximation Algorithm (IPAA)

Step 0. Select a sample $\mathcal{N}_1$ and solve ($\mathrm{P}_{\mathcal{N}_1}$) to optimality, yielding a solution $\left(z^{(3)},y^{(3)}\right)$.

Step 1. Set z equal to $z^{(3)}$.

Step 2. Set y equal to $y^{(3)}$.

Similarly to RAA, Step 3 could be replaced by one of the following:

Step 3′. Select a sample $\mathcal{N}_2$ and set y equal to $\hat{y}^{(3)} \equiv \arg\min_{y\in\mathcal{Y}} \Psi_{\mathcal{N}_2}\left(z^{(3)},y\right)$.

or

Step 3″. Set y equal to $y^{*(3)} \equiv \arg\min_{y\in\mathcal{Y}} \Psi\left(z^{(3)},y\right)$.

We will refer to these modified algorithms as $\widehat{\mathrm{IPAA}}$ and IPAA^*, respectively.

Note that RAA^*, ORAA^*, and IPAA^* will only rarely be computationally tractable. Moreover, it can be expected that IPAA and its variants will be much more computationally inefficient than RAA and its variants.

3.4.2 Approximation results

In this section we will explore required sample sizes for the discussed approaches and combine these results with those of Section 3.3 to obtain approximation algorithms with a probabilistic guarantee. Since (R) has an optimal solution we can, without loss of optimality, restrict the feasible region in the optimization problem (R) to a bounded set $\bar{\mathcal{Y}} \subseteq \mathcal{Y}$. Then define

- $\Delta =$ diameter of $[0,1]^m \times \bar{\mathcal{Y}}$
- $L =$ Lipschitz constant of the function $G(\Gamma(z, y; \mathbf{R}, \mathbf{F}, \mathbf{D}))$ on $[0,1]^m \times \bar{\mathcal{Y}}$

where our assumptions imply that all of these values are finite. While it may be difficult in general to compute these values, in Appendix A.3 we derive explicit bounds for Example 1 from Section 3.2.2.1, with $\rho = E$. We also note that for any fixed $z \in [0,1]^m$ we may be able to compute smaller bounds $\bar{\mathcal{Y}}(z)$, $\Delta(z)$, and $L(z)$.

3.4.2.1 Single sample

We first provide sample size results for the algorithm variants that require only a single sample, $\mathcal{N}_1$.

Theorem 3.4.1. *Let $\mathcal{N}_1$ be a sample satisfying*

$$|\mathcal{N}_1| \geq O(1) \left(\frac{\Delta L}{\tau_1}\right)^2 \left[(m + \dim(\mathcal{Y})) \log\left(\frac{\Delta L}{\tau_1}\right) + \log\left(\frac{O(1)}{\delta_1}\right)\right]$$

where $\tau_1 > 0$ is an absolute cost error measure and $1 - \delta_1$ is a confidence level. Then

(i) RAA and RAA* yield solutions satisfying

$$\Psi\left(z^{(1)}, y^{*(1)}\right) \leq \Psi\left(z^{(1)}, y^{(1)}\right) \leq b\left(\Psi^* + \tau_1\right)$$

with probability at least $1 - \delta_1$;

(ii) ORAA* yields a solution satisfying

$$\Psi\left(z^{(2)}, y^{*(2)}\right) \leq \frac{1}{1 - e^{-1}}\left(\Psi^* + \tau_1\right)$$

with probability at least $1-\delta_1$*;*

(iii) IPAA and IPAA* *yield solutions satisfying*

$$\Psi\left(z^{(3)}, y^{*(3)}\right) \leq \Psi\left(z^{(3)}, y^{(3)}\right) \leq \Psi^* + \tau_1$$

with probability at least $1-\delta_1$*.*

Proof. The first inequalities in (i) and (ii) are obvious. Furthermore,

(i) Theorem 3.3.3 says that

$$\Psi\left(z^{(1)}, y^{(1)}\right) = \Psi\left([\bar{z}^{(1)}]_\beta, \frac{1}{1-\beta}\bar{y}^{(1)}\right) \leq b\Psi\left(\bar{z}^{(1)}, \bar{y}^{(1)}\right).$$

By Theorem 2 of [50] we have that the bound on the sample size $\mathcal{N}_1$ implies that

$$\Psi(\bar{z}^{(1)}, \bar{y}^{(1)}) \leq \underline{\Psi} + \tau_1$$

with probability at least $1-\delta_1$. Finally, it is easy to see that $\underline{\Psi} \leq \Psi^*$. Combining these inequalities yields the first result.

(ii) This follows in a similar way from Theorem 3.3.3 and Corollary 3.3.5.

(iii) This follows immediately from Theorem 2 of [50].

□

3.4.2.2 Two samples

We now provide sample size results for the algorithm variants which call for two samples $\mathcal{N}_1$ and $\mathcal{N}_2$. For our performance guarantees to hold, the two samples should be independently generated since it is unclear how dependence may affect the quality of the decisions.

Theorem 3.4.2. *Let $\mathcal{N}_2$ be a sample satisfying*

$$|\mathcal{N}_2| \geq O(1)\left(\frac{\Delta L}{\tau_2}\right)^2 \left[(\dim(\mathcal{Y}))\log\left(\frac{\Delta L}{\tau_2}\right) + \log\left(\frac{O(1)}{\delta_2}\right)\right]$$

where $\tau_2 > 0$ is an absolute cost error measure and $1-\delta_2$ is a confidence level. In addition, let $\mathcal{N}_1$ meet the condition of Theorem 3.4.1. Then

(i) $\widehat{\mathrm{RAA}}$ *yields a solution satisfying*

$$\Psi\left(z^{(1)}, \hat{y}^{(1)}\right) \leq b\left(\Psi^* + \tau_1\right) + \tau_2$$

with probability at least $(1-\delta_1)(1-\delta_2)$;

(ii) ORAA and $\widehat{\mathrm{ORAA}}$ *yield solutions satisfying*

$$\Psi\left(z^{(2)}, y^{(2)}\right) \leq \frac{1}{1-e^{-1}}\left(\Psi^* + \tau_1\right) + \tau_2$$

and

$$\Psi\left(z^{(2)}, \hat{y}^{(2)}\right) \leq \frac{1}{1-e^{-1}}\left(\Psi^* + \tau_1\right) + \tau_2$$

respectively with probability at least $(1-\delta_1)(1-\delta_2)$.

(iii) $\widehat{\mathrm{IPAA}}$ *yields a solution satisfying*

$$\Psi\left(z^{(3)}, \hat{y}^{(3)}\right) \leq \Psi^* + \tau_1 + \tau_2$$

with probability at least $(1-\delta_1)(1-\delta_2)$.

Proof. (i) (a) By Theorem 2 of [50] we have that the bound on the sample size $\mathcal{N}_2$ implies that

$$\Psi\left(z^{(1)}, \hat{y}^{(1)}\right) \leq \Psi^*\left(z^{(1)}\right) + \tau_2$$

with probability at least $1-\delta_2$.

(b) Theorem 3.4.1(i) says that

$$\Psi^*\left(z^{(1)}\right) \leq b(\Psi^* + \tau_1)$$

with probability at least $1 - \delta_1$.

Combining (a)–(b) yields the desired result.

(ii) Recall the optimization problem over β that ORAA solves in Step 2 obtains the heuristic solution $\left(z^{(2)}, y^{(2)}\right)$. By Theorem 2 of [50] and a similar argument as above we have that the provided bound on the sample sizes $\mathcal{N}_1$ and $\mathcal{N}_2$ imply that

$$\begin{aligned}
\Psi\left(z^{(2)}, y^{(2)}\right) &\leq \inf_{\beta\in[0,1)} \Psi\left([\bar{z}^{(2)}]_\beta, \frac{1}{1-\beta}\bar{y}^{(2)}\right) + \tau_2 \\
&\leq E_{\boldsymbol{\beta}}\left[\Psi\left([\bar{z}^{(2)}]_{\boldsymbol{\beta}}, \frac{1}{1-\boldsymbol{\beta}}\bar{y}^{(2)}\right)\right] + \tau_2 \\
&\leq \frac{1}{1-e^{-1}}\Psi\left(\bar{z}^{(2)}, \bar{y}^{(2)}\right) + \tau_2 \leq \frac{1}{1-e^{-1}}(\Psi^* + \tau_1) + \tau_2
\end{aligned}$$

with probability at least $1 - \delta_2$.

Similarly the optimization problem over β that $\widehat{\text{ORAA}}$ solves in Step 2 to obtain the heuristic solution $\left(z^{(2)}, \hat{y}^{(2)}\right)$ can alternatively be formulated as

$$\min_{\beta\in[0,1), y\in\mathcal{Y}} \Psi_{\mathcal{N}_2}\left([\bar{z}^{(2)}]_\beta, y\right).$$

By Theorem 2 of [50] we then have that the provided bound on the sample size $\mathcal{N}_2$ implies that

$$\begin{aligned}
\Psi\left(z^{(2)}, \hat{y}^{(2)}\right) &\leq \inf_{\beta\in[0,1)} \min_{y\in\mathcal{Y}} \Psi\left([\bar{z}^{(2)}]_\beta, y\right) + \tau_2 \leq \inf_{\beta\in[0,1)} \Psi\left([\bar{z}^{(2)}]_\beta, \frac{1}{1-\beta}\bar{y}^{(2)}\right) + \tau_2 \\
&\leq \frac{1}{1-e^{-1}}(\Psi^* + \tau_1) + \tau_2
\end{aligned}$$

with probability at least $1-\delta_2$.

(iii) (a) By Theorem 2 of [50] we have that the bound on the sample size $\mathcal{N}_2$ implies that

$$\Psi\left(z^{(3)}, \hat{y}^{(3)}\right) \leq \Psi^*\left(z^{(3)}\right) + \tau_2$$

with probability at least $1-\delta_2$.

(b) Theorem 3.4.1(iii) says that

$$\Psi^*\left(z^{(3)}\right) \leq \Psi^* + \tau_1$$

with probability at least $1-\delta_1$.

Combining (a)–(b) yields the desired result.

□

3.4.2.3 Discussion

Note that, perhaps surprisingly, the approximation bound for RAA* is identical to that for RAA, and the result for $\widehat{\mathrm{RAA}}$ is weaker than that for RAA (both in terms of bound and level of confidence). However, it can be expected that, in general, $\widehat{\mathrm{RAA}}$ will perform (much) better than RAA (and similar to the often intractable RAA*) since $y^{(1)}$ is likely to be a far worse solution to the problem with selection vector $z^{(1)}$ than $\hat{y}^{(1)}$ (and the latter is likely to be comparable to $y^{*(1)}$). We also see this same pattern for the IPAA set of algorithms, albeit without the additional factor of 2 on the performance bound relative to the rounded solutions. It is also interesting to note that while implementing $\widehat{\mathrm{ORAA}}$ requires two samples, we are able to obtain a significant performance bound improvement by comparison to $\widehat{\mathrm{RAA}}$ without significantly more computational effort.

The sample size bounds in Theorems 3.4.1 and 3.4.2 show us that the complexity of the problem grows quadratically as a function of how the diameter and Lipschitz constant compare to the absolute gap we allow. We also see that the confidence parameters only appear as the denominator within a log function. Similarly the number of decision variables only appear linearly, except as they directly affect the diameter and implicitly affect the Lipschitz constant. In Appendix A.3 we work through explicit values of the parameters for Example 1 and $\rho = E$. In Appendix A.3.1 we provide an alternative bound on $|\mathcal{N}_2|$ by using the work of [49]. In Appendix A.3.2 we provide both deterministic and probabilistic bounds on the diameter Δ. We also could revise Theorem 3.4.2 to instead use $\Delta(z)$ and $L(z)$ instead of the global parameters, which may be significantly smaller. As an illustration we provide both a bound on L and $L(z)$ in Appendix A.3.3.

Depending on the $O(1)$ term, these approximation bounds may seem to suggest that we need very large samples in order to achieve solutions that are not practical. However, our numerical experiments in Section 3.5 demonstrate that for our first example, the performance of these algorithms can be experimentally very good along with the information provided by the theoretical guarantees. This is consistent with the existing literature on SAA which focuses more on how the sample size grows as a function of problem parameters than computing an exact number of samples to use.

3.5 Computational results

In this section, we will study the performance of three core algorithms (RAA, ORAA, and IPAA) by evaluating and comparing three variants of each of these algorithms (the original one-sample version, a two-sample version, and one which uses the true distribution information after the selection decision has been made), for solving selection problems of the form (P). In particular, we will empirically investigate the sample sizes that are required in practice as well as the value of using true demand information if available.

3.5.1 Problem instances

We will consider risk-neutral and risk-averse versions of the selective newsvendor problem (Example 1 from Section 3.2.2.1), with $\rho = E$ and $\rho =$ CVaR respectively. We assume that the true (joint) demand distribution has finite support, which will allow us to compare the solutions found by the approximations with the true optimal solutions. In particular, we generate each problem instance by sampling a support of the joint demand distribution of finite cardinality Ξ uniformly from $\prod_{i=1}^{m}[0, u_i]$ where u_i is sampled uniformly from $[0, 10]$ (for $i = 1, \ldots, m$). The true demand distribution is then taken to be the uniform distribution on that support. We let the procurement cost, salvage value, and expediting costs be $c = 0.8$, $v = 0.6$, and $e = 1$. The per-unit market revenues r_i are generated uniformly from $[0.8, 1]$ and $\mathbf{R}_i = r_i \mathbf{D}_i$ ($i = 1, \ldots, m$). Finally, we sample the (deterministic) fixed market selection costs $\mathbf{F}_i$ uniformly from $[0, 0.5]$ ($i = 1, \ldots, m$).

3.5.2 Results

In the remainder of this section we will demonstrate how the different algorithms compare in efficiency and performance. All tests were performed on a PC with an Intel Xeon Quad Core 3.2 GHz processor with 8 GB RAM using C++ and CPLEX version 12.5. When needed by the algorithms, CPLEX was used to solve for (z, y), as well as a particular y given a fixed z. We also used CPLEX to find the optimal solution to the true integer problem (which we refer to simply as using CPLEX) For each set of problem dimensions (characterized by m and Ξ) we generate 100 problem instances. Most of our comparisons are based on the "percent error" observed. For a specific decision (z, y), the percent error is defined as

$$100\left(\frac{\Psi(z, y) - \Psi^*}{\Psi^*} - 1\right).$$

3.5.2.1 $\rho = E$

We begin our study by considering problem instances with $\rho = E$ (the expectation measure), $m = 10$, and $\Xi = 1,000$. In Figure 3.1, we compare how the variants of each algorithm perform as the sample size grows. For the two-sample variants, we use two samples of the same size, where the first is used to find the market selection, and the second for the order quantity. We compute the percent error for each problem instance, and then compute the mean and 95$^{\text{th}}$ percentile of that error for the 100 problem instances (i.e., the average and fifth-worst observed error). We see that the IPAA variants perform very similarly to each other, as do $\widehat{\text{RAA}}$ and RAA^*. RAA performs considerably worse than the other RAA variants, and does not result in better solutions as the sample size grows. As a contrast to the other two, ORAA shows different performance for the variants. We see that ORAA^* has considerably better performance than $\widehat{\text{ORAA}}$, which itself does better than ORAA.

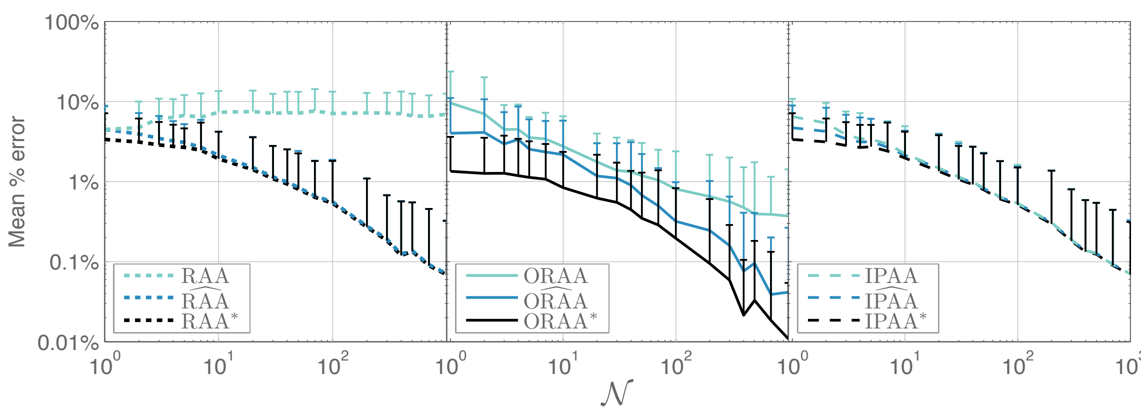


Figure 3.1: Algorithm performance by average % error for $\rho = E$, $m = 10$, and $\Xi = 1,000$.

With the same data as in Figure 3.1, we see in Figure 3.2 how the one-sample algorithms that use scaling to determine the order quantity and the two-sample algorithms that either approximately or exactly solve for the order quantity (indicated by $\widehat{}$ and *, respectively), compare. In this figure, we see that the one-sample algorithms perform qualitatively differently, with RAA performing poorly and IPAA performing well. The algorithms which approximately solve for the order quantity all perform about equally well. In the final comparison, while RAA* and IPAA* perform almost identically, ORAA* obtains considerably better solutions.

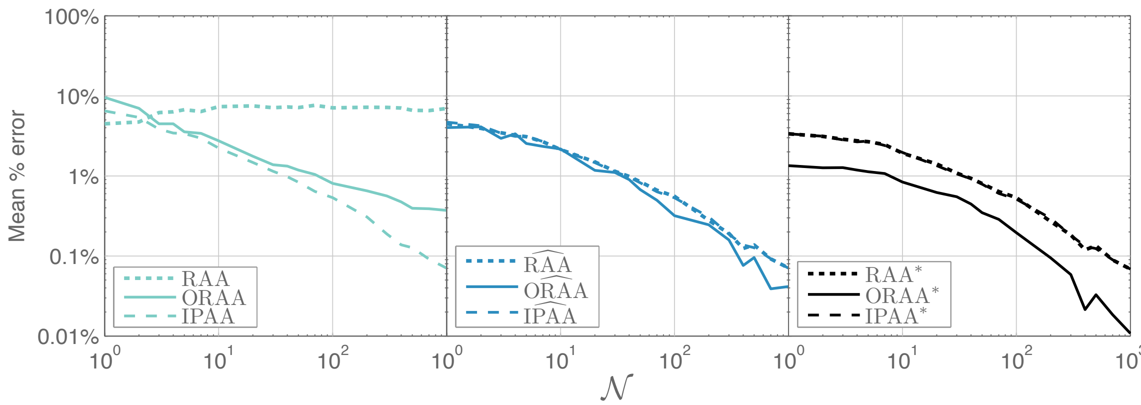


Figure 3.2: Variant performance by average % error for $\rho = E$, $m = 10$, and $\Xi = 1,000$.

In these figures, in addition to the comparative performance, we also see that only relatively small sample sizes are needed to obtain very small errors for our problem. This suggests that while simple rounding is generally not sufficient to obtain high-quality solutions, having knowledge of the true distribution is not necessary either. With the exception of RAA, a sample size of 10 yields an average error of less than 5%, and a sample size of 100 yields an average error below 1%. This is encouraging since using the true demand distribution to determine the order quantity for a fixed set of selected markets is often intractable. Further, the distributions may be continuous and the cardinality of the state space increases exponentially in m. We do see, however, that simple rounding is not competitive: even with a sample size of 1,000 the average error exceeds 10%. Even when we are able to choose the best rounding parameter, as is the case for ORAA, the performance is considerably worse than the other successful algorithms.

If we instead consider larger problem instances with $m = 100$, we discover a slightly different story. While most of the results are similar, ORAA* performs differently. In Figure 3.3, we show the two plots that include ORAA*, and note that it now performs very similarly to $\widehat{\text{ORAA}}$ and the other algorithms which use the true distribution to compute the order quantity. This suggests that when there are relatively few binary variables, choosing the correct rounding parameter is more critical.

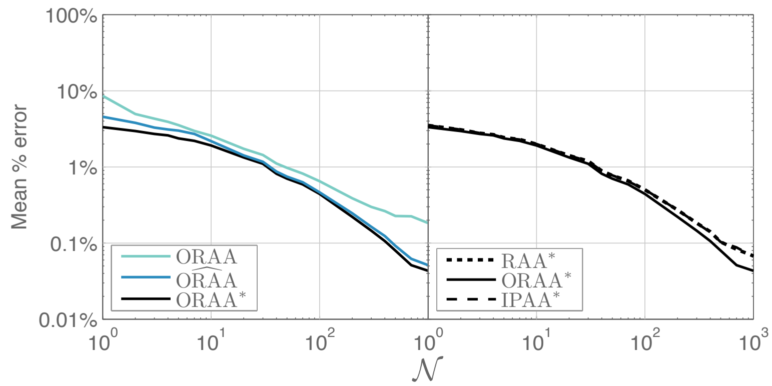


Figure 3.3: Comparison of performance of ORAA* for $\rho = E$, $m = 100$, and $\Xi = 1,000$.

In Table 3.1 we show computation times for $m \in \{10, 100, 500\}$ and $\Xi \in \{1,000; 10,000\}$. The latter values are chosen so that we can compare the approximate solutions to the exact optimal ones. However, note that typically the true distributions will be continuous or discrete with a support that is exponential in m, so that in practice solving the exact problem can be expected to be intractable. As observed earlier, sample sizes of 10 or 100 yield high-quality solutions to our problem, so we include computational times for those numbers of samples. First note that CPLEX takes considerably longer than the proposed approximation algorithms even for unrealistically small values of Ξ. More importantly, and perhaps surprisingly, we see that the IPAA variants solve the problem almost as fast as the RAA variants. This is likely due to the relatively simple nature of the supply chain costs (consisting only of newsvendor costs in this example) and is not expected to extend to more general settings. Finally, note that in this example, finding the optimal order quantity for a given market selection vector only requires computing the $\frac{e-c}{e-v}$-fractile of demand in selected markets. This is an easy task, and results in most versions of the approximation algorithms taking similar time. As we will later see in the example with risk-averse objective in Section 3.5.2.2, there is a much larger difference between the efficiency of the algorithms when, given the market selection vector z, it is still hard to compute the optimal y.

	$\mathcal{N} = 10$						$\mathcal{N} = 100$					
S	1,000			10,000			1,000			10,000		
m	10	100	500	10	100	500	10	100	500	10	100	500
RAA	0.00	0.01	0.04	0.01	0.03	0.14	0.01	0.02	0.27	0.01	0.04	0.37
$\widehat{\mathrm{RAA}}$	0.00	0.01	0.04	0.01	0.03	0.14	0.01	0.02	0.27	0.01	0.04	0.37
RAA^*	0.00	0.01	0.05	0.01	0.04	0.17	0.01	0.02	0.28	0.01	0.05	0.40
ORAA	0.01	0.02	0.07	0.01	0.04	0.16	0.02	0.05	0.41	0.02	0.07	0.50
$\widehat{\mathrm{ORAA}}$	0.02	0.03	0.11	0.02	0.07	0.30	0.02	0.08	0.69	0.03	0.12	0.86
ORAA^*	0.02	0.06	0.31	0.12	0.53	2.86	0.03	0.10	0.69	0.14	0.70	4.46
IPAA	0.01	0.03	0.10	0.01	0.06	0.19	0.02	0.04	0.31	0.02	0.06	0.41
$\widehat{\mathrm{IPAA}}$	0.01	0.03	0.10	0.01	0.06	0.19	0.02	0.04	0.31	0.02	0.06	0.41
IPAA^*	0.01	0.03	0.10	0.01	0.07	0.23	0.02	0.04	0.32	0.02	0.07	0.45
CPLEX	0.12	0.41	3.85	6.02	10.61	51.99	0.12	0.41	3.85	6.02	10.61	51.99

Table 3.1: Time required to solve the problem with $\boldsymbol{\rho = E}$ in seconds.

3.5.2.2 $\rho =$ CVaR

Our second set of numerical experiments consider a risk-based objective. In particular, we use $\rho =$ CVaR with $\theta = 0.05$, i.e., we minimize the upper 5% tail average of cost. We consider $m = 10$ and $\Xi = 1,000$, and implement the same collection of algorithms.

Comparing Figures 3.4 and 3.5 to Figures 3.1 and 3.2 we see that the average error for $\rho =$ CVaR for all algorithms and for a given sample size is considerably higher than for $\rho = E$. We continue to see that RAA performs very poorly, and ORAA also does not perform as well as the versions with two-samples or exact information after 50 samples. We also see that the algorithm variants which use information from the true problem to choose the y vector offer a significant improvement over the other algorithms until there are a relatively large number of samples. This observation is consistent with our understanding of risk measures, since it requires a large number of samples to even observe one of the high risk realizations (there is some work to address this issue, see for example [51]). The two-sample algorithms continue to have virtually identical performance over the range we have studied, however, the algorithms which exactly solve for the order quantity diverge as the number of samples increase. These results show that, for a risk averse decision maker, we need a larger number of samples than for a risk neutral decision maker. Intuitively this can be explained by observing that a larger number of samples will be required to accurately estimate a tail average (CVaR) than the full population average.

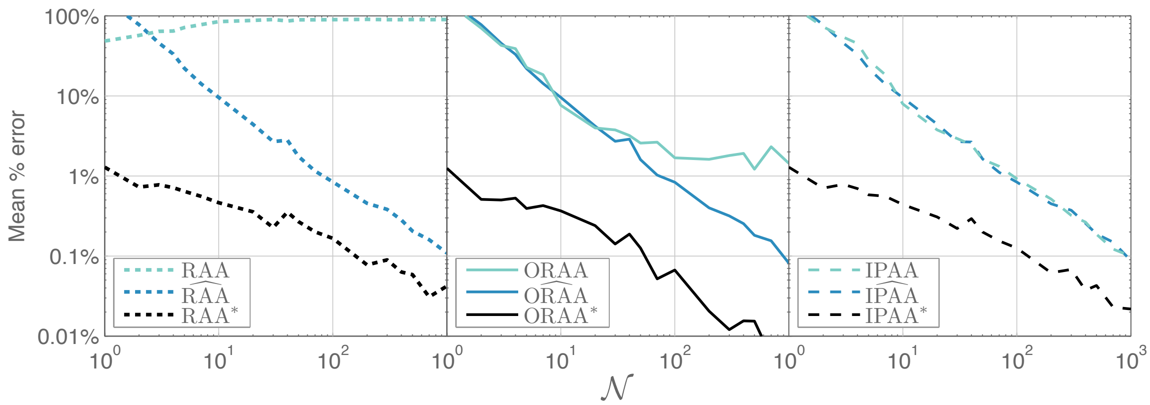


Figure 3.4: Algorithm performance by average % error for ρ = CVaR, $m = 10$, and $\Xi = 1,000$.

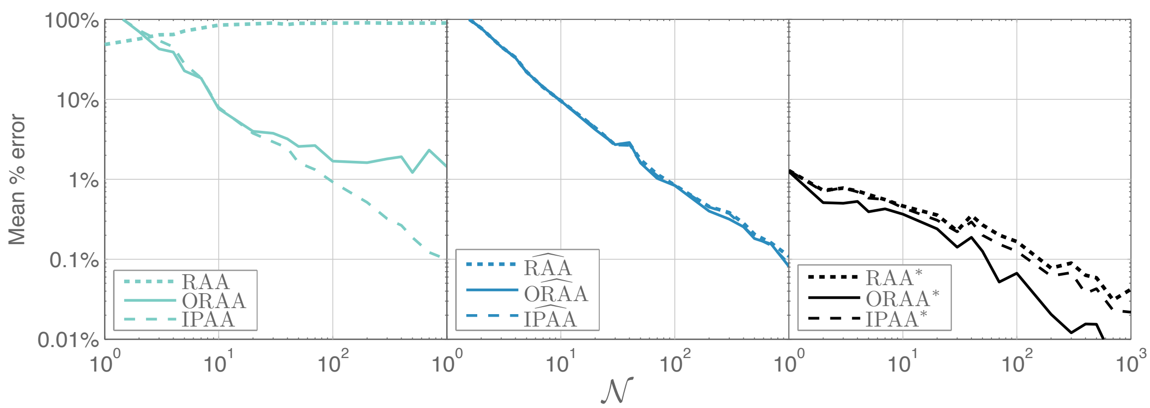


Figure 3.5: Variant performance by average % error for ρ = CVaR, $m = 10$, and $\Xi = 1,000$.

Finally, Table 3.2 shows the running times of the various algorithms for this more difficult class of problems. We again observe that CPLEX takes considerably longer than our algorithms, even though the support of the true distribution is small (as in the previous examples). Further, in this case we are not able to use a critical-fractile type solution for the y variables. This seemingly small change makes a big difference in the efficiency of the algorithms, and it becomes much more attractive to consider approximate methods.

	$\mathcal{N} = 10$						$\mathcal{N} = 100$					
S	1,000			10,000			1,000			10,000		
m	10	100	500	10	100	500	10	100	500	10	100	500
RAA	0.00	0.01	0.07	0.01	0.05	0.27	0.01	0.04	0.60	0.01	0.08	0.79
$\widehat{\text{RAA}}$	0.01	0.02	0.14	0.02	0.11	0.52	0.02	0.07	1.05	0.03	0.16	1.42
RAA^*	0.06	0.27	4.26	1.89	3.35	43.02	0.06	0.30	4.79	1.94	3.36	43.58
ORAA	0.01	0.03	0.16	0.03	0.12	0.54	0.02	0.11	1.21	0.04	0.19	1.60
$\widehat{\text{ORAA}}$	0.01	0.03	0.16	0.03	0.12	0.54	0.03	0.11	1.22	0.04	0.20	1.60
ORAA^*	0.13	0.48	5.39	4.80	7.08	56.92	0.17	0.77	7.09	6.71	12.26	72.98
IPAA	0.01	0.03	0.11	0.01	0.08	0.30	0.02	0.26	1.13	0.02	0.30	1.38
$\widehat{\text{IPAA}}$	0.01	0.04	0.18	0.02	0.14	0.56	0.03	0.28	1.57	0.03	0.38	2.01
IPAA^*	0.06	0.29	4.30	1.87	3.34	43.08	0.07	0.51	5.32	1.91	3.57	44.18
CPLEX	0.20	2.05	20.67	6.42	36.49	527.38	0.20	2.05	20.67	6.42	36.49	527.38

Table 3.2: Time required to solve the problem in seconds with $\boldsymbol{\rho}$ =CVaR and 5% tail

3.5.3 Discussion

In this section we have examined the performance of three variants of our three approximation approaches. Throughout our results, we see that the error decreases very rapidly in sample size for our problem. We also see that, while RAA is an obvious poor choice and ORAA is probably not a practical option in most cases, the remainder of the algorithms perform about equally well on our test problems. Especially of value is that, with very few samples and hence limited computational effort, we are able to obtain solutions that are close to optimal. These approximation results are promising for related problems since in practice the true distribution may be unknown or it may be impractical to exactly solve the problem with a fixed selection vector. Finally, we see that for our problem, like other experiences with SAA, smaller sample sizes than those dictated by the theory are sufficient to get excellent performance.

3.6 Concluding remarks

This chapter proposes and analyses approximation algorithms to solve a wide class of selection problems with applications in supply chain management, resource allocation, and other fields. Our approach consists of a novel combination of rounding and sampling for which we can derive explicit performance guarantees. The approximation techniques that we employ are able to handle a class of risk measures which includes expected value, CVaR, and other coherent risk measures. We show empirically that our approximation algorithms perform very well even for small sample sizes for an example within our problem class. These results are encouraging since in practice the cardinality of these problems may be very large making them computationally intractable.

In the context of this thesis, both this chapter and Chapter 2 consider specific generalizations of the selective newsvendor problem. In Chapter 2 we extended the SNP to a class of problems which included a multiperiod version of the problem presented in Taaffe et al. [15] as a special case. In this chapter we presented a more general selection problem and a larger variety of demand distributions as well as several classes of approximation algorithms for solving the problem. The next two chapters take more novel approaches to the supplier's market selection problem.

CHAPTER 4

Efficient Multi-unit Auctions for Selling to Newsvendor-type Cost Retailers

4.1 Introduction

In this chapter we take a very different approach than the previous two chapters. While before we considered a supplier who selected a subset of all possible markets, here we study a supplier with a fixed and limited quantity to sell to a set of retailer-customers. We assume that the retailers have newsvendor-type costs and participate in a supplier-run auction to order the good. Using an auction to sell a limited resource is not a new concept. However, those buying the goods are typically the end users. In the case where bidders are retailers and are attempting to satisfy possibly unknown demand from their own customers, the amount the retailer wants to order will depend directly on the price they must pay. This trade-off between the amount desired and the price each retailer must pay provides motivation to use auctions to allocate the limited inventory.

In this chapter we consider only a supplier maximizing the efficiency of the allocation. A maximum-efficiency auction ensures that those who have the highest value for the goods, receive them. This kind of objective may be desirable when the supplier represents a government entity or as a means to ensure that the resulting allocation maximizes total welfare. For example, departments of natural resources across the US use auctions to allocate timber rights. Similarly, a supplier who has experienced a supply disruption could use an efficient auction to decide how best to allocate their limited inventory. In settings like these, running an efficient auction may be one way to still receive the benefit of information revelation in auctions without alienating customers by being profit-focused.

The efficient auction we present here is modeled after the work of Ausubel [7] which presents a multi-unit efficient auction. This auction is the multi-unit equivalent to the classic Vickrey auction which allocates a single unit by auction to the highest-value bidder at the price of the second-highest bidder's valuation. In Ausubel and Cramton [52], the authors showed that typically a multi-unit auction with all objects sold at the clearing price results in an inefficient allocation which motivates the slightly more complicated auction presented in [7]. In addition to devising efficient discrete and continuous auctions for the supply-chain setting, this chapter also introduces an approximately efficient auction. Lehmann et al. [53] cover an approximately efficient combinatorial auction where the bidders approximately solve for the right bids. In our setting, it is easy for the bidders to determine the right bid for any particular bid price, but participating in the full auction may involve excessive computations. In this chapter, we consider a subset of all desirable prices that maintains a certain level of efficiency.

While this chapter focuses on the efficient auction, future research may consider a revenue-maximizing objective function for the supplier or instead specify practically relevant auction mechanisms. Zhan [54] does a review of the efficient and optimal auction literature. Maskin and Riley [6] develop an optimal multi-unit auction with multi-unit demands. Deshpande and Schwarz [55] apply that work to a supply-chain setting similar to the one we study here. In addition, Kastl [56] covers the discriminatory and uniform price auctions which are well-studied as they relate to treasury bond auctions.

In the next section, we introduce the basic model. We then present the discrete, continuous, and ϵ-efficient auctions. Finally, we conclude and discuss future research directions.

4.2 Model

A supplier has a fixed capacity of W units to allocate among m retailers with $M \equiv \{1, 2, \ldots, m\}$. Each retailer i can be assigned any quantity $Q \geq 0$ such that total allocations do not exceed the capacity W. Each retailer i's expected profit is quasilinear and equal to $\Pi_i(Q) = V_i(Q) - p_i$, with $V_i(Q)$ as the value function which gives the expected value of receiving Q units, and p_i as the procurement cost paid to the supplier.

We choose $V_i(Q)$ to be defined as the expected net retail-side profits of retailer i (not including the procurement costs). We consider retailers who have income of $r_i \geq 0$ for each unit sold and then pay newsvendor-type costs depending on how the realized demand compares to the allocated quantity Q. Specifically, let D_i be the random demand of retailer i and Q be an allocated quantity. Similarly to the earlier chapters retailer i will have

$$V_i(Q) = E[r_i \min\{D_i, Q\} - e'_i(D_i - Q)^+ + v_i(Q - D_i)^+]$$

where e'_i is a penalty cost per unit of unmet demand, v_i is the salvage value for any excess units, and $(y)^+ = \max\{0, y\}$. $D_i \sim H_i$ is chosen as a random variable which has nonnegative support and has pdf h_i where the D_i are independent of each other. We assume that H_i is continuous and privately known by retailer i and satisfies the conditions presented in Ausubel [7] which prevent the other retailers from guessing the private information based on bids (we will be more specific about these assumptions later). We also note that while the description above is for the lost-sales version of the newsvendor problem, the value function can be rearranged by setting $e_i = r_i + e'_i$ to have the form:

$$V_i(Q) = E[r_i D_i - e_i(D_i - Q)^+ + v_i(Q - D_i)^+].$$

The supplier is interested in developing an auction which allocates the limited supply in a useful way. In this chapter we specifically consider a supplier interested in maximizing retailer efficiency. However, maximizing supplier profits is also a useful criterion which could be studied in future work. The supplier's problem is to decide on a selling mechanism $(\mathcal{B}, \mathcal{A}, \mathcal{P})$ made up of the set of *feasible bids* $\mathcal{B}$ for the collection of retailers where $\mathcal{B} \in \beta$ and β is the set of all possible definitions of feasible bids, an *allocation rule* $\mathcal{A} : \mathcal{B} \to \mathbb{R}^n_+$, and a *payment rule* $\mathcal{P} : \mathcal{B} \to \mathbb{R}^n$. The choice of the set of feasible bids defines how bidding occurs. The allocation rule $\mathcal{A} \in A$ determines, as a function of all submitted bids, the allocated units x_i to each retailer i. Finally, the payment rule $\mathcal{P} \in P$ determines, as a function of all submitted bids, the amount each retailer i will pay p_i for their received quantity.

An efficient auction allocates the W available units in a way that maximizes total expected retailer profits, where any unsold items provide zero value to the supplier. The supplier's objective is then to choose an efficient auction mechanism ($\mathcal{B}^E, \mathcal{A}^E, \mathcal{P}^E$ where E denotes efficient) which optimizes the following problem:

$$\begin{aligned} \max_{(\mathcal{B}^E,\mathcal{A}^E,\mathcal{P}^E)\in(\beta,A,P)} \quad & \sum_{i=1}^{n} E[V_i(x_i)] \\ \text{s.t.} \quad & x = \mathcal{A}^E(B). \end{aligned} \tag{PE}$$

And the retailer's objective is:

$$\begin{aligned} \max_{B_i\in\mathcal{B}^E} \quad & E[V_i(x_i)] - p_i \\ \text{s.t.} \quad & x = \mathcal{A}^E(B) \\ & p = \mathcal{P}^E(B) \\ & B \in \mathcal{B}^E. \end{aligned} \tag{PR}$$

We will use the auction mechanism presented in Ausubel [7] with minor modifications to devise an efficient auction for our setting. In that paper, the marginal values are assumed to be discrete and to come from a publicly known set. That structure allows the auctioneer to consider a discrete set of prices. In addition, the bidders are then indifferent to receiving the unit or not when paying their marginal value for that unit. In this chapter, the retailers do not have a discrete or finite set of marginal values. We address this issues in two ways. In Section 4.3 we present a finite efficient auction by considering a finite set of candidate demand distributions, and then consider a carefully constructed set of marginal values for the ascending-price auction. These modifications allow the original discrete mechanism from Ausubel [7] to be implemented with no additional modifications. In Section 4.4 we present an ascending price auction with a continuous set of prices which allows a continuous version of the mechanism from Ausubel [7] to result in an efficient allocation.

While the above approaches allow us to use the auction mechanism from Ausubel [7], the author's specific results regarding information no longer hold. The author is able to demonstrate that private values ensure bidders submit truthful bids. This works since there is enough uncertainty to iteratively eliminate other possible bids. In this chapter, the same approach no longer holds since the auctioneer needs full demand distribution information to construct the set of auction prices. Depending on the specifics of the problem instance, it may be possible to allow the bidders to observe the bids or allocation quantities while still maintaining an efficient auction. However, for the purposes of this chapter we will simply assume that while the candidate demand distributions are publicly known, the entire bid history and allocations are not known until the auction has terminated.

4.3 Finite efficient auction

We implement a simplified version of the discrete Ausubel auction mechanism from [7]. This is an ascending-price auction where the actual price a retailer pays for marginal unit allocations are determined by the bids of the other retailers alone. Because of certain modeling choices of ours, the simplified version of the algorithm is still guaranteed to result in an efficient allocation.

The outline of the auction is that the supplier announces a finite set of per-unit prices $\mathcal{C}$

that will be considered in ascending order. At each price, each retailer submits a bid of an order quantity. This continues until the total orders are no more than the capacity, at which point allocations and payments are finalized according to a special rule. We construct a set $\mathcal{C}$ to ensure that the resulting auction is efficient. We also note here that if W is so large as to exceed the sum of all demand, any excess units will have zero marginal value and the resulting allocation will still be efficient.

Let H_i for each retailer i come from a finite set of ℓ_i candidate CDFs where each CDF is denoted by $H_{i,k}$ and $k \in \{1, 2, \ldots, \ell_i\}$. We then assume that each $H_{i,k}^{-1}(\rho)$ which is the inverse CDF of the ρ fractile both exists and is continuous in ρ. For notational simplicity, we let $H_i^{-1}(\rho)$ denote the true inverse CDF for retailer i. We also assume that each retailer i only knows their own set of candidate demand distributions $H_{i,k}$ for $k = 1, 2, \ldots, \ell_i$.

We now define the efficient price c^E in the following lemma:

Lemma 4.3.1. *Let each retailer i have newsvendor-type costs of $V_i(Q)$ and one of the candidate inverse CDFs denoted $H_i^{-1}(\rho)$, which are all assumed to be continuous in ρ. Then there exists a per-unit price $c^E > 0$ such that each retailer will optimally choose an order quantity of $H_i^{-1}\left(\frac{e_i - c^E}{e_i - v_i}\right)$ and such that*

$$\sum_{i=1}^{n} H_i^{-1}\left(\frac{e_i - c^E}{e_i - v_i}\right) = W.$$

Proof. The retailer's expected profit function when the amount paid is a set per-unit price c is:

$$\Pi_i(Q) = E[r_i D_i - e_i(D_i - Q)^+ + v_i(Q - D_i)^+] - cQ.$$

This is exactly the newsvendor problem with expediting costs. The well known optimal order quantity for the retailer is then the critical-fractile of the demand distribution, namely $Q^* = H_i^{-1}\left(\frac{e_i - c}{e_i - v_i}\right)$.

Since the $H_i^{-1}(\rho)$ are continuous in ρ, the efficient price c^E is guaranteed to exist. Therefore, if the supplier were to offer the good at a per-unit price of c^E, each retailer would optimally order Q^* units which then sum exactly to W. $\square$

This lemma suggests that if the supplier could a priori select the per-unit price c^E, they could consider only that price while still efficiently allocating the W units since the retailers have decreasing marginal value for the goods. However, since the supplier only knows the set of possible retailer demand distributions, they must include prices in $\mathcal{C}$ other than c^E. One might think that the solution would be a uniform-price auction where bidders pay the same price on all units they are allocated, and that price is simply c from the final auction round. However, as was demonstrated in Ausubel and Cramton [52], all objects simply being sold at the clearing price does not result in an efficient allocation.

Using the allocation and payment rules we discuss below ensures an efficient allocation as long as c^E is included in the set of considered prices. In this section we construct the smallest set of prices $\mathcal{C}$ that the supplier can guarantee will include c^E based on the finite set of candidate CDFs H_i for each retailer i. The difference between the amount paid to the supplier if $\mathcal{C} = c^E$ and the constructed $\mathcal{C}$ is known as the price of information in this auction. I.e., the amount the supplier must pay for the retailers to reveal their true value functions. Because we consider only an efficient auction, the total profit of the system is fixed and only how that profit is split is determined by the choice of $\mathcal{C}$.

In the remainder of the section we formally define the auction mechanism and prove that it results in an efficient allocation.

4.3.1 Feasible bids $\mathcal{B}$

We construct the set $\mathcal{C}$ in the following manner. Let $\mathcal{C}$ include any price c such that for any set of $k_i \in \{1, \ldots, \ell_i\}$:

$$\sum_{i=1}^{n} H_{i,k_i}^{-1}\left(\frac{e_i - c}{e_i - v_i}\right) = W.$$

The supplier begins the auction at an initial per-unit price $c_0 > 0, \in \mathcal{C}$ which is the smallest element of $\mathcal{C}$, and then accepts a bid of a desired order quantity $b_{i,0} \geq 0$ from each retailer i. The supplier than selects the next lowest per-unit price $c_1 > c_0, \in \mathcal{C}$, and accepts a bid of a desired order quantity $b_{i,1} \leq b_{i,0}$ from each retailer i. This continues with strictly increasing and sequential per-unit prices of $c_j \in \mathcal{C}$ for any round number j. The bidding process terminates when the sum of the order quantities in a particular round J is no more than the available quantity, i.e. $\sum_{i=1}^{n} b_{i,J} \leq W$.

4.3.2 Allocation rule $\mathcal{A}$

Let $x_{i,j}$ denote the total allocated quantity to retailer i in rounds 0 through j. The initially allocated quantity to retailer i in the first round is

$$x_{i,0} = \min\left\{ b_{i,0}, \left(W - \sum_{k \neq i} b_{k,0} \right)^{+} \right\}.$$

In the subsequent rounds, a similar concept carries forward. Namely, in round $0 < j \leq J$ we have

$$x_{i,j} - x_{i,j-1} = \min\left\{ (b_{i,j} - x_{i,j-1})^{+}, \left(\left(W - \sum_{k \neq i} b_{k,j} \right)^{+} - x_{i,j-1} \right)^{+} \right\}.$$

Ausubel describes this as each retailer "clinching" quantity as the demand from the other retailers decreases below the currently available quantity. Because of the structure of the auction, any clinched units are guaranteed to the retailers in future round.

4.3.3 Payment rule $\mathcal{P}$

The price at which a quantity is clinched is the price the retailer must pay for that quantity. Therefore, each retailer will pay to the supplier the following amount:

$$p_i = \sum_{j=1}^{J} c_j (x_{i,j} - x_{i,j-1}).$$

Any units which are allocated in the first round are also sold for c_0 per unit, which allows each retailer to bid independently of the other retailers. We could alternatively add a dummy first round for which no units are clinched, which would also preserve the retailer motivation to bid according to the same strategy in all periods.

4.3.4 Efficiency of the multi-unit auction

The auction mechanism described above is a simplified version of the mechanism originally presented in Ausubel [7]. Our work is distinct in that we consider continuous order quantities while [7] only considers discrete quantities. Further, through the specific choice of the set $\mathcal{C}$ presented above and the assumption that the $H_{i,k}^{-1}(\rho)$ are continuous in ρ, we can guarantee that for any possible collection of the private information values $k_i \in \ell_i$ there exists a price $c \in \mathcal{C}$ such that the total orders when retailers have newsvendor-type costs is exactly W units.

Outside of the auction framework, the fact that the retailers simply order their newsvendor quantity is not surprising. As we will see in the following theorem, following the same strategy during the auction for each price $c \in \mathcal{C}$ remains an optimal choice for each retailer. This is an interesting result since according to the payment rule, the retailers will generally only clinch a small number of units in the final round. Practically, this means they will order the newsvendor order quantity even though they are paying less than the final per-unit price for most of their allocation. However, while the retailer saved money early in the auction, the newsvendor order quantity balances the marginal utility for the next unit with the given per-unit prices. Effectively, the auction provides a way for the supplier to identify the efficient price c^E from Lemma 4.3.1. We now present the theorem.

Theorem 4.3.2. *Consider retailers with newsvendor-type costs who have no knowledge of the other retailer's valuations other than their previous bids. Assuming the retailers do not collaborate, it is an optimal strategy for each retailer to choose a bid of* $b_{i,j} = H_i^{-1}\left(\frac{e_i - c_j}{e_i - v_i}\right)$ *in round* j *which will result in an efficient allocation of the capacity.*

Proof. Suppose that retailer i has clinched a cumulative quantity of $\tilde{Q}_i^j$ in rounds $j = 1, ..., \ell - 1$ (for $i = 1, ..., m$). In round ℓ, retailer i's expected profit if retailer i were to receive the maximum of their previous clinched quantity and their bid in this round of Q is given by

$$\Pi_i(Q) = E[r_i D_i - e_i(D_i - \max\{Q, \tilde{Q}_i^{\ell-1}\})^+ + v_i(\max\{Q, \tilde{Q}_i^{\ell-1}\} - D_i)^+] - \sum_{j=1}^{\ell} c_j(\tilde{Q}_i^j - \tilde{Q}_i^{j-1})^+$$

where $\tilde{Q}_i^\ell = Q$.

Now let $(Q_i^\ell)^* = H_i^{-1}((e_i - c_\ell)(e_i - v_i))$ since retailer i faces newsvendor-type costs and a per-unit purchase price of c_ℓ in the current round.

CASE A. $(Q_i^\ell)^* <= \tilde{Q}_i^{\ell-1}$: Since $(Q_i^\ell)^*$ is decreasing in ℓ, this will also be the case in subsequent rounds. This means that $\tilde{Q}_i^{\ell-1}$ maximizes the expected profit $\Pi_i(Q)$. Bidding $\tilde{Q}_i^{\ell-1}$ yields that $\tilde{Q}_i^\ell = \tilde{Q}_i^{\ell-1}$ (so no additional items will be clinched). The only feasible alternative is to bid a quantity strictly larger than $\tilde{Q}_i^{\ell-1}$. But this makes the expected profit lower, and (by repeating this argument) can only make future expected profits lower (or equal). So it is not beneficial for the retailer to bid a quantity more than $\tilde{Q}_i^{\ell-1}$.

CASE B. $(Q_i^\ell)^* > \tilde{Q}_i^{\ell-1}$: Bidding $(Q_i^\ell)^*$ may result in the clinching of additional items $\tilde{Q}_i^\ell >= \tilde{Q}_i^{\ell-1}$. Now there are two alternatives:

i) suppose that retailer i bids a quantity strictly smaller than $(Q_i^\ell)^*$. Since the amount clinched can easily be seen to be nondecreasing in the amount bid, we know that in this case retailer i will not clinch more than their current bid of Q. Therefore, the total expected profit is either identical to or worse than when the retailer bids $(Q_i^\ell)^*$ (since the immediate expected profit cannot be higher, and any additional items purchased will be more costly).

ii) suppose that retailer i bids a quantity strictly larger than $(Q_i^\ell)^*$. As a result, retailer i will either clinch up to $(Q_i^\ell)^*$ (i.e., the same as if the retailer would bid $(Q_i^\ell)^*$ to begin with) or a larger quantity. In the former case, there is no benefit to bidding the larger amount. In the latter case since $\tilde{Q}_i^\ell > (Q_i^\ell)^*$ then the immediate expected profit will be lower, and it will also be undesirable to clinch additional items in future rounds.

In addition, since retailers only know the prior bids and cannot collaborate, they are unable to predict when the auction will terminate. This eliminates the opportunity to reduce their bids in an attempt to receive slightly fewer units at a slightly lower per-unit price. Therefore, it is always optimal for retailer i to bid $(Q_i^\ell)^*$.

From Lemma 4.3.1 the efficient price c^E will be offered, and as a result of the above bidding strategy, the auction will terminate exactly when that price is offered. Since the marginal utility is decreasing in the received quantity for each retailer, we can also see that no retailer would in fact find themselves with $\tilde{Q}_i^j < Q_i^*$ for any round j. Therefore this allocation is also efficient.

□

As we mentioned in Section 4.2, these results depend in part on relatively strong information assumptions. Ausubel [7] provides much milder conditions because the marginal values are from a discrete set. Depending on the strategic behavior of the retailers, those milder conditions may also hold in our setting.

4.4 Continuous efficient auction

In the previous section, we assumed that the demand distribution H_i was privately known but the supplier knew of a finite set ℓ_i of candidate CDF functions for each retailer i. This allowed us to consider only a finite set of possible prices $\mathcal{C}$. However, in some cases the private information may not have a finite number of values, or it may be challenging to compute the set $\mathcal{C}$ (either due to the cardinality of the set, or the complexity of finding the price c^E for a particular set of candidate distributions). In that case, we implement a variation of the continuous Ausubel auction mechanism from Ausubel [7]. Since we continue to suppose that retailer valuations and order quantities are continuous, some of the complexity of the continuous version of the Ausubel auction can be eliminated.

Let H_i for each retailer i come from a non-finite set of possible CDFs which are distinguished by a real parameter k. We then assume that $H_{i,k}^{-1}(\rho)$ which is the inverse CDF of the ρ fractile both exists and is continuous in ρ. We continue to assume that the individual retailers only know their own set of possible CDFs. In addition, Lemma 4.3.1 continues to hold since it only depends on the underlying newsvendor-type costs and not the private information.

4.4.1 Feasible bids $\mathcal{B}$

While in the previous section we constructed a special set of prices, in this section we can simplify the choice of $\mathcal{C}$ by starting at a price of $\max_i\{v_i\}$ and increasing the price continuously until exactly W units are ordered, which is bounded above by $\max_i\{e_i\}$. If the supplier knows the candidate demand distributions of the retailers, we could use the same approach as before to find the smallest possible (but now continuous) set $\mathcal{C}$ which includes all possible efficient prices c^E.

For notational simplicity, suppose the supplier begins the auction at a lowest per-unit price of $c_0 = \max_i\{v_i\}$, and accepts a bid of a desired order quantity $b_{i,c_0} \geq 0$ from each retailer i. The supplier then continuously increases the per-unit price c and $b_{i,c}$ is retailer i's bid at price c. At all per-unit prices between c_0 and when the auction terminates, retailer i must choose their desired order quantity $b_{i,c}$ where $b_{i,c}$ is nonincreasing in c. The bidding process terminates at the first price c^E such that total demand is no more than the capacity, i.e., $\sum_{i=1}^{m} b_{i,c^E} \leq W$.

4.4.2 Allocation rule $\mathcal{A}$

Let $x_{i,c}$ denote the allocated quantity to retailer i through any price $c \geq c_0$. We also make the following assumption:

Assumption 4.4.1. *Retailers must satisfy the activity rule*

$$b_{i,c} \geq x_{i,c'}$$

for any $c \geq c'$.

We could instead define the auction as we did in Section 4.3 such that clinched quantities never decrease. However, in this section, it greatly simplifies the later definitions to include the assumption above that retailers cannot bid for a lower quantity than they have already been allocated.

The allocated quantity to retailer i at any price $c \geq c_0$ is then defined as:

$$x_{i,c} = \min \left\{ b_{i,c}, \left(W - \sum_{k \neq i} b_{k,c} \right)^+ \right\}.$$

This is simply a continuous version of "clinching" as described in the previous section. The auction rules continue to state that any clinched units are guaranteed to the retailers in future rounds.

4.4.3 Payment rule $\mathcal{P}$

The price at which a quantity is clinched is the price the retailer must pay for that quantity. Therefore, each retailer will pay to the supplier the following amount:

$$p_i = \int_{c=c_0}^{c^E} c \left(\frac{dx_{i,c}}{dc} \right) dc$$

where c^E is the last considered price (which is not known a priori). The payment rule also requires that any units which are allocated in the first round are sold for c_0 per unit.

4.4.4 Efficiency of the multi-unit auction

We begin with the following strategy for each retailer.

Definition 4.4.2. *We let the sincere order of bidder i at a per-unit price of $c \in \mathcal{C}$ be*

$$H_i^{-1}\left(\frac{e_i - c}{e_i - v_i}\right).$$

Next we identify conditions such that the allocation and payments exist and are computable.

Assumption 4.4.3. *Suppose all retailers make sincere bids at every price $c \in [0, \bar{c}]$. For each retailer i and the true inverse CDFs, the following function only has finitely many zeros.*

$$g(i,c) = H_i^{-1}\left(\frac{e_i - c}{e_i - v_i}\right) - \left(W - \sum_{k \neq i} H_k^{-1}\left(\frac{e_k - c}{e_k - v_k}\right)\right)^+.$$

This is a more restrictive assumption than necessary as all we really need is that $x_{i,c}$ is differentiable at all but finitely many discontinuities when the retailers make sincere bids in the range $\left[c_0, c^E\right]$. In that case, both the allocated quantity and the payments are computable since the inverse CDFs are continuous and the derivative exists and is finite for all but finitely many points.

Lemma 4.4.4. *Consider retailers with newsvendor-type costs who have no knowledge of the other retailer's valuations other than their previous bids. Assuming the retailers do not collaborate and Assumptions 4.4.1 and 4.4.3 hold, for the ascending-bid auction for our model, sincere bidding is an ex post optimal strategy.*

The proof follows equivalently to the finite case since Assumption 4.4.3 ensures that all the relevant quantities exist and are computable. Based on this result we can present the theorem for the continuous auction case:

Theorem 4.4.5. *Consider retailers with newsvendor-type costs who have no knowledge of the other retailers valuations other than their previous bids. Assuming the retailers do not collaborate and Assumptions 4.4.1 and 4.4.3 hold, for the ascending-bid auction for our model, the allocation is efficient.*

Proof. Since the allocation is ultimately just the order quantity at the price c^E when demand is exactly W, this is an efficient allocation. □

This demonstrates that with only small modifications to our assumptions, we are able to generalize the auction rules to the continuous case. However, in most cases such an auction would not be implementable as the retailers would effectively need to submit their complete demand curve. This reality motivates the remaining section.

4.5 Finite ϵ-efficient auction

In the previous sections we presented an auction mechanism which is efficient if $\mathcal{C}$ is chosen correctly. In general, the supplier is free to choose $\mathcal{C}$ in any way they wish, and is therefore not guaranteed to have an efficient allocation. For example, if the supplier chose $\mathcal{C} = \{\ \alpha\}$ where $0 > \alpha > \min_i v_i$, each retailer i would choose $b_{i,\alpha} \geq W$, which provides the supplier with no information to base an efficient allocation on. Further, which choices of $\mathcal{C}$ will lead to an efficient allocation will depend on the retailers' incentives.

As was presented before, for the discrete auction to be guaranteed to be efficient, the set $\mathcal{C}$ should includes any candidate clearing price if the retailers bid their newsvendor quantities. If the demand distribution for any retailer is continuous, now all potential prices up to the maximum realization of demand to that retailer are distinct. Even if the demands are discrete, the set $\mathcal{C}$ may contain a large number of elements when combined with the retailer-specific parameters r_i, e_i, and v_i.

Rather than considering the full set of values, the supplier could select a collection of prices $\mathcal{C}$ that allows retailers to reveal enough of their private information to approximate a fully efficient auction. For example, the supplier may consider ten-cent or ten-dollar intervals, depending on the scale of prices being considered. There has been some work on costly auctions such as Parkes [57], however they do not provide bounds on the result of the auction. Another justification for the use of a discrete set $\mathcal{C}$ is that the retailers may only be willing to submit a limited number of quantity bids. Kastl [56] discusses the practical constraint that bidders will not submit a large number of bids but will instead approximate their needs with a limited number of bids. While the supplier might want a completely efficient auction, the retailers may decide to submit the same bid for several neighboring prices in order to save computational effort. Since the supplier chooses the set $\mathcal{C}$ a priori, the retailers can decide how to respond based on the supplier's choice. In addition to this auction related literature, there are also papers on communication complexity in general (e.g., [58]).

In order to present an ϵ-efficient auction, we use the same auction rules as in Section 4.3 for the discrete auction. Since c^E might not be included in the set $\mathcal{C}$, the final round may have total demand strictly less than the available units W and therefore the auction may not result in an efficient allocation. We begin this section by observing that the optimal retailer strategy continues to be making sincere bids.

Lemma 4.5.1. *Consider retailers with newsvendor-type costs who have no knowledge of the other retailers valuations other than their previous bids. Assuming the retailers do not collaborate and the supplier has chosen a set of prices $\mathcal{C}$, it is an optimal strategy for each retailer to choose a bid of $b_{i,j} = H_i^{-1}\left(\frac{e_i - c_j}{e_i - v_i}\right)$ for each $c_j \in \mathcal{C}$.*

Proof. This follows immediately from the proof of Theorem 4.3.2. □

Given the auction rules and bidding strategy above, we want to connect the choice of $\mathcal{C}$ to a bound on the deviation from an efficient allocation. Guided by this goal, we present a special case by assuming each retailer i has uncertain demand which comes from a uniform distribution on $[\gamma_i, \xi_i]$, where γ_i and ξ_i are privately known but are bounded by upper and lower bounds $\gamma_i \in \left[\underline{\gamma}_i, \bar{\gamma}_i\right]$ and $\xi_i \in \left[\underline{\xi}_i, \bar{\xi}_i\right]$. If the other retailers could see the bid history, the retailers could use that information to improve their expected return. Therefore, we make the restrictive assumption that the bids and allocations are confidential to the supplier until the auction terminates. We present this case because it demonstrates the relationships between the key parameters and ϵ. Further, if a supplier did in fact have retailers with uniformly distributed demands and wanted to run an ϵ-efficient auction, they could choose to run a sealed bid auction for the set of prices $\mathcal{C}$.

We now suppose that $\mathcal{C}$ is simply defined by an initial lowest price c_0, and then have each subsequent price increment by Δc. Given this initial naive selection of $\mathcal{C}$, we identify an issue in the following lemma.

Lemma 4.5.2. *Consider a supplier with W units running an auction with $\mathcal{C} = \{c_0, c_0 + \Delta c, c_0 + 2\Delta c, \ldots\}$ and let the bidders be retailers with newsvendor-type costs e_i and v_i and facing demand which is uncertain but uniformly distributed on the privately known interval $[\gamma_i, \xi_i]$ with publicly known bounds $\gamma_i \in \left[\underline{\gamma}_i, \bar{\gamma}_i\right]$ and $\xi_i \in \left[\underline{\xi}_i, \bar{\xi}_i\right]$. Then, the maximum possible change in quantity ordered ΔQ in a single step of the auction is at least $\underline{\gamma}_i$ if $c^E > e_i$ for any retailer i.*

Proof. From Lemma 4.5.1, each retailer simply submits a bid of $b_{i,j} = H_i^{-1}\left(\frac{e_i - c_j}{e_i - v_i}\right)$ for each $c_j \in \mathcal{C}$ until the auction terminates. Consider the retailer i' for which $c^E > e_{i'}$. Let j be the highest price c_j which is less than $e_{i'}$. Then, retailer i' will have a drop in their bid from some $b_{i',j} \geq \underline{\gamma}_{i'}$ to a bid $b_{i',j+1} = 0$. $\square$

This lemma demonstrates that if retailers have a demand distribution which is bounded below but by a strictly positive number then a simple grid of prices may lead to large discontinuities. One solution then is to simply include the publicly known e_i values in the set $\mathcal{C}$. Since the retailers are indifferent between receiving γ_i and zero when the price is e_i, the supplier may be able to use the bid information at e_i to produce an exactly efficient allocation.

For simplicity we now assume that for all retailers i, $\gamma_i = 0$. Given this setting, we present the following theorem:

Theorem 4.5.3. *Consider a supplier with W units running an auction with $\mathcal{C} = \{c_0, c_0 + \Delta c, c_0 + 2\Delta c, \ldots\}$ and let the bidders be retailers with newsvendor-type costs e_i and v_i and facing demand which is uncertain but uniformly distributed on the privately known interval $[0, \xi_i]$ with publicly known bounds $\xi_i \in \left[\underline{\xi}_i, \bar{\xi}_i\right]$. If $\Delta c \leq \epsilon \sum_{i=1}^{m} \frac{(e_i - v_i)}{e_i \underline{\xi}_i}$, the resulting auction will be ϵ-efficient.*

Proof. For each retailer i, $h_i = \frac{1}{\xi_i}$. Therefore, for a step in price of Δc, each retailer will decrease their order quantity by $(\xi_i)\frac{\Delta c}{e_i - v_i}$. Summing over the m retailers we then have

$$\Delta Q \leq \sum_{i=1}^{m} (-\underline{\xi}_i) \frac{\Delta c}{e_i - v_i}.$$

Since each retailer i's allocation has per unit value at most e_i, we then have that the gap from the efficient allocation is at most $\sum_{i=1}^{m} e_i(-\underline{\xi}_i)\frac{\Delta c}{e_i - v_i}$. Solving for Δc with ϵ as the maximum gap:

$$\Delta c \leq \epsilon \sum_{i=1}^{m} \frac{(e_i - v_i)}{e_i \underline{\xi}_i}.$$

□

This result is informative as it demonstrates that as h_i decreases in the relevant region, the jumps between bid quantities and therefore maximum loss in efficiency, grows.

4.6 Concluding remarks

In this chapter we consider a supplier selling a fixed quantity of units to a set of newsvendor-type retailers. We saw that when the supplier chooses to implement an efficient or ϵ-efficient auction, the retailers have a straightforward strategy of choosing their truthful order which is simply the well known newsvendor critical fractile. In the future, we believe other objectives or auction mechanisms can be used in a similar setting and for other supply chain problems. Multi-unit auctions are a powerful tool for matching supply and demand. In addition, we hope others find the definition of an ϵ-efficient auction useful as it allows the decision maker to relate a discretization in considered prices to the resulting loss in efficiency. This chapter provides an alternative way of thinking about the underlying market selection problem discussed in the previous two chapters. In the next chapter we consider another model and one more perspective suppliers may consider.

CHAPTER 5

The Outsourcing Newsvendor Problem

5.1 Introduction

An important issue for multinational corporations is how to allocate resources across international boundaries. In Chapters 2 and 3 we considered the selective newsvendor problem which helps to answer these kinds of questions by identifying a desirable subset of all candidate markets to choose to serve. In this chapter we build on that problem by considering an additional set of subsidiary suppliers who can work with Supplier 1 to meet demand in the full set of markets. We call this problem the outsourcing newsvendor problem (ONP) and consider how the opportunity to work with subsidiaries affects Supplier 1. At a high level, solving both the SNP and the equivalent ONP can provide managerial insights into the structure of market profitability and production costs.

A car manufacturer which produces cars in several countries faces a difficult challenge when trying to match the (candidate) set of production facilities with a set of markets with uncertain demand. Generally, multi-national companies operate with a headquarters which makes overall strategic decisions while branches of the company in specific regions will make more local decisions. How to appropriately model these settings is a challenging problem as each company is different in how they handle decision making between the parent company and subsidiaries. In this chapter we provide one reasonable relationship between the companies. There are a number of empirical papers and discussions of these kinds of problems. For example, Kenyon et al. [59] use empirical data to study the effects on the firm and customers of a company outsourcing production. While their paper considers the more traditional interpretation of outsourcing to an external company, we are more interested in working with subsidiaries. Prahalad and Doz [60] consider the underlying idea that a supplier should consider both the global and subsidiary-level problems, which includes our case. While their work provides an overview of things to consider when making these kinds of decisions, there are limited papers which look at outsourcing decisions as an optimization problem.

De Kok [61] consider outsourcing as a tool to accomodate supplier capacity constraints when faced with deterministic demand. Nazari-Shirkouhi et al. [62] consider a multi-period, multi-product outsourcing problem constrained resources, centralized decision making, continuous quantities, also with deterministic demand. Chen and Xiao [25] consider outsourcing as a response to disruption risk and uncertain capacity. The authors then use a meta-heuristic to provide high-quality solutions to their outsourcing problem. Kermani et al. [63] also use a meta-heuristic to look at supplier selection in a competitive environment while considering price, quality, and delivery performance. One other related set of papers includes those considering make-or-buy decisions such as Vrat [64].

In this chapter, we consider a supplier with a collection of markets and uncertain demand in those markets. As can be seen from above, most if not all of the existing literature on outsourcing assumes deterministic demands. We further suppose that the supplier has a number of subsidiaries who can be allocated any subset of the markets to serve. Each supplier / each subsidiary faces newsvendor-type costs depending on the uncertain demand for their market allocation as well as their own procurement decision. Because of the parent-subsidiary relationship, we also require that each subsidiary should have nonnegative expected profit through the use of a transfer payment from the parent company if necessary. This basic structure provides a starting point for understanding multi-national companies that outsource to subsidiaries. It also complements the existing production planning literature which considers the centralized or decentralized models of coordination.

In the next section we present our model formulation. In Section 5.3 we present structure to the solutions of the ONP and also make some observations about the optimal solution for both the outsourcing problem and the SNP. In Section 5.4 we then present computational experiments and finally conclude.

5.2 Problem formulation

Consider a stochastic market allocation problem as before with m markets, one primary supplier, and a new set of $S-1$ subsidiary suppliers. We let Supplier 1 denote the primary supplier who is also the decision maker for the market allocation problem. The remaining $S-1$ subsidiaries are then allocated a set of markets to serve by Supplier 1. Recall the market selection decisions $z_{ij} \in \{0,1\}^{mS}$ with $z_{ij} = 1$ denoting that supplier i serves market j and $z_{ij} = 0$ meaning that they do not serve that market. Each supplier i may then choose a procurement quantity as before of $y_i \geq 0$ to meet the uncertain demand in their allocation of markets. We then suppose that since Supplier 1 makes the allocation decisions, they are responsible for ensuring that each subsidiary have nonnegative expected profit by providing a transfer payment ζ_i to subsidiary i if necessary. As mentioned in the introduction, this is meant to replicate some aspects of multi-national corporations which have not been addressed in the past literature.

After the allocation decisions and procurement quantity have been determined, the demand in each market is realized. Recall that $\mathbf{D}$ denotes the random demand vector and $\mathbf{D}_j$ is the random demand in market j. We then have the previously defined newsvendor-type costs with v_i as the per-unit salvage value for leftover units, c_i the per-unit procurement cost for ordering y_i units, and e_i the per-unit expediting cost for any excess demand not met by y_i. Each unit of demand in market j is also associated with r_j revenue. The resulting optimization problem is then:

$$\max_{z,y,\zeta} \; E\left[\sum_{j=1}^{m} r_j \mathbf{D}_j z_{1j} - e_1[\mathbf{D}^\top z_1 - y_1]^+ + v_1[y_1 - \mathbf{D}^\top z_1]^+\right] - c_1 y_1 - \sum_{i=2}^{S} \zeta_i \tag{ONP}$$

$$\begin{aligned}
\text{subject to: } & \sum_{i=1}^{S} z_{ij} = 1 \qquad j \in \{1,2,\dots,m\} \\
& 0 \le E\left[\sum_{j=1}^{m} r_j \mathbf{D}_j z_{ij} - e_i[\mathbf{D}^\top z_i - y_i]^+ + v_i[y_i - \mathbf{D}^\top z_i]^+\right] \\
& \qquad\qquad + \zeta_i - c_i y_i \quad i \in \{2,\dots,S\} \\
& z_{ij} \in \{0,1\} \qquad i \in \{1,\dots,S\}, j \in \{1,\dots,m\} \\
& \zeta_i \ge 0 \qquad i \in \{1,\dots,S\} \\
& y_i \ge 0 \qquad i \in \{1,\dots,S\}
\end{aligned}$$

where $e_i \ge c_i \ge v_i \ge 0$.

This formulation is the most natural based on the previous description. We will also find it useful to consider a discretized MILP version of the problem. Suppose the demand distribution is discrete and made up of Ξ equal probability demand vectors where for each $k \in \{1,\dots,\Xi\}$ there is a demand vector $\mathbf{D}^k$ with each $\mathbf{D}_j^k$ as the demand in market j. The MILP version of the problem is then:

$$\max_{z,y,\zeta} \; \sum_{k=1}^{\Xi}\left[\sum_{j=1}^{m} r_j D_j^k z_{1j} - e_1 s_{1k} + v_1 t_{1k}\right] - c_1 y_1 - \sum_{i=2}^{S} \zeta_i \tag{OMILP}$$

$$\begin{aligned}
\text{subject to: } & \left(\sum_{j=1}^{m} D_j^k z_{ij}\right) - y_i \le s_{ik} \qquad k \in \{1,\dots,\Xi\}, i \in \{1,\dots,S\} \\
& y_i - \sum_{j=1}^{m} D_j^k z_{ij} \ge t_{ik} \qquad k \in \{1,\dots,\Xi\}, i \in \{1,\dots,S\} \\
& \zeta_i - c_i y_i + \sum_{k=1}^{\Xi}\left[\sum_{j=1}^{m} r_j D_j z_{ij} - e_i s_{ik} + v_i t_{ik}\right] \ge 0 \quad i \in \{2,3,\dots,S\}
\end{aligned}$$

$$\sum_{i=1}^{S} z_{ij} = 1 \qquad j \in \{1, 2, \ldots, m\}$$

$$z_{ij} \in \{0, 1\} \qquad i \in \{1, \ldots, S\}, j \in \{1, \ldots, m\}$$

$$\zeta_i, y_i \geq 0 \qquad i \in \{1, \ldots, S\}$$

$$s_{ik}, t_{ik} \geq 0 \qquad k \in \{1, \ldots, \Xi\}, i \in \{1, \ldots, S\}.$$

In the following section we present our results.

5.3 Outsourcing newsvendor problem results

The outsourcing newsvendor problem introduced in the previous section has similar structure to the selective newsvendor problem presented in Chapters 2 and 3. However, ONP adds additional complexity as no market can remain unassigned and the supplier needs to optimize for their own costs as well as those of the subsidiaries. We will show in section 5.3.2 that in fact the SNP can be viewed as a special case of ONP. We expect that like for the SNP, computational methods will be very effective at quickly solving (ONP). As we will show in Section 5.4, we are able to solve problems which are equivalent to those studied in Section 3.5 in reasonable time. Given these observations, the goal in the remainder of this section will be to show what can be learned about the underlying demand and procurement structure using the outsourcing model described above as well as the SNP model.

5.3.1 Solutions to the outsourcing newsvendor problem

Suppose we have solved (ONP) and found that for each supplier i the optimal allocation is z_i^* with the procurement quantity y_i^*. We now provide the following results and discussion based on the problem structure and the given optimal solution. In this section we begin with two lemmas which explore interesting features of how we model the outsourcing decision. We then present one important implication of uncertain demands and newsvendor-type costs on the optimal solution.

Lemma 5.3.1. *If $c_1, e_1,$ and v_1 are low enough that all markets are profitable, it will be optimal for Supplier 1 to choose to serve all markets regardless of the cost parameters of the subsidiaries (i.e., $z_1^* = \mathbf{1}$ and $z_{i\neq 1}^* = \mathbf{0}$ where $\mathbf{1}$ is the vector of m ones and $\mathbf{0}$ is the vector of m zeros).*

Proof. The objective of (ONP) consists of revenues and newsvendor-type costs for Supplier 1's allocation z_1, the procurement costs $c_1 y_1$ and any transfer payments to the subsidiaries ζ_i. Suppose Supplier 2 has the lowest procurement costs $e_2 = c_2 = v_2 = 0$. Any markets which are not allocated to Supplier 1, can then be assumed to be served by Supplier 2. Therefore, (ONP) reduces to:

$$\max_{z,y,\zeta} \; E\left[\sum_{j=1}^{m} r_j \mathbf{D}_j z_{1j} - e_1[\mathbf{D}^\top z_1 - y_1]^+ + v_1[y_1 - \mathbf{D}^\top z_1]^+\right] - c_1 y_1 \qquad \text{(S1)}$$

$$\text{subject to: } z_{1j} \in \{0,1\} \qquad j \in \{1,\ldots,m\}$$

$$y_1 \geq 0 \qquad .$$

This reduced problem may have $z_1^* = \mathbf{1}$. In this case, clearly any set of costs $c_2, e_2,$ and v_2 will not result in a change to z_1^* since it would at best negatively effect the expected profit. $\square$

This lemma emphasizes the fact that when the issue of outsourcing is modeled as we have in this paper, Supplier 1 is primarily interested in maximizing their own profits rather than those of their subsidiaries. Therefore, if Supplier 1 has low enough costs relative to the other suppliers, they will not choose to use the option of allocating some markets to the subsidiaries. Further, if their costs are low enough with regards to the variability in the market demands, it will be optimal for Supplier 1 to serve all markets for any set of subsidiary cost parameters.

If Supplier 1 were to solve (ONP) for a product that already has some distribution, finding that the optimal solution is to simply serve all markets themselves may inform future choices. In particular, if the expected profit to Supplier 1 is also high, it would indicate that the product in question is highly profitable including any variable demand. On the other hand, if the expected profit is low, we would instead infer that there are no subsidiaries who are well equipped to produce the good. This may motivate investment in the procurement process either for themselves or a subsidiary to improve the profitability of the product.

While Lemma 5.3.1 highlights one aspect of our choice to model the objective primarily in terms of Supplier 1's profit, the following lemma emphasizes a complementary aspect.

Lemma 5.3.2. *Supplier 1 is indifferent between any solutions to (ONP) which have the same z_1^* allocation, the associated optimal order quantity for them of y_1^* and total transfer payment $\sum_{i=2}^{S} \zeta_i^*$.*

Proof. The objective is only affected by Supplier 1's allocation z_1 and procurement quantity y_1, and how much any transfer payments may be. They are only interested in optimizing the subsidiary's situations in as much as it affects the total required transfer payment. Therefore, any feasible solution with z_1^*, y_1^* and $\sum_{i=2}^{S} \zeta_i = \sum_{i=2}^{S} \zeta_i^*$ is optimal. □

In many cases, the above lemma points to there being multiple optimal solutions, often with different suppliers meeting the demand in the markets not allocated to Supplier 1. If the transfer payment is not zero, the given optimal solution is likely to be unique as long as the suppliers have unique cost parameters. However, there still may be other solutions which are near optimal and may be beneficial for other reasons. When the transfer payment is actually zero, it is highly possible that there are in fact multiple optimal solutions, especially if the expected profit for some subsidiaries is large.

Multiple optimal or near-optimal solutions are important in practice since there are often external factors which affect market allocations. For example, Supplier 1 may have a favorite subsidiary because of quality concerns or capacity constraints which are not presently modeled. After solving ONP, Supplier 1 has a good idea of what z_1^* should be, and that any allocation of the remaining markets that the subsidiaries will accept for the same total transfer payments, is also optimal.

While the previous lemmas discuss interesting features of how we model the outsourcing decision in this paper, the following lemma demonstrates the effect of newsvendor-type costs on the structure of optimal solutions. In this lemma we assume as in Chapter 2 that demands are normally distributed with mean μ_j and variance s_j in market j.

Lemma 5.3.3. *Consider (ONP) with 2 suppliers and m markets. Suppose the market demands are independent and normally distributed (i.e., for each market j we have $D_j \sim \mathcal{N}(\mu_j, \sqrt{s_j})$). Let z^*, ζ_2^* be the optimal solution to ONP and assume that for the particular problem instance $z_1^* \neq \mathbf{1}$ and $\zeta_2 > 0$.*

If an addiitional market $m+1$ must be allocated to one of the suppliers without otherwise changing z and we have

$$\zeta_2 > r_{m+1}D_{m+1} - K_1\left(\sqrt{s^\top z_1 + s_{m+1}} - \sqrt{s^\top z_1}\right)$$

then the market will optimally be assigned to Supplier 2.

Proof. We begin by referring to Chapter 2 and note that for a particular market selection z_i, the optimal expected profit for Supplier i is:

$$\sum_{j=1}^{m} r_j D_j z_{ij} - K_i\sqrt{s^\top z_i},$$

where s is the vector of demand variance and $K_i = (c_i - v_i)\Phi^{-1}(\rho_i) + (e_i - v_i)L(\Phi^{-1}(\rho_i)$, and $\rho_i = (e_i - c_i)/(e_i - v_i)$.

If Supplier 1 serves market $m+1$, the change to their expected profit will be

$$r_{m+1}D_{m+1} - K_1\left(\sqrt{s^\top z_1 + s_{m+1}} - \sqrt{s^\top z_1}\right).$$

If Supplier 2 serves market $m+1$, Supplier 1 may have a lower transfer payment ζ_2.

Since $z_1^* \neq \mathbf{1}$, we know that

$$K_2\sqrt{s^\top z_2} < K_1\left(\sqrt{s^\top \mathbf{1}} - \sqrt{s^\top z_1}\right).$$

This ensures that Supplier 2 will have lower newsvendor-type costs for serving market $m+1$ than Supplier 1. Therefore, the only case where Supplier 1 would be better off serving market $m+1$ is if the expected profit from adding that market to their allocation is larger than the transfer payment ζ_2. $\square$

While this lemma was written assuming an additional market became available and the rest of the market allocation must remain the same, the implications are more general. Supplier 1 would prefer to allocate additional markets to a subsidiary versus simply making a transfer payment, as long as the benefit from serving the market themselves is not larger than the current transfer payment. This depends on the subsidiary having lower newsvendor-type costs for the marginal markets, which will always be the case if the subsidiary is allocated any markets at all.

5.3.2 Outsourcing newsvendor problem and the SNP

In the previous section we considered a few features of the ONP. In this section we use both the ONP and the SNP to improve our understanding of market selection-type problems in practice.

Theorem 5.3.4. *The SNP with no fixed costs can be seen as a special case of the ONP.*

Proof. Consider the outsourcing problem with just two suppliers. Let $v_2 = c_2 = e_2 = 0$. Then, (ONP) reduces to (S1) from above. This is exactly the Selective Newsvendor Problem with no fixed costs as introduced in Section 2.2.1. Therefore, any optimal solution z_1^* and y_1^* to (S1) will also be an optimal solution to the equivalent SNP. $\square$

While we see from this theorem that the SNP with no fixed costs can be thought of as a special case of the ONP (though fixed costs could also be added to the ONP), solving both problems gives us useful insight into the supplier and subsidiaries, the market demands, and how those problem features interact. Either the SNP or the outsourcing solutions could be implemented in practice depending on the company in question, but typically these kinds of strategic decisions are not simply the implementation of optimization models. Instead, we hope that the solutions may help decision makers consider what forms of flexibility they do have in serving a collection of markets and then use that flexibility to improve profits. In the remainder of the section we identify ways in which the solutions to the SNP and ONP inform us about the underlying supply chain setting.

Corollary 5.3.5. *Suppose the optimal solution to the SNP z^* and ONP z_1^* are not the same. If a subsidiary experiences a sufficient reduction in costs, the optimal solutions will be the same.*

Proof. This result follows immediately from the proof of Theorem 5.3.4. □

While this corollary follows immediately from the connection between the SNP and the ONP, it also provides a useful practical perspective on the problem. Specifically, a company which finds the solutions to SNP and ONP are different for their supply chain can see that Supplier 1 has reduced profits from the constraint that demand in all markets must be met. In this case there are two connected explanations for the reduced profitability:

- The supplier would rather not meet demand in all markets, but there are no subsidiaries with low enough costs to allocate the remaining markets to.
- The demand is highly variable but pooling reduces that variability. If there is enough total demand, there can be "support" for multiple suppliers having the benefits of pooling. If there is less total demand relative to the variability, we will only be able to support a single supplier.

Ultimately, depending on the circumstances, Supplier 1 may decide to invest in their own or subsidiary procurement, take actions to reduce the relative variability in demand, or simply accept the reduction in expected profit as a cost of meeting demand in all markets. For settings where Supplier 1 is unwilling or unable to modify costs or demand distributions, we have the following result:

Lemma 5.3.6. *If the optimal solution to the SNP z^* and ONP z_1^* are the same and there are no transfer payments, the constraint that all markets are served does not affect the optimal solution to (ONP).*

Proof. Since z_1^* is the same as z^*, that indicates that either $z^* = \mathbf{1}$ and therefore all markets increase total expected profit or $z^* \neq \mathbf{1}$ but a subsidiary has nonnegative expected profits for meeting demand in the other markets. □

As mentioned above, if Supplier 1 has the option to invest in reducing their own procurement costs or demand variability relative to total demand, they may still be able to increase their expected profits. However, in cases where the benefits would not offset the investment, this is a best-case-scenario for Supplier 1.

5.3.3 Discussion

In this section we presented both some results for the outsourcing newsvendor problem and the equivalence between the selective newsvendor problem with no fixed costs and a special case of the ONP. It is clear from the discussion that while solving the SNP and the ONP each give some information about the underlying market structure, we can see even more by comparing the solution to the ONP to the solution to the equivalent SNP.

5.4 Computational results

In Section 5.3 we mentioned that while the outsourcing newsvendor problem is computationally challenging, we believe that practical-sized problems can be solved in reasonable time. In this section, we present the results of numerical experiments which support that belief and also provide some comparisons between the solutions to the ONP, SNP, and requiring Supplier 1 to serve all markets.

5.4.1 Problem instances

We consider effectively the same set of problem instances as in Chapter 3. We create each problem instance by generating the joint demand distribution of finite cardinality Ξ uniformly from $\prod_{i=1}^{m}[0, u_i]$ where u_i is randomly chosen from the uniform distribution on $[0, 10]$ (for $i = 1, \ldots, m$). We let the procurement cost, salvage value, and expediting costs for Supplier 1 be $c_1 = 0.9$, $v_1 = 0.6$, and $e_1 = 1$ (this is a small deviation from Chapter 3 which had $c = 0.8$). We then have the procurement cost, salvage value, and expediting costs for the subsidiaries be selected randomly from the following uniform distributions: $c_i \in U[0.8, 0.9], v_i \in U[0.6, 0.7]$, and $e_i \in U[0.9, 1.0]$ for each i in $\{2, \ldots, S\}$. Finally, the per-unit market revenues r_j are uniformly chosen from $[0.8, 1.0]$ for each market $j \in \{1, \ldots, m\}$. Because of the change to c_1, we note that on average half of markets will for sure not be profitable for Supplier 1 to serve because the cost of procurement c_1 will be larger than the revenue r_j. However, we believe the resulting problem instances are interesting and illustrative.

We consider the number of suppliers $S \in \{2, 4, 6, 8, 10\}$ which includes Supplier 1 and consider numbers of markets including $m \in \{10, 20, 30, 40, 50\}$. For each pair (S, m) we considered 100 problem instances and any average values were based on the full set of instances. The problem was coded as an MILP and each problem instance was solved using CPLEX version 12.5 with default values. All tests were performed on a PC with an

Intel Xeon Quad Core 3.2 GHz processor with 8 GB RAM using C++. In order to compute the percent profit, the newsvendor-type costs (including any salvage value) experienced by all suppliers were summed and then compared to the absolute profit enjoyed by Supplier 1 or the total system depending on the setting.

5.4.2 Results

We begin this section by showing how quickly CPLEX is able to solve ONP and SNP for our problem instances. We will then present a comparison of the average and per-instance percent profit for each of the models on equivalent problem instances.

In Figures 5.1 and 5.2 we see the computational time required for CPLEX to solve ONP and then SNP. Our results for SNP are consistent with Table 3.1, which is reassuring. We see that with $\Xi = 1000$, up to 10 suppliers and 50 markets can be solved in reasonable time. We do note that the range in solution times across the 100 problem instances showed a lot of variability which we believe explains the shape of the curves in the figues. As an example, it took between 4 and 300 seconds to solve ONP for $S = 10$ and $m = 50$.

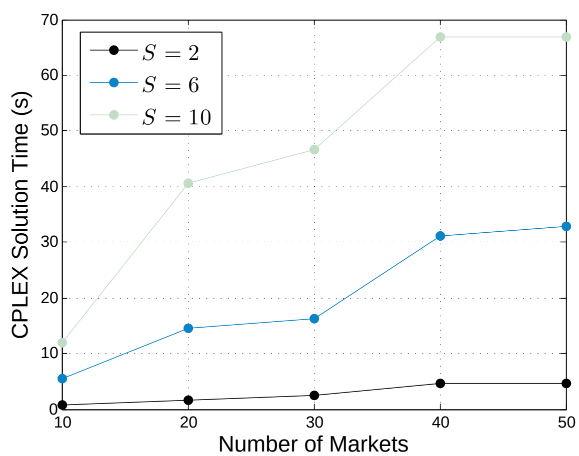


Figure 5.1: CPLEX run time as a function of m and S for ONP.

These results demonstrate that for these instances ONP is a more computationally challenging problem to solve than SNP. Computational complexity is important since, just as for SNP, $\Xi = 1000$ may be impractically small. Consider that having 20 markets and just 2 demand levels in each independent market would involve 2^{20} different possible demand vectors. In these cases, improved methods for solving ONP would be needed. However, as we mentioned in Section 5.3, SAA and rounding could potentially be used to resolve some of the computational challenges.

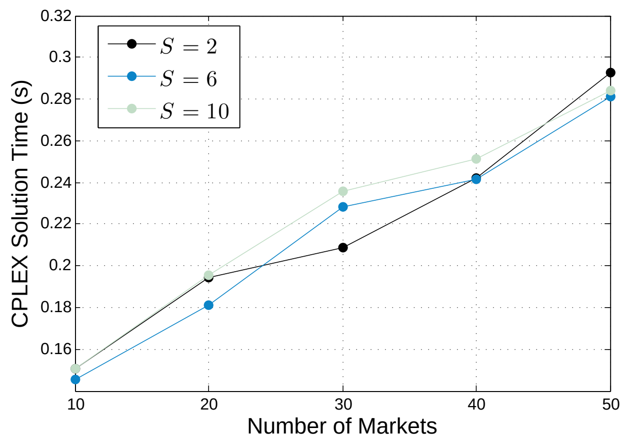


Figure 5.2: CPLEX run time as a function of m and S for SNP.

Beyond computational complexity, the focus of this section is on the quality of solutions to the outsourcing newsvendor problem. We provide three comparisons in order to understand both ONP and SNP for our particular set of problem instances. While the details will necessarily be different in practice, we hope these comparisons can be illustrative of how the different modeling choices could affect expected profit. We begin with two figures as representative examples of the differences in performance and follow with tables with the complete comparisons.

In Figure 5.3 and 5.4 we see the average percent profit experienced by Supplier 1 in three different settings as well the total system profit for ONP. The settings we consider include if they can simply choose their preferred set of markets (SNP), if they must serve all markets themself (ONP constraint with no subsidiaries), and if they work with $S - 1$ subsidiaries. Figure 5.3 shows that when there is only one subsidiary, the total system profit is about the same as when Supplier 1 solves the SNP. However, requiring Supplier 1 to meet all demand results in negative expected profit, while their expected profit under an ONP objective is much closer to SNP. In Figure 5.4 we see the same information when $S = 10$ and observe that the total system profit improves significantly when there are more subsidiaries. These figures highlight that for our problem instances, the difference in procurement costs between Supplier 1 and the subsidiaries is fairly small. In a setting where subsidiary costs are significantly lower than for Supplier 1, we would expect total system profit to be relatively high while Supplier 1 is able to enjoy the same expected profit as in the SNP case.

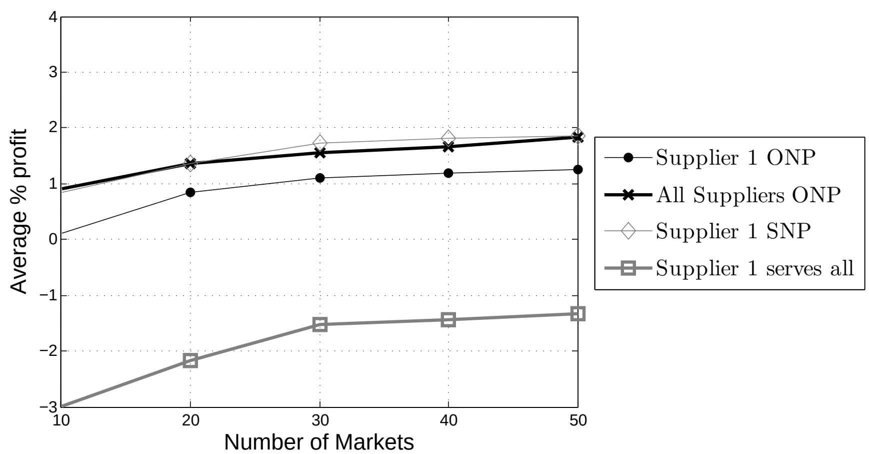


Figure 5.3: Average percent profit for different models with $S = 2$.

In Table 5.1 we compare the average percent profit for Supplier 1 under the ONP model as well as the equivalent SNP. We find that when there are more markets and suppliers, Supplier 1 is able to come close to the solution to the SNP. This is a result of increased likelihood of a subsidiary with lower procurements costs and the increased total demand. We note that in Table 5.1, in fact there are some settings where the average percent profit is higher for ONP than SNP. This is a result of our choice to consider percent profit rather than absolute profit and the fact that ONP can have lower total costs than SNP.

While Supplier 1 can hope that at best their absolute expected profit for ONP will equal

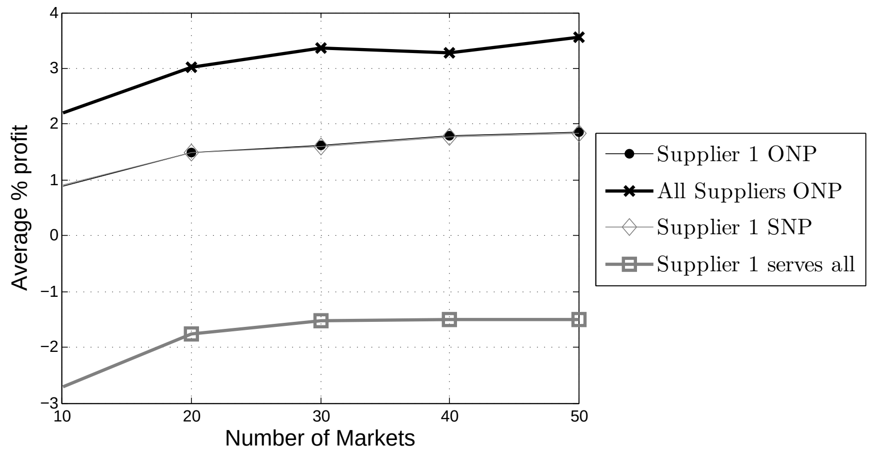


Figure 5.4: Average percent profit for different models with $S = 10$.

	Supplier 1 ONP					SNP				
	m					m				
S	10	20	30	40	50	10	20	30	40	50
2	0.1%	0.8%	1.1%	1.2%	1.3%	0.8%	1.4%	1.7%	1.8%	1.9%
4	0.8%	1.4%	1.6%	1.7%	1.8%	1.0%	1.4%	1.6%	1.8%	1.9%
6	0.8%	1.4%	1.6%	1.7%	1.9%	0.9%	1.4%	1.6%	1.7%	1.9%
8	0.8%	1.4%	1.6%	1.8%	2.0%	0.9%	1.4%	1.6%	1.8%	1.9%
10	0.9%	1.5%	1.6%	1.8%	1.9%	0.9%	1.5%	1.6%	1.8%	1.8%

Table 5.1: Supplier 1 average percent profit for ONP and SNP.

	SNP					All Supplier ONP				
	m					m				
S	10	20	30	40	50	10	20	30	40	50
2	0.8%	1.4%	1.7%	1.8%	1.9%	0.9%	1.4%	1.6%	1.7%	1.8%
4	1.0%	1.4%	1.6%	1.8%	1.9%	1.8%	2.4%	2.5%	2.8%	2.8%
6	0.9%	1.4%	1.6%	1.7%	1.9%	1.6%	2.6%	3.0%	2.9%	3.2%
8	0.9%	1.4%	1.6%	1.8%	1.9%	1.9%	2.7%	3.2%	3.3%	3.4%
10	0.9%	1.5%	1.6%	1.8%	1.8%	2.2%	3.0%	3.4%	3.3%	3.6%

Table 5.2: Supplier 1 average percent profit for SNP and system profit for ONP.

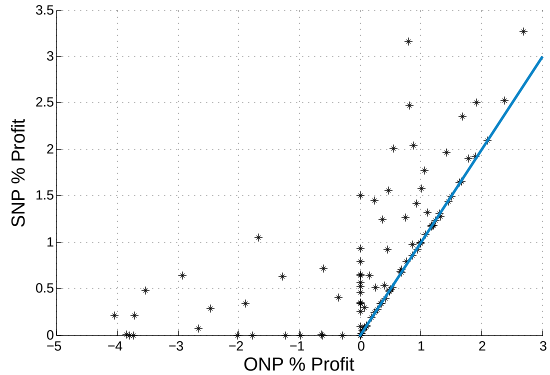
that for SNP, the system as a whole can do much better. In Table 5.2 we compare the average percent profit for the SNP to the system average percent profit for the ONP. When there is only one subsidiary, there is a relatively small difference between the two models. For more subsidiaries, the average system percent profit increases substantially over the SNP values.

One of the goals of this chaper in the context of this thesis is to address the notion that suppliers should choose to meet demand in all candidate markets. While Chapter 1 covered the perspective that in some sense suppliers are always choosing only a subset of demand, this chapter uses a different strategy. In Table 5.3 we directly make the comparison between the average percent profit for Supplier 1 using the ONP model by comparison to simply requiring Supplier 1 to serve all markets. It is clear immediately that using outsourcing as a tool is a significant improvement over simply meeting all demand and can be the difference between making and losing money.

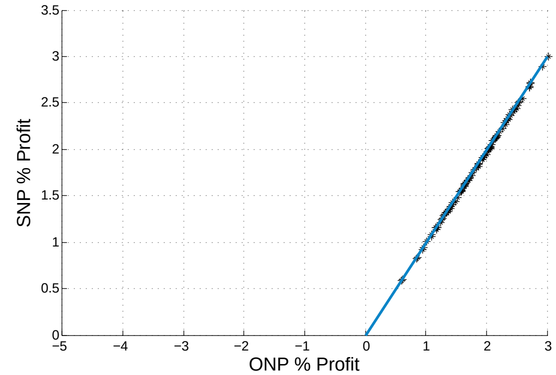
	Supplier 1 ONP					Supplier 1 serves all				
	m					m				
S	10	20	30	40	50	10	20	30	40	50
2	0.1%	0.8%	1.1%	1.2%	1.3%	-3.0%	-2.2%	-1.5%	-1.4%	-1.3%
4	0.8%	1.4%	1.6%	1.7%	1.8%	-3.0%	-2.0%	-1.7%	-1.5%	-1.3%
6	0.8%	1.4%	1.6%	1.7%	1.9%	-3.1%	-2.1%	-1.9%	-1.6%	-1.3%
8	0.8%	1.4%	1.6%	1.8%	2.0%	-3.1%	-2.2%	-1.7%	-1.5%	-1.3%
10	0.9%	1.5%	1.6%	1.8%	1.9%	-2.7%	-1.8%	-1.5%	-1.5%	-1.5%

Table 5.3: Supplier 1 average percent profit for ONP and serving all markets.

We conclude this section by showing the per-instance variation between different models. In Figure 5.5 we see the 100 instances for both the smallest pair $S = 2, m = 10$ and the largest pair $S = 10, m = 50$. For the smaller problem instances, there is substantial variability and ONP has several instances with negative expected profit. For larger instances, in every case Supplier 1 with the ONP model is able to match the performance of the SNP. In Figure 5.6, we instead see a comparison between Supplier 1 serving all markets and ONP. For the smaller problem instances, there is a lot of variability between the two models. For

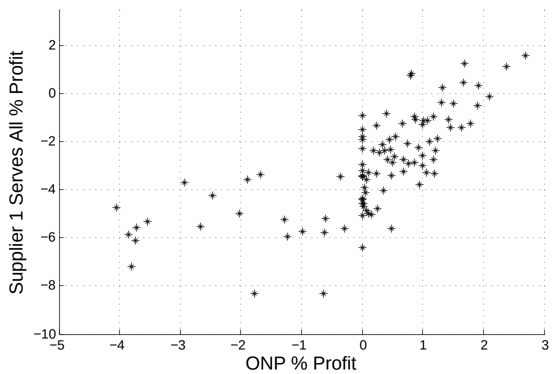


(a) $S = 2, m = 10$

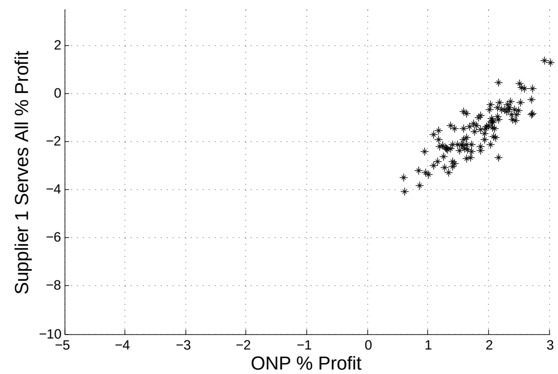


(b) $S = 10, m = 50$

Figure 5.5: Percent profit across instances for ONP vs. SNP



(a) $S = 2, m = 10$



(b) $S = 10, m = 50$

Figure 5.6: Percent profit across instances for ONP vs. Supplier 1 serving all markets

larger problem instances the correlation is much higher, but as we saw before Supplier 1 serving all markets has much lower average profits.

5.5 Concluding remarks

In this chapter we presented the outsourcing newsvendor problem which is a new model and includes the selective newsvendor problem as a special case. We showed that solving both the ONP and the SNP can provide a decision maker with relevant insights into the market demand and supplier / subsidiary structure. We also demonstrated that the version of the problem considered here could be solved in reasonable time using CPLEX, and we believe the methods developed in Chapter 3 of this thesis could be modified to apply to this problem as well. We saw that while solving the ONP results in a loss of expected profit relative to the SNP, for our computational experiments we were able to significantly outperform the solutions when Supplier 1 simply met demand in all markets. In the future we hope that the perspective presented in this chapter can help decision makers assess whether for a particular product they should work with subsidiaries.

CHAPTER 6

Conclusion

6.1 Thesis summary

In this dissertation I considered several models related by the same goal of supporting a supplier who needs to choose a subset of all possible stochastic demand to meet. This work included generalizations of the previously-studied selective newsvendor problem as well as an auction setting and the newly-introduced outsourcing newsvendor problem. Deciding how to choose a subset of candidate demand to meet is a difficult problem which cannot simply be modeled as an optimization. However, optimization tools can help decision makers consider which issues are most important to them, and how those concerns can be used. By presenting several optimization tools in the same thesis, I hope that readers are able to think more broadly about demand and market selection for their practical setting.

Beyond the big-picture contributions, this thesis includes a variety of interesting methodological developments. In Chapter 2 we developed both exact and heuristic methods to solve a multi-period generalization of the SNP and also presented numerical experiments demonstrating the effectiveness of the approaches. In Chapter 3 we studied a very general selection problem and developed an algorithm which combined sample average approximation and rounding to find a high-quality solution with high probability. Combining multiple approximation algorithms allowed us to solve the stochastic selection problem for very general (nonnegative support with finite mean) stochastic distributions. We also showed through numerical experiments that for the SNP, the approximation algorithm was very effective even for very few distribution samples. In Chapter 4 we shifted gears and used an efficient auction mechanism to allocate a limited supply of goods to a set of newsvender-type cost retailers. We also introduced an ϵ-efficient auction which allows the auctioneer to run a discrete auction which is not guaranteed to be efficient, but is guaranteed to be within a fixed difference from efficient. Finally, in Chapter 5 we consider a supplier who may work with subsidiaries to meet demand in a set of markets. The newly-introduced Outsourcing Newsvendor Problem provides an alternative to the SNP with important implications in practice.

In conclusion, this thesis has included several research problems which try to answer the very general question of how suppliers choose a subset of all demand to meet. The models have focused on stochastic demands because many industries have significant effects from demand uncertainty, but that choice make the optimization models much more challenging to solve. This work has included a variety of methodological tools as well as illustrative numerical experiments and practical discussions.

6.2 Future Work

This thesis is connected largely by the high-level perspective on the problems studied. From that perspective, we introduced two models which are entirely new to supply chain optimization in addition to the generalizations of the SNP. There are many directions for future work related to the content of this thesis.

Chapter 2 included a multi-period multi-market selective newsvendor problem as a special case. However, the model did not include a way to have finite but nonzero backlogging and carryover costs. Similarly, finite but nonzero costs for changing the market selection from period to period are an important concern for practical problems. Addressing these issues even with heuristic solutions may significantly improve the implementability of the optimal solution in practice. In Chapter 3 we considered a very general market selection problem with minimal assumptions on the stochastic demand. The combined approach of using both rounding and Sample Average Approximation was an effective tool based on our numerical experiments to solve the SNP quickly and with high-quality solutions. Given these observations, we believe this approach could be used in other settings with expected-value objectives.

While Chapters 2 and 3 built on the literature for selective newsvendor problems, Chapters 4 and 5 are much more of a departure from existing work. In Chapter 4 we presented an efficient auction for a fixed quantity to a set of newsvendor-type cost retailers. However, efficient auctions are just one objective and are unlikely to be the best approach for every seller. A natural next step is to consider Suppliers interested in maximizing their expected revenues from the auction, which is known as an optimal auction. Maskin and Riley [6] develop an optimal multi-unit auction with multi-unit demands and Deshpande and Schwarz [55] apply that work to a supply-chain setting similar to the one studied in this thesis. In addition, most of the similar work in the auction literature and practice is in the area of treasury bond auctions. Kastl [56] covers the discriminatory and uniform price auctions which may also be useful to suppliers with a fixed inventory. Chapter 5 also has many possible future research directions. Developing an exact algorithm for independent and normally distributed demands is an ideal next step. Using SAA and rounding to efficiently solve the more general problem would also be beneficial. In the future, it would also be interesting to consider a multi-period version of the problem as well as one which included investment opportunities to improve the procurement costs of Supplier 1 or the subsidiaries.

APPENDIX A

Paper 2 appendicies

A.1 Proof of Lemma 3.2.4

(i) *Convexity:*

Let $0 \leq \lambda \leq 1$. Then

$$\begin{aligned}
\rho\left[\lambda X_1 + (1-\lambda)X_2\right] &= \inf_{\theta\in\Theta} E\left[G\left(\lambda X_1 + (1-\lambda)X_2;\theta\right)\right] \\
&= \inf_{\theta_1,\theta_2\in\Theta} E\left[G\left(\lambda X_1 + (1-\lambda)X_2;\lambda\theta_1 + (1-\lambda)\theta_2\right)\right] \\
&\leq \inf_{\theta_1,\theta_2\in\Theta} \left\{\lambda E\left[G(X_1;\theta_1)\right] + (1-\lambda)E\left[G(X_2;\theta_2)\right]\right\} \\
&= \lambda \inf_{\theta_1\in\Theta} E\left[G(X_1;\theta_1)\right] + (1-\lambda) \inf_{\theta_2\in\Theta} E\left[G(X_2;\theta_2)\right] \\
&= \lambda\rho[X_1] + (1-\lambda)\rho[X_2].
\end{aligned}$$

(ii) *Monotonicity:*

Suppose that $X_2 \succeq X_1$. Then

$$\begin{aligned}
\rho[X_2] &= \inf_{\theta\in\Theta} E\left[G(X_2;\theta)\right] \\
&\geq \inf_{\theta\in\Theta} E\left[G(X_1;\theta)\right] \\
&= \rho[X_1].
\end{aligned}$$

(iii) *Positive Homogeneity:*

Let $\lambda > 0$. Then

$$\begin{aligned}\rho[\lambda X] &= \inf_{\theta\in\Theta} E\left[G(\lambda X;\theta)\right] \\ &= \inf_{\theta\in\Theta} E\left[G(\lambda X;\lambda\theta)\right] \\ &= \inf_{\theta\in\Theta} E\left[\lambda G(X;\theta)\right] \\ &= \inf_{\theta\in\Theta} \lambda E\left[G(X;\theta)\right] \\ &= \lambda\rho[X].\end{aligned}$$

(iv) *Translation Equivariance:*

Let $a \in \mathbb{R}$. Then

$$\begin{aligned}\rho[X+a] &= \inf_{\theta\in\Theta} E\left[G(X+a;\theta)\right] \\ &= \inf_{\theta\in\Theta} E\left[G(X+a;\theta+a\mathbf{1})\right] \\ &= \inf_{\theta\in\Theta} E\left[G(X;\theta)+a\right] \\ &= \inf_{\theta\in\Theta} E\left[G(X;\theta)\right]+a \\ &= \rho[X]+a.\end{aligned}$$

A.2 Proof of Theorem 3.3.4

We start by fixing a value $\beta \in [0,\delta]$. Since $\mathcal{Y}$ is a cone we have that $\frac{1}{1-\beta}\hat{y} \in \mathcal{Y}$ so that $\Psi\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}\right)$ is well-defined. Condition (ii) of the theorem implies that

$$\Gamma_2\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y};\mathbf{F},\mathbf{D}\right) \preceq \frac{1}{1-\beta}\Gamma_2(\hat{z},\hat{y};\mathbf{F},\mathbf{D}).$$

Together with Condition (i) of the theorem in turn implies that

$$\begin{aligned}\Psi\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}\right) &= \rho\left[\Gamma\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}; \mathbf{F}, \mathbf{D}\right)\right] \\ &= \rho\left[\bar{r}^\top\left(\mathbf{1} - [\hat{z}]_\beta\right) + \Gamma_2\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}; \mathbf{F}, \mathbf{D}\right)\right] \\ &= \bar{r}^\top\left(\mathbf{1} - [\hat{z}]_\beta\right) + \rho\left[\Gamma_2\left([\hat{z}]_\beta, \frac{1}{1-\beta}\hat{y}; \mathbf{F}, \mathbf{D}\right)\right]\end{aligned}$$

(by the translation invariance property of ρ)

$$\leq \bar{r}^\top\left(\mathbf{1} - [\hat{z}]_\beta\right) + \rho\left[\frac{1}{1-\beta}\Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right]$$

(by the monotonicity property of ρ)

$$= \bar{r}^\top\left(\mathbf{1} - [\hat{z}]_\beta\right) + \frac{1}{1-\beta}\rho\left[\Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right]$$

(by the positive homogeneity property of ρ). Now replace β by the random variable $\boldsymbol{\beta} \sim U[0, \delta]$. Then

$$\begin{aligned}E\left[1 - [z_i]_{\boldsymbol{\beta}}\right] &= \Pr\left(1 - [z_i]_{\boldsymbol{\beta}} = 1\right) = \Pr\left([z_i]_{\boldsymbol{\beta}} = 0\right) \\ &= \Pr\left(z_i < 1 - \boldsymbol{\beta}\right) = \Pr\left(\boldsymbol{\beta} < 1 - z_i\right) \\ &= \min\left\{1, \frac{1-z_i}{\delta}\right\} \leq \frac{1-z_i}{\delta}.\end{aligned}$$

We use the last inequality to obtain

$$\begin{aligned}E_{\boldsymbol{\beta}}\left[\Psi\left([\hat{z}]_{\boldsymbol{\beta}}, \frac{1}{1-\boldsymbol{\beta}}\hat{y}\right)\right] &\leq E_{\boldsymbol{\beta}}\left[\bar{r}^\top\left(\mathbf{1} - [\hat{z}]_\beta\right)\right] + E_{\boldsymbol{\beta}}\left[\frac{1}{1-\boldsymbol{\beta}}\right]\rho\left[\Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right] \\ &\leq \frac{1}{\delta}\bar{r}^\top(\mathbf{1} - \hat{z}) + \frac{1}{\delta}\ln\left(\frac{1}{1-\delta}\right)\rho\left[\Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right] \\ &\leq \frac{1}{\delta}\max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\}\left(\bar{r}^\top(\mathbf{1} - \hat{z}) + \rho\left[\Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right]\right)\end{aligned}$$

$$= \frac{1}{\delta} \max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\} \rho\left[\left(\bar{r}^{\top}(\mathbf{1} - \hat{z}) + \Gamma_2(\hat{z}, \hat{y}; \mathbf{F}, \mathbf{D})\right]\right)$$
$$= \frac{1}{\delta} \max\left\{1, \ln\left(\frac{1}{1-\delta}\right)\right\} \Psi\left(\hat{z}, \hat{y}\right).$$

A.3 Alternative and explicit bounds

In this appendix we consider Example 1 from Section 3.2.2.1 with $\rho = E$ (i.e., the selective newsvendor problem). For this class of problems, an alternate required sample size $|\mathcal{N}_2|$ can be used by employing an approximation pertaining to the newsvendor problem that is a slight modification of a result from [49]. We also derive explicit bounds on the global problem parameters Δ and L for this case.

A.3.1 Alternative value for $\mathcal{N}_2$

We first provide an approximation result for the newsvendor problem that is based on a result from [49] but accounts for the slightly different formulation of our cost function that is due to the selection component of our problem.

Lemma A.3.1. *Let $\mathcal{N}_2$ be a sample satisfying*

$$\|\mathcal{N}_2| \geq \frac{9}{2\epsilon^2}\left(\frac{\min(e-c, c-v)}{e-v}\right)^{-2} \log\left(\frac{2}{\delta_2}\right)$$

where $\epsilon \in (0, 1]$ is a relative cost error measure and $1 - \delta_2$ is a confidence level. Then, for a given $z \in [0, 1]^m$ let $\hat{y}^{(0)} = \arg\min_{y \in \mathcal{Y}} \Psi_{\mathcal{N}_2}(z, y)$. Then

$$\Psi(z, \hat{y}^{(0)}) \leq (1 + \epsilon)\Psi^*(z)$$

with probability at least $1 - \delta_2$.

Proof. Lemma 2.2 in [49] say that $\hat{y}^{(0)}$ satisfies

$$\left| F_z\left(\hat{y}^{(0)}\right) - \frac{e-c}{e-v} \right| \leq \frac{\epsilon}{3} \min\left(\frac{e-c}{e-v}, \frac{c-v}{e-v}\right) \equiv \alpha$$

where F_z is the c.d.f. of $\mathbf{D}^\top z$, i.e., $\hat{y}^{(0)}$ is α-*accurate*. The cost function considered by [49] is given by $C(y) = \Psi(z,y) - E[\mathbf{R}]^\top(1-z) - E[\mathbf{F}]^\top z - cE[\mathbf{D}]$. Corollary 2.1 in [49] show that $C\left(\hat{y}^{(0)}\right) \leq (1+\epsilon)\max_{y\in\mathbb{R}} C(y)$. Since $E[\mathbf{F}], E[\mathbf{R}], E[\mathbf{D}] \geq 0$ we obtain

$$\Psi\left(z, \hat{y}^{(0)}\right) \leq (1+\epsilon)\Psi^*(z).$$

The desired result then follows directly from Theorem 2.2 in [49]. □

The result of Lemma A.3.1 can then replace the additive approximation result of the form $\Psi(z,\hat{y}) \leq \Psi(z,\hat{y}) + \tau$ that we used earlier in this paper, yielding alternative approximation results.

A.3.2 Explicit upper bounds on Δ

We provide two approaches to bound $\Delta = \sup_{(z,y)\in([0,1]^m\times\mathbb{R}_{\geq 0})} \|z-y\|$. While bounding the binary portion is trivial, $\mathcal{Y}$ is in general unbounded. We derive a bounded set $\bar{\mathcal{Y}}$ which includes an optimal solution (possibly with a confidence guarantee).

A.3.2.1 Deterministic upper bound

Lemma A.3.2. *Consider problem (R), and suppose there exists some function $\Xi(y)$ such that $\Xi(y) \leq \Psi(z,y)$ for any $(z,y) \in (\{0,1\}^m \times \mathbb{R}_+)$. Then, without loss of optimality, we may restrict the feasible region of (R) to the set*

$$\bar{\mathcal{Y}} = \{y \mid \Xi(y) \leq \Psi^*,\ i = 1,\ldots,n\}.$$

Proof. This follows immediately from the fact that Ξ is a lowerbounding function to Ψ. □

It is easy to see that for the expected-value newsvendor we have

$$\begin{aligned}\Psi(z,y) &= E\left[\mathbf{R}^\top(\mathbf{1}-z)+\mathbf{F}^\top z + cy + e\left(\mathbf{D}^\top z - y\right)^+ - v\left(y-\mathbf{D}^\top z\right)^+\right]\\ &= E\left[\mathbf{R}^\top(\mathbf{1}-z)+\mathbf{F}^\top z + (c-v)y + v\mathbf{D}^\top z + (e-v)\left(\mathbf{D}^\top z - y\right)^+\right]\\ &\geq (c-v)y\end{aligned}$$

so that we can choose $\Xi(y) = (c-v)y$, which satisfies the conditions of Lemma A.3.2. This means that we can limit ourselves to order quantities $0 \leq y \leq \frac{\Psi^*}{c-v}$. Now consider the feasible solution $(z,y) = (\mathbf{0},0)$. It is easy to see that $\Psi(\mathbf{0},0) = E[\mathbf{R}^\top\mathbf{1}] \geq \Psi^*$, so that we can definitely restrict ourselves to $\bar{\mathcal{Y}} = [0,\bar{y}]$ with $\bar{y} = \frac{E[\mathbf{R}^\top\mathbf{1}]}{c-v}$, so that $\Delta \leq \frac{E[\mathbf{R}^\top\mathbf{1}]}{c-v}$.

A.3.2.2 Probabilistic upper bound

Let $y^*(\mathbf{1})$ be the optimal order quantity when all markets are selected ($z = \mathbf{1}$) and F denote the c.d.f. of the associated aggregate demand $D^\top\mathbf{1}$. Now note that the aggregate demand is stochastically largest when all markets are selected (since demands are nonnegative) and therefore $y^*(\mathbf{1})$ is a valid upper bound on the optimal order quantity for any $z \in \{0,1\}^m$. If $y^*(\mathbf{1})$ can easily be computed, we have a candidate bound on Δ. However, in general, this task is difficult, so we may apply Lemma 2.2 in [49] to provide a probabilistic upper bound, i.e., a bound that is valid with some guaranteed probability.

Theorem A.3.3. *Let $\alpha, \delta_3 \in (0,1)$ and let $\mathcal{N}_3$ be a sample satisfying*

$$|\mathcal{N}_3| \geq \frac{1}{2\alpha^2}\ln\left(\frac{2}{\delta_3}\right).$$

Then the $\left(\frac{e-c}{e-v}+\alpha\right)$-quantile of the sample of $D^\top\mathbf{1}$ is an upper bound on Δ with probability at least $1-\delta_3$.

Proof. Let $\hat{y}$ be the $\left(\frac{e-c}{e-v}+\alpha\right)$-quantile of a sample $\mathcal{N}_3$ of $D^\top\mathbf{1}$ values as in the theorem. From Lemma 2.2 in [49] we have that $\hat{y}$ is α-accurate with probability at least $1-\delta_3$, and therefore

$$\left|F(\hat{y}) - \left(\frac{e-c}{e-v} + \alpha\right)\right| \leq \alpha$$

which implies that

$$F(\hat{y}) \geq \frac{e-c}{e-v}$$

so that $\hat{y}$ is an upper bound for $y^*(\mathbf{1})$, and hence for Δ, with probability at least $1 - \delta_3$. $\square$

If the sample $\mathcal{N}_3$ is chosen independently from the samples $\mathcal{N}_1$ and $\mathcal{N}_2$ used in the approximation algorithms, this probabilistic bound can be used to derive approximation results similar to the ones provided earlier in the paper, at the expense of an additional loss in confidence due to the factor $1 - \delta_3$.

A.3.3 Explicit upper bound on L

A.3.3.1 Upper bound on the global L

Lemma A.3.4. *For any* $(z, y), (z', y') \in [0,1]^m \times \mathbb{R}_+$ *we have that*

$$|\Psi(z,y) - \Psi(z',y')| \leq L\|(z,y) - (z',y')\|_2$$

with

$$L = \sqrt{m}\max\left(e - v, \|E[\mathbf{F} + c\mathbf{D} - \mathbf{R}]\|_\infty + (e-v)\|E[\mathbf{D}]\|_\infty\right).$$

Proof. Let $(z, y), (z', y') \in [0,1]^m \times \mathbb{R}_+$ and consider the following representation of the cost function:

$$\Psi(z,y) = E\left[\mathbf{R}^\top(\mathbf{1} - z) + (\mathbf{F} + c\mathbf{D})^\top z + (c-v)\left(y - \mathbf{D}^\top z\right)^+ + (e-c)\left(\mathbf{D}^\top z - y\right)^+\right].$$

The difference in costs between these two solutions is then:

$$\Psi(z,y) - \Psi(z',y') = E[(\mathbf{F} + c\mathbf{D} - \mathbf{R})^\top(z - z')] + (c-v)\left(E\left[\left(y - \mathbf{D}^\top z\right)^+\right]\right)$$

$$
\begin{aligned}
&\quad - (c-v)\left(E\left[\left(y' - \mathbf{D}^\top z'\right)^+\right]\right) \\
&\quad + (e-c)\left(E\left[\left(\mathbf{D}^\top z - y\right)^+\right] - E\left[\left(\mathbf{D}^\top z' - y'\right)^+\right]\right) \\
&\leq \|E[\mathbf{F} + c\mathbf{D} - \mathbf{R}]\|_\infty |z - z'|_1 + (e-v)|y - y'| \\
&\quad + (e-v)\|E[\mathbf{D}]\|_\infty \|z - z'\|_1 \\
&\leq \max\Big\{e - v, \\
&\quad \|E[\mathbf{F} + c\mathbf{D} - \mathbf{R}]\|_\infty + (e-v)\|E[\mathbf{D}]\|_\infty\Big\} \cdot \|(z,y) - (z'', y'')\|_1 \\
&\leq \sqrt{m}\max\Big\{e - v, \\
&\quad \|E[\mathbf{F} + c\mathbf{D} - \mathbf{R}]\|_\infty + (e-v)\|E[\mathbf{D}]\|_\infty\Big\} \cdot \|(z,y) - (z'', y'')\|_2
\end{aligned}
$$

which proves the desired result. □

A.3.3.2 Upper bound on $L(z)$

If z is fixed, we can instead compute a bound $L(z)$.

Lemma A.3.5. *For any $(z) \in [0,1]^m$ and $y, y' \in \mathbb{R}_+$ we have that*

$$
|\Psi(z,y) - \Psi(z',y')| \leq L\|(z,y) - (z',y')\|_2
$$

where $L = \max\{(e-c),(c-v)\}$.

Proof. Because $\Psi(z,y)$ is simply the expected value of the newsvendor cost function once z is fixed, we can immediately see that in the worst case, $y - y'$ represents either buying a unit that is used with probability 1 (leading to a change of $e - c$ in the objective), or buying a unit that is used with probability 0 (leading to a change of $c - v$ in the objective). Therefore, $L(z)$ is the max of these two values. □

BIBLIOGRAPHY

[1] Krishnan, V. and Ulrich, K., “Product development decisions: A review of the literature,” *Management Science*, Vol. 47, No. 1, 2001, pp. 1–21.

[2] Chiang, W., Chen, J., and Xu, X., “An overview of research on revenue management: current issues and future research,” *Int. J. Revenue Management*, Vol. 1, No. 1, 2007, pp. 97–128.

[3] Yano, C. and Gilbert, S., “Coordinated Pricing and Production/Procurement Decisions: A Review,” *Managing Business Interfaces: Marketing and Engineering Issues in the Supply Chain and Internet Domains*, edited by A.K. Chakravarty and J. Eliashberg , Vol. 16 of *International Series in Quantitive Marketing*, 2005, pp. 65–103.

[4] Petruzzi, N. and Dada, M., “Pricing and the newsvendor problem: A review with extensions,” *Operations Research*, Vol. 47, No. 2, 1999, pp. 183–194.

[5] Krishna, V., “Auction Theory, 2nd Edition,” *Auction Theory, 2nd Edition*, Elsevier, 2010.

[6] Maskin, E. and Riley, J., “Optimal multi-unit auctions,” *The economics of missing markets, information, and games*, 1989.

[7] Ausubel, L., “An Efficient Ascending-Bid Auction for Multiple Objects,” *The American Economic Review*, Vol. 94, No. 5, 2004, pp. 1452–1475.

[8] Chen, F., “Auctioning supply contracts,” *Management Science*, Vol. 53, No. 10, 2007, pp. 1562–1576.

[9] Duenyas, I., Hu, B., and Beil, D., “Simple Auctions for Supply Contracts,” *Management Science*, Vol. 59, No. 10, 2013, pp. 2332–2342.

[10] “Competitive location, production, and market selection,” *European Journal of Operational Research*, Vol. 149, No. 1, 2003, pp. 211 – 228.

[11] Geunes, J., Merzifonluoğlu, Y., Romeijn, H., and Taaffe, K., “Demand selection and assignment problems in supply chain planning,” *TutORials in Operations Research*, edited by J. Smith, INFORMS, 2005, pp. 124–141.

[12] Van den Heuvel, W., Kundakcioglu, E., Geunes, J., Romeijn, H., Sharkey, T., and Wagelmans, A., "Integrated market selection and production planning: complexity and solution approaches," *Mathematical Programming*, Vol. 134, No. 2, 2012, pp. 395–424.

[13] Geunes, J., Levi, R., Romeijn, H., and Shmoys, D., "Approximation algorithms for supply chain planning and logistics problems with market choice," *Mathematical Programming*, Vol. 130, No. 1, 2011, pp. 85–106.

[14] Carr, S. and Lovejoy, W., "The inverse newsvendor problem: Choosing an optimal demand portfolio for capacitated resources," *Management Science*, Vol. 46, No. 7, 2000, pp. 912–927.

[15] Taaffe, K., Geunes, J., and Romeijn, H., "Target market selection and marketing effort under uncertainty: the selective newsvendor," *European Journal of Operational Research*, Vol. 189, No. 3, 2008, pp. 987–1003.

[16] Chahar, K. and Taaffe, K., "Risk averse demand selection with all or nothing orders," Vol. 37, No. 5, 2009, pp. 996–1006.

[17] Taaffe, K. and Chahar, K., "Risk averse selective newsvendor problems," *International Journal of Operational Research*, Vol. 3, No. 6, 2008, pp. 681–703.

[18] Waring, A., *Risk-Averse Selective Newsvendor Problems*, Ph.D. thesis, University of Michigan, 2012.

[19] Khouja, M., "The single-period (news-vendor) problem: literature review and suggestions for future research," *Omega*, Vol. 27, No. 5, 1999, pp. 537 – 553.

[20] Qin, Y., Wang, R., Vakharia, A., Chen, Y., and Seref, M., "The newsvendor problem: Review and directions for future research," *European Journal of Operational Research*, Vol. 213, No. 2, 2011, pp. 361 – 374.

[21] Kraiselburd, S. and andA. Raman, V. N., "Contracting in a Supply Chain with Stochastic Demand and Substitute Products," *Production and Operations Management*, Vol. 13, No. 1, 2004, pp. 46–62.

[22] Gotoh, J. and Takano, Y., "Newsvendor solutions via conditional value-at-risk minimization," *European Journal of Operational Research*, Vol. 179, No. 1, 2007, pp. 80–96.

[23] Ahmed, S., Çakmak, U., and Shapiro, A., "Coherent risk measures in inventory problems," *European Journal of Operational Research*, Vol. 182, 2007, pp. 226–238.

[24] Weng, Z., "Coordinating order quantities between the manufacturer and the buyer: A generalized newsvendor model," *European Journal of Operational Research*, Vol. 156, No. 1, 2004, pp. 148 – 161.

[25] Chen, K. and Xiaon, T., "Outsourcing strategy and production disruption of supply chain with demand and capacity allocation uncertainties," *International Journal of Production Economics*, Vol. 170, No. A, 2015, pp. 243–257.

[26] Strinka, Z., Romeijn, H., and Wu, J., "Exact and heuristic methods for a class of selective newsvendor problems with normally distributed demands," *Omega*, Vol. 41, No. 2, 2013, pp. 250–258.

[27] Freling, R., Romeijn, H., Morales, D. R., and Wagelmans, A., "A branch-and-price algorithm for the multi-period single-sourcing problem," *Operations Research*, Vol. 51, No. 6, 2003, pp. 922–939.

[28] Huang, W., Romeijn, H., and Geunes, J., "The continuous-time single-sourcing problem with capacity expansion opportunities," *Naval Research Logistics*, Vol. 52, 2005, pp. 193–211.

[29] Shen, Z., Coullard, C., and Daskin, M., "A Joint Location-Inventory Model," *Transportation Science*, Vol. 37, No. 1, 2003, pp. 40–55.

[30] Shu, J., Teo, C., and Shen, Z., "Stochastic transportation-inventory network design problem," *Operations Research*, Vol. 53, No. 1, 2005, pp. 48–60.

[31] Taaffe, K., Geunes, J., and Romeijn, H., "Supply capacity acquisition and allocation with uncertain customer demands," *European Journal of Operational Research*, Vol. 204, No. 2, 2010, pp. 263–273.

[32] Porteus, E., "Stochastic inventory theory," *Handbooks in Operations Research and Management Science*, edited by D. Heyman and M. Sobel, Vol. 2, chap. 12, North-Holland, Amsterdam, The Netherlands, 1990, pp. 605–652.

[33] Eppen, G., "Effects of centralization on expected costs in a multi-location newsboy problem," *Management Science*, Vol. 25, No. 5, 1979, pp. 498–501.

[34] Taaffe, K., Romeijn, H., and Tirumalasetty, D., "A selective newsvendor approach to order management," *Naval Research Logistics*, Vol. 55, No. 8, 2008, pp. 769–784.

[35] Geunes, J., Shen, Z.-J., and Romeijn, H., "Economic ordering decisions with market selection flexibility," *Naval Research Logistics*, Vol. 51, No. 1, 2004, pp. 117–136.

[36] Chen, X. and Zhang, J., "A stochastic Programming Duality Approach to Inventory Centralization Games," *Operations Research*, Vol. 57, No. 4, 2009, pp. 840–851.

[37] Huang, X. and Sošić, G., "Repeated newsvendor game with transshipments under dual allocations," *European Journal of Operational Research*, Vol. 204, 2010, pp. 274–284.

[38] Bazaraa, M., Sherali, H., and Shetty, C., *Nonlinear Programming: Theory and Algorithms*, 2nd ed., 1993.

[39] Shu, J., Li, Z., and Huang, L., "Demand selection decisions for a multi-echelon inventory distribution system," *Journal of the Operational Research Society*, Vol. 64, No. 9, 2013, pp. 1307–1313.

[40] Alptekinoğlu, A. and Tang, C., "A model for analyzing multi-channel distribution systems," *European Journal of Operational Research*, Vol. 163, No. 3, 2005, pp. 802–824.

[41] Lin, J. and Ng, T., "Robust multi-market newsvendor models with interval demand data," *European Journal of Operational Research*, Vol. 212, No. 2, 2011, pp. 361–373.

[42] Artzner, P., Delbaen, F., Eber, J., and Heath, D., "Coherent measures of risk," *Mathematical Finance*, Vol. 9, No. 3, 1999, pp. 203–228.

[43] Rockafellar, R. and Uryasev, S., "Optimization of Conditional Value-at-Risk," *Journal of Risk*, Vol. 2, 2000, pp. 21–41.

[44] Pflug, G., "Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk," *Probabilistic Constrained Optimization: Methodology and Applications*, edited by S. Uryasev, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000, pp. 272–281.

[45] Ogryczak, W. and Ruszczyński, A., "Dual stochastic dominance and related mean-risk models," *SIAM Journal on Optimization*, Vol. 13, 2002, pp. 60–78.

[46] Hardy, G., Littlewood, J., and Pólya, G., *Inequalities*, Cambridge [Eng.]: University Press, 2nd ed., 1952.

[47] Kleywegt, A., Shapiro, A., and de Mello, T. H., "The Sample Average Approximation Method for Stochastic Discrete Optimization," *SIAM Journal on Optimization*, 2001, pp. 479 502.

[48] Shapiro, A., "Monte Carlo Sampling Methods," *Stochastic Programming*, edited by A. Ruszczyński and A. Shapiro, Vol. 10 of *Handbooks in OR and MS*, Elsevier, 2003, pp. 353 – 425.

[49] Levi, R., Roundy, R., and Shmoys, D., "Provably Near-Optimal Sampling-Based Policies for Stochastic Inventory Control Models," *Mathematics of Operations Research*, Vol. 32, No. 4, 2007, pp. 821–839.

[50] Shapiro, A. and Nemirovski, A., "On complexity of stochastic programming problems," *Continuous Optimization: Current Trends and Applications*, edited by V. Jeyakumar and A. Rubinov, Springer, 2005, pp. 111–144.

[51] Espinoza, D. and Morena, E., "A primal-dual aggregation algorithm for minimizing Conditional-Value-at-Risk in linear programs," 2013, Working paper.

[52] Ausubel, L. M., Cramton, P., Pycia, M., Rostek, M., and Weretka, M., "Demand Reduction and Inefficiency in Multi-Unit Auctions," *Review of Economic Studies*, Vol. 81, No. 4, 2014, pp. 1366–1400.

[53] Lehmann, D., O'Callaghan, L., and Shoham, Y., "Truth revelation in approximately efficient combinatorial auctions," *Journal of the ACM*, Vol. 49, No. 5, SEP 2002, pp. 577–602.

[54] Zhan, R., "Optimality and Efficiency in Auctions Design: A Survey," *PARETO OPTIMALITY, GAME THEORY AND EQUILIBRIA*, edited by Chinchuluun, A and Pardalos, PM and Migdalas, A and Pitsoulis, L, Vol. 17 of *Springer Series in Optimization and Its Applications*, 2008, pp. 437–454.

[55] Deshpande, V. and Schwarz, L., "Optimal capacity choice and allocation in decentralized supply chains," Tech. rep., 2002, Working paper.

[56] Kastl, J., "On the properties of equilibria in private value divisible good auctions with constrained bidding," *Journal of Mathematical Economics*, Vol. 48, No. 6, DEC 2012, pp. 339–352.

[57] Parkes, D., "Auction design with costly preference elicitation," *Annals of Mathematics and Artificial Intelligence*, Vol. 44, No. 3, JUL 2005, pp. 269–302.

[58] Kushilevitz, E., "Communication Complexity," Vol. 44 of *Advances in Computers*, Elsevier, 1997, pp. 331 – 360.

[59] Kenyon, G., Meixell, M., and Westfall, P., "Production outsourcing and operational performance: An empirical study using secondary data," *International Journal of Production Economics*, Vol. 171, No. 3, SI, 2016, pp. 336–349.

[60] Prahalad, C. and Doz, Y., *The Multinational Mission: Balancing Local Demands and Global Vision*, New York: Free Press, 1987.

[61] de Kok, T., "Capacity allocation and outsourcing in a process industry," *International Journal of Production Economics*, Vol. 68, No. 3, 2000, pp. 229–239.

[62] Nazari-Shirkouhi, S., Eivazy, H., Ghodsi, R., K. Rezaie, K., and Atashpaz-Gargari, E., "Solving the integrated product mix-outsourcing problem using the Imperialist Competitive Algorithm," *Expert Systems with Applications*, Vol. 37, No. 12, 2010, pp. 7615–7626.

[63] Kermani, M., Aliahmadi, A., Salamat, V., Barzinpour, F., and Hadiyan, E., "Supplier selection in a single-echelon supply chain with horizontal competition using Imperialist competitive algorithm," *International Journal of Computer Integrated Manufacturing*, Vol. 28, No. 6, 2015, pp. 628–638.

[64] Vrat, P., "Make or Buy Decisions: Outsourcing Strategy," *Materials Management: An Integrated Systems Approach*, Springer Texts in Business and Economics, 2014, pp. 303–316.