ROBUST INVENTORY OPTIMIZATION

SEE CHUEN TECK

NATIONAL UNIVERSITY OF SINGAPORE

2009

ROBUST INVENTORY OPTIMIZATION

SEE CHUEN TECK

(MSc. NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF

PHILOSOPHY

DEPARTMENT OF INDUSTRIAL AND

SYSTEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible. First and foremost, this work would not have been possible without the guidance of my supervisors Associate Professor Poh Kim Leng and Associate Professor Melvyn Sim, under whose supervision I chose the topic and began the thesis. I was particularly inspired by Melvyn's enthusiasm and passion for research, his clear explanation of theory, his excellent teaching and good company. I would also like to thank to Mr Lee Yeaw Lip for his support, and A/P Poh for providing me with an excellent and conducive environment to carry out this research. I cannot end without thanking my family, on whose constant encouragement I have relied throughout the course of the research. I dedicate this thesis to them.

Contents

1 Introduction **1**

1.1 Motivation . . . 2

1.2 Organization of Thesis . . . 3

1.3 Notation . . . 3

2 Literature Review **4**

2.1 Second-order Cone Programming . . . 4

2.2 Multiperiod Inventory Model . . . 7

2.3 Robust Inventory Models . . . 8

2.4 Fill Rate Models . . . 11

3 Robust Optimization of Multiperiod Inventory **13**

3.1 Stochastic Inventory Model . . . 13

3.1.1 Factor-based Demand Model . . . 17

3.2 Robust Inventory Model . . . 18

3.2.1 Bound on $E((\cdot)^+)$ . . . 21

3.2.2 Tractable Replenishment Policies . . . 25

3.3 Extensions . . . 31

3.3.1 Fixed Ordering Cost . . . 31

3.3.2 Supply Chain Networks . . . 32

4 Computational Studies of Multiperiod Inventory Model **34**

4.1 Experimental Setup . . . 35

4.2 Comparison of Policies . . . 38

4.3 Influence of Directional Deviations . . . 43

4.4 Effects of Demand Variability . . . 43

4.5 Analysis of Policies 44
4.6 Computational Time 48
4.7 Summary 51

5 Safeguarding Fill Rate Against Distributional Uncertainty 54

5.1 Multiproduct Fill Rate Model 55
5.1.1 Definition of Fill Rate 55
5.1.2 Distributional Uncertainty 57
5.1.3 Descriptive Statistics & Uncertainty Sets 59
5.2 Fill Rate Bounds 61
5.2.1 Bounds using the Pdf Range 61
5.2.2 Bounds using the Median 63
5.2.3 Bounds using Modal Information 66

6 Computational Studies of Fill Rate Model 69

6.1 Technical Tests 69
6.2 Effectiveness of Robust Fill Rate Model 73
6.3 An Example with Real-life Demand 76
6.4 Summary 80

7 Conclusions 82

7.1 Contributions 83
7.1.1 Multiperiod Inventory Model 83
7.1.2 Fill Rate Model 84
7.2 Concluding Remarks 85

A Proof of Proposition 3 92

B Proof of Theorem 2 94

C Proof of Theorem 3 95

D Proof of Theorem 4 97

E Proof of Theorem 5 100

F Sample Formulation in *PROF* 104

G Raw Data of Example **106**

Summary

This thesis proposes methodologies to optimize uncertain inventory in a robust manner for two important settings. The first is a multiperiod inventory control problem where we trade-off the cost of holding excess inventory against the cost of backlog under ambiguous demands. The second setting is a service-level scenario where we propose bounds to guarantee a high level of expected fill rate against demands where the distribution is uncertain.

More specifically, the front half of this thesis proposes a robust optimization approach to address a multiperiod inventory control problem under ambiguous demands, where only limited information of the demand distributions such as mean, support and some measures of deviations are available. The approach is developed around a factor-based model, which has the ability to incorporate business factors as well as time-series forecast of trend, seasonality and cyclic variations. We obtain the parameters of the replenishment policies by solving a tractable deterministic optimization problem in the form of a second-order cone optimization problem (SOCP), with solution time; unlike dynamic programming approaches, is polynomial and independent on parameters such as replenishment lead time, demand variability, and correlations. The proposed truncated linear replenishment policy, which is piecewise-linear with respect to demand history, improves upon static and linear policies and achieves objective values that are reasonably close to optimal.

While traditional fill rate optimization of inventory assumes a known distribution, in reality demand distributions are seldom known exactly, only approximately. This is the motivation for the latter half of the thesis where we propose an approach to optimize fill rate using descriptive statistics so as to assure that a high fill rate is achieved even when there is distributional uncertainty. That is, the order quantity needs to achieve an expected fill rate target for a family of distributions with the same demand range, demand median and range of the probability density function. We develop bounds for the expected fill rate function, which enables the multiproduct problem to be approximated by linear programming formulation.

List of Tables

3.1 Forward and backward deviation of some common probability distributions 19

3.2 Directional deviations for truncated exponential variable with support $[0, \bar{z}]$. 20

4.1 Performance of truncated linear replenishment policy $T = 5$ 39

4.2 Performance of truncated linear replenishment policy $T = 10$ . . . 40

4.3 Performance of truncated linear replenishment policy $T = 20$ . . . 41

4.4 Performance of truncated linear replenishment policy $T = 30$ . . . 42

4.5 Performance of truncated linear replenishment policy $T = 5$ with and without directional deviations 43

4.6 Performance of truncated linear replenishment policy $T = 5$ with various demand range 45

4.7 Performance of truncated linear replenishment policy under a 2-point demand distribution 46

4.8 A sample path of the truncated linear replenishment policy 47

4.9 TLRP and LRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 40, S_t = 240$ 48

4.10 TLRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 80, S_t = 240$ 49

4.11 LRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 80, S_t = 240$ 50

4.12 Size of the TLRP model, where $\mathcal{L}^n = \{(x_0, \boldsymbol{x}) \in \Re \times \Re^{n-1} : \|\boldsymbol{x}\|_2 \leq x_0\}$. 52

4.13 Computational time 52

5.1 Errors in the estimation of distribution (obtained from simulation) . 58

6.1 Robust fill rate model versus traditional approach 77

6.2 Monthly electricity generation (descriptive statistics) 79

F.1 Formulation of Problem (3.17) using *PROF* 105

G.1 Monthly electricity generation data 107

List of Figures

3.1 Comparing the bounds of $\mathrm{E}((\tilde{z}-a)^+)$ . . . 25

5.1 Uncertainty in distribution type . . . 58
5.2 Two possible distributions . . . 60
5.3 Fill rate cases . . . 63
5.4 Demand cases . . . 64
5.5 Fill rate for three demand distributions with mean of 25 . . . 64
5.6 Fill rate cases with median constraint . . . 65
5.7 Unimodal demand . . . 66

6.1 Case 1 - $h_1(t) = 1$, $h_2(t) = 2t$ . . . 70
6.2 Case 2 - $h_1(t) = 2t$, $h_2(t) = 3t^2$ . . . 71
6.3 Case 3 - $h_1(t) = 2(1-t)$, $h_2(t) = 3(1-t)^2$ . . . 71
6.4 Results with unimodal Information . . . 72
6.5 Fill rate with buffer to the pdf ranges . . . 73
6.6 Histogram of Expected Fill Rate Achievements . . . 76
6.7 Monthly electricity generation (histogram) . . . 78
6.8 Fill rate of electricity demand . . . 79
6.9 Aggregate fill rate of electricity demand . . . 80

Chapter 1

Introduction

Inventory management has been an area of active research, even before operations research emerged as a scientific discipline. In particular, Zipkin (2000) attributed the advent of modern inventory theory to the economic order quantity formula which was proposed about 100 years ago by Harris (1913). He mentioned that Raymond (1931) is the first published book on inventory management. Notice that these developments pre-dated Dantzig's seminal work on linear programming. Graves et al. (1993) commented that even with the long and fruitful history, the major issues and problems in the area of inventory management have not all been resolved. They expect research on the area to continue to flourish, namely because of the following factors.

- "First, many logistics systems are extremely complex. It is difficult, both for researchers and managers, to achieve a clear, coherent picture of how such systems work. It has sometimes taken decades to obtain satisfactory solutions to the technical problems inherent in these systems, and even now many such problems remain open. Moreover, we continue to witness the development of fundamentally new approaches to the subject and a lively debate over the basic terms, premises and issues."

- "Second, the practical world of logistics has changed markedly over the past few decades. . . . Even more striking has been the explosion of information technologies, which has utterly transformed the very nature of logistics management."

1.1 Motivation

Inventory ties up working capital and incurs holding costs, reducing profit every day that excess stock is held. Good inventory management is hence crucial to businesses as they seek to continually improve their customer service and profit margins, in the light of global competition and demand variability. Baldenius and Reichelstein (2005) offered perhaps the most convincing study on the contribution of good inventory management to profitability. They studied inventories of publicly traded American manufacturing companies between 1981 and 2000, and they concluded: "Firms with abnormally high inventories have abnormally poor long-term stock returns. Firms with slightly lower than average inventories have good stock returns, but firms with the lowest inventories have only ordinary returns."

The ability to incorporate more realistic assumptions about product demand into inventory models is one key factor to profitability. Practical models of inventory need to address the issue of demand forecasting while ensuring sufficient robustness against uncertainty and maintaining tractability. In most industrial contexts, demand is uncertain. Many demand histories have factors that behave like random walks that evolve over time with frequent changes in directions and rates of growth or decline. In practice, for such demand processes, inventory managers rely on forecasts based on a time-series of prior demands, which are often correlated over time. For example, a product demand may depend on factors such as market outlook, oil prices, and so forth, and contains effects of trend, seasonality, cyclic variation and randomness.

Motivated by the need to explicitly address uncertainty, this thesis proposes methodologies to optimize uncertain inventory in a robust manner for two important settings. The first is a multiperiod inventory control problem where we trade-off the cost of holding excess inventory against the cost of backlog under ambiguous demands. The second setting is a service-level scenario where we propose bounds to guarantee a high level of expected fill rate against demands where the distribution is uncertain. The latter is motivated by the fact that in most practical settings, the distribution of demand is seldom known exactly but only approximately, and therefore it is difficult to derive the expected fill rate function. In particular, given a set of empirical data, it is common to find not one but several possible fits to the distribution. Uncertainty in the type of demand distribution may also arise with changing trends. For example, many consumer goods are known to exhibit seasonal variation with demand distributed differently over time.

1.2 Organization of Thesis

The thesis is organized as follows. We begin with a review of the relevant literature in Chapter 2. In Chapter 3 we describe the model to optimize multiperiod inventory robustly. This is followed by computational experiments in Chapter 4. Chapter 5 describes the approach to safeguard fill rate against distributional uncertainty. It is followed by the computational experiments in Chapter 6. We then conclude the thesis in Chapter 7 with a summary of our contributions.

1.3 Notation

Throughout this thesis, a random variable is denoted with the tilde sign such as $\tilde{y}$. We denote vectors with bold face lower-case letters such as $\boldsymbol{y}$ and matrices with bold face upper-case such as $\boldsymbol{A}$. We use $\mathbf{y}'$ to denote the transpose of vector $\mathbf{y}$. Also, denote $y^+ = \max(y, 0)$, $y^- = \max(-y, 0)$, and $\|\boldsymbol{y}\|_2 = \sqrt{\sum y_i^2}$. We generally use the bar and underline signs to denote the range or bound of a variable. For example, the upper and lower support of random variable $\tilde{d}$ are denoted as $\underline{d}$ and $\bar{d}$, respectively. We use $m(\tilde{d})$ to denote the median of the random variable $\tilde{d}$.

Chapter 2

Literature Review

This chapter contains a review of the relevant literature. Specifically, we highlight previous work that is directly related to our models. Before proceeding with the literature review proper, we provide some background material on second-order cone programming and discuss how it can be used to represent robust optimization problems. The material originated from Boyd and Vandenberghe (2004) and is included to facilitate the exposition of the multiperiod inventory control models, which are essentially second-order cone programs.

2.1 Second-order Cone Programming

Conic optimization problems are a class of convex nonlinear optimization problems, lying between linear programming (LP) problems and general convex nonlinear problems. A conic optimization problem can be written as an LP (with a linear objective and linear constraints) plus one or more cone constraints. A cone constraint specifies that the vector formed by a set of decision variables is constrained to lie within a closed convex pointed cone. The simplest example of such a cone is the nonnegative orthant, the region where all variables are nonnegative: the normal situation in an LP. Conic optimization allows for more general cones, with second-order cone being the more common case.

A second-order cone program (SOCP) takes the following form:

$$\begin{aligned} \min \quad & \boldsymbol{f}'\boldsymbol{x} \\ \text{s.t.} \quad & ||\boldsymbol{A}_i\boldsymbol{x} + \boldsymbol{b}_i||_2 \leq \boldsymbol{c}_i'\boldsymbol{x} + d_i, \qquad i = 1 \dots m, \\ & \boldsymbol{F}\boldsymbol{x} = \boldsymbol{g}, \end{aligned}$$

where $\boldsymbol{x} \in \Re^n$ is the decision variable, $\boldsymbol{A}_i \in \Re^{n_i \times n}$, and $\boldsymbol{F} \in \Re^{p \times n}$. We call a constraint of the form

$$||A\boldsymbol{x} + \boldsymbol{b}||_2 \leq \boldsymbol{c}'\boldsymbol{x} + d$$

where $\boldsymbol{A} \in \Re^{k \times n}$, a *second-order cone constraint*, since it is the same as requiring the affine function $(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{c}'\boldsymbol{x} + d)$ to lie the second-order cone in $\Re^{k+1}$. When $\boldsymbol{c}_i = \boldsymbol{0},\ i = 1 \dots m$, the SOCP is equivalent to a quadratic constrained quadratic programming (which is obtained by squaring each of the constraints). Similarly if $\boldsymbol{A}_i = \boldsymbol{0},\ i = 1 \dots m$, then the SOCP reduces to a linear program. Second-order cone programs are more general than quadratic constrained quadratic programming (and of course linear programs). SOCPs are known to be tractable, and can be solved with great efficiency by interior point methods, see for instance Boyd and Vandenberghe (2004). A number of commercial solvers are able to solve SOCP efficiently. The noteworthy ones are CPLEX, LOQO, MOSEK and PENSDP.

Consider a quadratic constraint on the form

$$\boldsymbol{x}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}'\boldsymbol{x} + c \leq 0.$$

This is equivalent to the SOC constraint

$$\left\| \begin{array}{c} (1 - \boldsymbol{b}'\boldsymbol{x} - c)/2 \\ \boldsymbol{A}\boldsymbol{x} \end{array} \right\|_2 \leq (1 - \boldsymbol{b}'\boldsymbol{x} - c)/2.$$

We now provide two examples on how SOCP can be used to represent robust optimization problems.

(a) Robust linear programming. We consider a linear program in inequality form,

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{a}_i'\boldsymbol{x} \leq b_i, \qquad i = 1 \dots m, \end{aligned}$$

in which there is some uncertainty or variation in the parameters $\boldsymbol{c}, \boldsymbol{a}_i, b_i$. To simplify the exposition we assume that $\boldsymbol{c}$ and $\boldsymbol{b}_i$ are fixed, and that $\boldsymbol{a}_i$ are known to lie in given ellipsoids:

$$\boldsymbol{a} \in \mathcal{E}_i = \{\boldsymbol{a}_i + \boldsymbol{P}_i\boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leq 1\},$$

where $\boldsymbol{P}_i \in \Re^{n \times n}$. (If $\boldsymbol{P}_i$ is singular we obtain 'flat' ellipsoids, of dimension rank $\boldsymbol{P}_i$; $\boldsymbol{P}_i = \mathbf{0}$ means that $\boldsymbol{a}_i$ is known perfectly.) We will require that the constraints be satisfied for all possible values of the parameters $\boldsymbol{a}_i$, which leads us to the robust linear program

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{a}_i'\boldsymbol{x} \leq b_i, \qquad \forall \boldsymbol{a}_i \in \mathcal{E}_i \qquad i = 1 \ldots m. \end{array} \tag{2.1}$$

The robust linear constraint, $\boldsymbol{a}_i'\boldsymbol{x} \leq b_i \; \forall \boldsymbol{a}_i \in \mathcal{E}_i$, can be expressed as

$$\sup\{\boldsymbol{a}_i'\boldsymbol{x} \mid \boldsymbol{a}_i \in \mathcal{E}_i\} \leq b_i,$$

the lefthand side of which can be expressed as

$$\begin{aligned} \sup\{\boldsymbol{a}_i'\boldsymbol{x} \mid \boldsymbol{a}_i \in \mathcal{E}_i\} &= \boldsymbol{a}_i'\boldsymbol{x} + \sup\{\boldsymbol{u}'\boldsymbol{P}_i'\boldsymbol{x} \mid \|\boldsymbol{u}\|_2 \leq 1\} \\ &= \boldsymbol{a}_i'\boldsymbol{x} + \|\boldsymbol{P}_i\boldsymbol{x}\|_2. \end{aligned}$$

Thus, the robust linear constraint can be expressed as

$$\boldsymbol{a}_i'\boldsymbol{x} + \|\boldsymbol{P}_i\boldsymbol{x}\|_2 \leq b_i,$$

which is evidently a second-order cone constraint. Hence Problem (2.1) can be expressed as the SOCP

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{a}_i' + \|\boldsymbol{P}_i\boldsymbol{x}\|_2 \leq b_i, \qquad i = 1 \ldots m. \end{array}$$

Note that the additional norm terms act as regularization terms; they prevent $\boldsymbol{x}$ from being large in directions with considerable uncertainty in the parameters $\boldsymbol{a}_i$.

(b) Linear programming with random constraints. The robust LP described above can also be considered in a statistical framework. Here we suppose that the parameters $\boldsymbol{a}_i$ are independent Gaussian random vectors, with mean $\bar{\boldsymbol{a}}_i$ and covariance $\boldsymbol{\Sigma}_i$. We require that each constraint $\boldsymbol{a}_i'\boldsymbol{x} \leq b_i$ should hold with a probability (or confidence) exceeding η, where $\eta \geq 0.5$, that is,

$$\text{prob}(\boldsymbol{a}_i'\boldsymbol{x} \leq b_i) \geq \eta. \tag{2.2}$$

We will show that this probability constraint can be expressed as a second-order cone constraint. Letting $u = \boldsymbol{a}_i'\boldsymbol{x}$, with σ^2 denoting its variance, this constraint can be written as

$$\text{prob}\left(\frac{u-\bar{u}}{\sigma} \leq \frac{b_i - \bar{u}}{\sigma}\right) \geq \eta.$$

Since $(u-\bar{u})/\sigma$ is a zero mean unit variance Gaussian variable, the probability above is simply $\Phi((b_i - \bar{u}/\sigma))$, where $\Phi(z)$ is the cumulative distribution function of a zero mean unit variance Gaussian random variable. Thus the probability constraint (2.2) can be expressed as

$$\frac{b_i - \bar{u}}{\sigma} \geq \Phi^{-1}(\eta),$$

or equivelently,

$$\bar{u} + \Phi^{-1}(\eta)\sigma \geq b_i.$$

From $\bar{u} = \bar{\boldsymbol{a}}_i'\boldsymbol{x}$ and $\sigma = (\boldsymbol{x}'\boldsymbol{\Sigma}_i\boldsymbol{x})^{1/2}$ we obtain

$$\bar{\boldsymbol{a}}_i'\boldsymbol{x} + \Phi^{-1}(\eta)\|\boldsymbol{\Sigma}_i^{1/2}\boldsymbol{x}\| \leq b_i.$$

By our assumption that $\eta \geq 1/2$, we have $\Phi^{-1}(\eta) \geq 0$, so this constraint is a second-order cone constraint. In summary, the problem

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \text{prob}(\tilde{\boldsymbol{a}}_i'\boldsymbol{x} \geq b_i) \geq \eta, \qquad i = 1\ldots m, \end{aligned}$$

can be expressed as the SOCP

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \bar{\boldsymbol{a}}_i'\boldsymbol{x} + \Phi^{-1}(1-\eta)\|\boldsymbol{\Sigma}_i^{1/2}\boldsymbol{x} \geq b_i\|_2 \qquad i = 1\ldots m. \end{aligned}$$

This concludes our exposition on second-order cone programming. For more applications of SOCP for robust optimization, the interested readers can refer to Boyd and Vandenberghe (2004).

2.2 Multiperiod Inventory Model

The multiperiod inventory control problem is a well-studied problem in operations research. For the single-product inventory control problem with history-independent demands, it is well known that the base-stock policy based on a critical fractile is optimum. See Scarf (1959, 1960), Azoury (1985), Miller (1986) and Zipkin (2000). For correlated demands, Veinott (1965) characterized conditions under which a myopic policy is optimal. Extending the results of Veinott, Johnson and Thompson (1975) considered an auto-regressive, moving-average (ARMA) demand process, zero replenishment lead time and no backlogs, and showed the optimality of a myopic policy when the demand in each period is bounded. Lovejoy (1990) showed that a myopic critical-fractile policy is optimum or near optimum in some inventory models with adaptive demand processes, citing exponential smoothing on the demand process and Bayesian updating on uniformly distributed demand as examples. Song and Zipkin (1993) addressed the case of Poisson demand, where the transition rate between states is governed by a Markov process.

Although optimum policies can be characterized in many interesting variants of inventory control problems, it is not easy to compute them efficiently, that is, in polynomial time with respect to the input size of the problem. In this thesis, the term *tractable replenishment policy* is used if the parameters of the policy are polynomial in size and can be obtained in polynomial time. For instance, the celebrated optimum base-stock policy may not necessarily be a tractable one. Sampling-based approximation has been applied to the inventory control problem; see, for instance, Levi et al. (2007). Using marginal cost accounting and cost-balancing techniques, Levi et al. (2007) proposed an elegant two-approximation algorithm for the inventory control problem. However, there is a lack of computational studies demonstrating the effectiveness of the approximation algorithm. Other sampling-based approaches include infinitesimal perturbation analysis (see Glasserman and Tayur (1995)), which uses stochastic gradient estimation technique, and the concave adaptive value estimation procedure, which successively approximates the objective cost function with a sequence of piecewise-linear functions (see Godfrey and Powell (2001) and Powell et al. (2004)). More recently, Iida and Zipkin (2006) and Lu et al. (2006) developed approximate solutions for demand following the martingale model of forecast evolution.

2.3 Robust Inventory Models

One of the fundamental assumptions of stochastic models, which has recently been challenged, is the availability of probability distributions in characterizing the uncertain parameters. Bertsimas and Thiele (2006) illustrated that an optimum inventory control policy that is heavily tuned to a particular demand distribution may perform poorly against another demand distribution bearing the same mean and variance. Assuming a demand distribution tacitly implies that we are able to obtain exact estimates of all the moments, which is practically prohibitive. It is a common practice to estimate the first two moments from data and fit the parameters to an assumed distribution. By doing so, we artificially extrapolate the rest of the moments using only the information from lower partial moments. Errors in estimating the first two moments will naturally propagate to higher moments. Therefore, it is not surprising that policies derived from assuming demand distributions may be less robust. One approach to account for distributional ambiguity is to consider a family of demand distributions, which can be characterized by their descriptive statistics such as partial moment information, support and so forth. Research on inventory control under ambiguous demand distributions dates back to Scarf (1958), where he considered a newsvendor problem and determined orders that minimized the maximum expected cost over all possible demand distributions with the same first and second moments and with nonnegative support. Various extensions of Scarf's single-period results have been studied by Gallego and Moon (1993). Although the solutions to these single-period models are in the form of a second-order cone optimization problem (SOCP), which are polynomial-time solvable, the minimax approach does not scale well computationally with the number of periods. Nevertheless, the optimum policies for multiperiod inventory control problems under various forms of demand ambiguity have been characterized by Kasugai and Kasegai (1960) and Gallego et al. (2001).

In recent years, robust optimization has seen an explosive growth and has become a dominant approach to address the optimization problem under uncertainty. Traditionally, the goal of robust optimization is to immunize uncertain mathematical optimization problems against infeasibility while preserving the tractability of the models. See, for instance, Ben-Tal and Nemirovski (1998, 1999, 2000), Bertsimas and Sim (2003, 2004), Bertsimas et al. (2004, 2009), El-Ghaoui and Lebret (1997), and El-Ghaoui et al. (1998). Many robust optimization approaches have the following two important characteristics:

(a) The model of data uncertainty in robust optimization permits distributional ambiguity. Data uncertainty can also be completely distributional free and specified by an uncertainty set parameterized by the "Budget of Uncertainty", which controls the size of the uncertainty set. Another model of uncertainty considers uncertain parameters whose distributions are unknown but are confined to a family of distributions that would generate the same descriptive statistics on the data, such as known means and variances.

(b) The solution (or approximate solution) to a robust optimization model can be obtained by solving a tractable deterministic mathematical optimization problem such as SOCP, whose associated solvers are commercially available, robust and efficiently optimized. Robust optimization methodology often decouples model formulation from the optimization engine, which enables the modeler to focus on modeling the actual problem and not to be hindered by algorithm design.

Based on the framework of robust optimization, Bertsimas and Thiele (2006) developed a new approach to address demand ambiguity for a multiperiod inventory control problem, which has the advantage of being computationally tractable. They considered a family of demand distributions similar to Scarf and enforced independence across time periods. Bertsimas and Thiele mapped the demand uncertainty model into a "Budget of Uncertainty" model of Bertsimas and Sim (2004) and proposed an open-loop inventory control approach in which the solutions can be obtained by solving a tractable linear optimization problem. They showed that the optimum solution of their robust model has a base-stock structure and the tractability of the problem readily extends to problems with capacity constraints and over a supply chain network, and their paper characterizes the optimum policies for these cases. The analysis of the robust models and computational experiments for independent demands suggests that robust approaches compare well against an optimum model under exact distribution and is yet robust against distributional ambiguity. Using a similar approach, Adida and Perakis (2006) proposed a deterministic robust optimization formulation for dealing with demand uncertainty in a dynamic pricing and inventory control problem for a make-to-stock manufacturing system. They developed a demand-based fluid model and showed that it is no more difficult to solve the robust formulation than it is to solve the nominal problem.

Other related work in the robust inventory control literature includes Bienstock and Ozbay (2008), where they proposed a robust model focusing on base-stock policy structure. Song et al. (2007) adopted a data-driven approach to robust inventory management.

To address the inadequacy of open-loop robust optimization models involving multistage decision processes, Ben-Tal et al. (2004) introduced the concept of adjustable robust counterpart, which permits decisions to be delayed until information is available. Unfortunately, with the additional flexibility in modeling, adjustable robust counterpart models are generally NP-hard, and the authors have proposed and advocated the use of linear decision rule as a tractable approximation. Ben-Tal et al. (2005) applied their model to a multiperiod inventory control problem and showed, by means of computational studies the advantages of the linear replenishment policy over the open-loop model which had a static replenishment policy. We emphasize that in contrast to stochastic models, the uncertainty considered in adjustable robust counterpart is completely distribution free, that is, the data uncertainty is characterized only by its support.

To bridge the gap between robust optimization and stochastic models, Chen et al. (2007) introduced the notions of directional deviations known as *forward and backward deviations*. They also proposed computationally tractable robust optimization models for immunizing linear optimization problems against infeasibility, which enhanced the modeling power of robust optimization in the characterization of ambiguous distributions. In a parallel work, Chen et al. (2008) proposed several piecewise-linear decision rules for approximating stochastic linear optimization problems that improve upon linear rules. These approaches have been unified by Chen and Sim (2009), where they proposed a general family of distributions characterized by the mean, covariance, directional deviations and support and showed how it can be extended to approximate the solution for a two-period stochastic model under a satisficing objective.

2.4 Fill Rate Models

On another front, while many classical inventory models are cost-based approaches trading off holding excess inventory against the penalty of shortage, service level approaches based on fully known distributions have also been popular. See for instance Chen and Krass (2001), Lee and Billinton (1992, 1995); Taylor (1997). Sherbrooke (1992) describes the use of fill rate as a measure of performance in inventory management of spare part. Fill rate models in the literature typically assume that the cumulative distribution function of demand is known. This is seen in, for example, Schwarz et al. (1985) and Ding et al. (2006). Distribution-free models involving fill rate include Song (1998) in which she developed bounds on order quantity using partial information such as first and second moments, and Agrawal and Seshadri (2000). While there are many distribution-free approaches in the literature, we are unaware of any counterpart that directly estimates the expected fill rate using only descriptive statistics. In some industries, the ability to incorporate service level considerations robustly into inventory models is crucial to good inventory management. It seems that models which address the issue of achieving high service standards while staying sufficiently immunized against uncertainty appear to be lacking. The model of Chapter 5 aims to address this gap.

Chapter 3

Robust Optimization of Multiperiod Inventory

We begin the exposition by describing the multiperiod inventory problem in detail. This is followed by a discussion of our robust inventory model. We then close the chapter with some possible extensions.

3.1 Stochastic Inventory Model

The stochastic inventory model involves the derivation of replenishment decisions over a discrete planning horizon consisting of a finite number of periods under stochastic demand. The demands in period $t = 1 \dots N$ form a sequence of random variables that are *not* necessarily identically distributed and *not* necessarily independent. We consider an inventory system with T planning horizons from $t = 1$ to $t = T$. External demands arrive at the inventory system and the system replenishes its inventory from some central warehouse (or supplier) with ample supply. The timeline of events is as follows.

1. At the beginning of the tth time period, before observing the demand, the inventory manager places an order of x_t at unit cost c_t for the product to arrive after a (fixed) order lead time of L periods. Orders placed at the *beginning* of the tth time period will arrive at the *beginning* of $t+L$th period. We assume that replenishment ceases at the end of the planning horizon, so that the last order is placed in period $T - L$. Without loss of generality, we

assume that purchase costs for inventory are charged at the time of order. The case where purchase costs are charged at the time of delivery can be represented by a straightforward shift of cost indices.

2. At the beginning of each time period t, the inventory manager faces an initial inventory level y_t and receives an order of x_{t-L}. The demand of inventory for the period is realized at the end of the time period. After receiving a demand of d_t, the inventory level at the end of the period is $y_t + x_{t-L} - d_t$.

3. Excess inventory is carried to the next period, incurring a per-unit overage (holding) cost. On the other hand, each unit of unsatisfied demand is backlogged (carried over) to the next period with a per-unit underage (backlogging) penalty cost. At the last period, $t = T$, the penalty of lost sales can be accounted through the underage cost.

We assume an inventory manager whose objective is to determine the dynamic ordering quantities x_t from period $t = 1$ to period $t = T - L$ so as to minimize the total expected ordering, inventory overage (holding), and inventory underage (backlog) costs in response to the uncertain demands. Observe that for $L \geq 1$, the quantities x_{t-L}, $t = 1, \ldots, L$ are known values. They denote orders made before period $t = 1$ and are inventories in the delivery pipeline when the planning horizon starts.

We introduce the following notation:

- $\tilde{d}_t$: stochastic exogenous demand at period t
- $\tilde{\boldsymbol{d}}_t$: a vector of random demands from period 1 to t, that is, $\tilde{\boldsymbol{d}}_t = (\tilde{d}_1, \ldots, \tilde{d}_t)$
- $x_t(\tilde{\boldsymbol{d}}_{t-1})$: order placed at the beginning of the tth time period after observing $\tilde{\boldsymbol{d}}_{t-1}$. The first-period inventory order is denoted by $x_1(\tilde{\boldsymbol{d}}_0) = x_1^0$
- $y_t(\tilde{\boldsymbol{d}}_{t-1})$: inventory level at the beginning of the tth time period
- h_t: unit inventory overage (holding) cost charged on excess inventory at the end of the tth time period
- b_t: unit underage (backlog) cost charged on backlogged inventory at the end of the tth time period

- c_t: unit purchase cost of inventory for orders placed at the tth time period
- S_t: the maximum amount that can be ordered at the tth time period.

We use $x_t(\tilde{\boldsymbol{d}}_{t-1})$ to represent the nonanticipative replenishment policy at the beginning of period t. That is, the replenishment decision is based solely on the observed information available at the beginning of period t, which is given by the demand vector $\tilde{\boldsymbol{d}}_{t-1} = (\tilde{d}_1, \ldots, \tilde{d}_{t-1})$. Given the order quantity $x_{t-L}(\tilde{\boldsymbol{d}}_{t-L-1})$ and stochastic exogenous demand $\tilde{d}_t$, the inventory level at the *end* of the t time period (which is also the inventory level at start of $t+1$ time period) is given by

$$y_{t+1}(\tilde{\boldsymbol{d}}_t) = y_t(\tilde{\boldsymbol{d}}_{t-1}) + x_{t-L}(\tilde{\boldsymbol{d}}_{t-L-1}) - \tilde{d}_t, \qquad t = 1, \ldots, T. \tag{3.1}$$

In resolving the initial boundary conditions, we adopt the following notation:

- The initial inventory level of the system is $y_1(\tilde{\boldsymbol{d}}_0) = y_1^0$.
- When $L \geq 1$, the orders that are placed before the planning horizon starts are denoted by
$$x_t(\tilde{\boldsymbol{d}}_{t-1}) = x_t^0, \qquad t = 1 - L, \ldots, 0.$$

Note that Equation (3.1) can be written using the cumulative demand up to period t and cumulative order received as follows:

$$y_{t+1}(\tilde{\boldsymbol{d}}_t) = \underbrace{y_1^0}_{\text{start inventory}} + \underbrace{\sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0}_{\text{committed orders}} + \underbrace{\sum_{\tau=L+1}^{t} x_{\tau-L}(\tilde{\boldsymbol{d}}_{\tau-L-1})}_{\text{order decisions}} - \underbrace{\sum_{\tau=1}^{t} \tilde{d}_\tau}_{\text{cumulative demands}}. \tag{3.2}$$

Observe that positive (respectively, negative) value of $y_{t+1}(\tilde{\boldsymbol{d}}_t)$ represents the total amount of inventory overage (respectively, underage) at the end of the period t after meeting demand. Thus, the total expected cost, including ordering, inventory overage, and inventory underage charges is equal to

$$\sum_{t=1}^{T} \left(\mathrm{E}\left(c_t x_t(\tilde{\boldsymbol{d}}_{t-1}) \right) + \mathrm{E}\left(h_t (y_{t+1}(\tilde{\boldsymbol{d}}_t))^+ \right) + \mathrm{E}\left(b_t (y_{t+1}(\tilde{\boldsymbol{d}}_t))^- \right) \right).$$

Therefore, the multiperiod inventory problem can be formulated as a T stage stochastic optimization model as follows:

$$\begin{array}{rll} Z_{STOC} = \min & \displaystyle\sum_{t=1}^{T}\Big(\mathrm{E}\left(c_t x_t(\tilde{\boldsymbol{d}}_{t-1})\right) + \mathrm{E}\left(h_t(y_{t+1}(\tilde{\boldsymbol{d}}_t))^+\right) + \mathrm{E}\left(b_t(y_{t+1}(\tilde{\boldsymbol{d}}_t))^-\right)\Big). & \\ \text{s.t.} & y_{t+1}(\tilde{\boldsymbol{d}}_t) = y_t(\tilde{\boldsymbol{d}}_{t-1}) + x_{t-L}(\tilde{\boldsymbol{d}}_{t-L-1}) - \tilde{d}_t \qquad t = 1,\ldots,T & \\ & 0 \le x_t(\tilde{\boldsymbol{d}}_{t-1}) \le S_t \qquad t = 1,\ldots,T-L. & \end{array} \tag{3.3}$$

The aim of the stochastic optimization model is to derive a feasible replenishment policy that minimizes the expected ordering and inventory costs. That is, we seek a sequence of action rules that advises the inventory manager of the action to take in time t as a function of demand history. Unfortunately, the decision variables in Problem (3.3), $x_t(\tilde{\boldsymbol{d}}_{t-1}), t = 1\ldots T-L$ and $y_t(\tilde{\boldsymbol{d}}_{t-1})$, $t = 2\ldots T+1$, are functionals, which means that Problem (3.3) is an optimization problem with an infinite number of variables and constraints, and hence generally intractable.

The stochastic optimization problem (3.3) can also be formulated as a dynamic programming problem. For simplicity, assuming zero lead time, the dynamic programming requires the following updates on the value function:

$$\begin{aligned} & J_t(y_t, d_1, \ldots, d_{t-1}) \\ = \; & \min_{x\in[0,S_t]} \mathrm{E}\Big(c_t x + G_t(y_t + x - \tilde{d}_t) + \\ & \qquad J_{t+1}(y_t + x - \tilde{d}_t, d_1, \ldots, d_{t-1}, \tilde{d}_t) \mid \tilde{d}_1 = d_1, \ldots, \tilde{d}_{t-1} = d_{t-1}\Big), \end{aligned}$$

where $G_t(u) = h_t \max(u, 0) + b_t \max(-u, 0)$. Maintaining the value function $J_t(\cdot)$ is computationally prohibitive, and hence most inventory control literature identify conditions such that the value functions are not dependent on past demand history, so that the state space is computationally amenable. For instance, it is well known that when the lead time is zero and the demands are independently distributed across time periods, there exist base-stock levels, q_t, such that the following replenishment policy,

$$x_t(\tilde{\boldsymbol{d}}_{t-1}) = \min\left\{\max\left\{q_t - y_t(\tilde{\boldsymbol{d}}_{t-1}), 0\right\}, S_t\right\} \tag{3.4}$$

is optimum. Hence, instead of being a function of the entire demand history, the optimum demand policy can be characterized by the inventory level as follows:

$$x_t(y_t) = \min\left\{\max\left\{q_t - y_t, 0\right\}, S_t\right\}.$$

$$J_t^{BSP}(y_t) = \min_{0 \le x \le S_t} \mathrm{E}_{\tilde{d}_t}\Big(c_t x + G_t(y_t + x - \tilde{d}_t) + J_{t+1}^{BSP}(y_t + x - \tilde{d}_t)\Big),$$

where $\tilde{d}_t = \tilde{z}_t + \alpha\tilde{z}_{t-1} + \alpha\tilde{z}_{t-2} + \cdots + \alpha\tilde{z}_1 + \mu$. The replenishment policy is given by

$$x_t^{BSP}(y_t) = \arg\min_{0 \le x \le S_t} \mathrm{E}_{\tilde{d}_t}\Big(c_t x + G_t(y_t + x - \tilde{d}_t) + J_{t+1}^{BSP}(y_t + x - \tilde{d}_t)\Big).$$

Under capacity limit on order quantities, the modified history-independent base-stock policy is optimum when the demands are independently distributed. This is discussed in Federgruen and Zipkin (1986).

3.1.1 Factor-based Demand Model

We adopt a *factor-based demand model* in which the uncertain demand is affinely dependent on zero mean random factors $\tilde{\boldsymbol{z}} \in \Re^N$ as follows:

$$d_t(\tilde{\boldsymbol{z}}) \triangleq \tilde{d}_t = d_t^0 + \sum_{k=1}^{N} d_t^k \tilde{z}_k, \qquad t = 1, \dots, T,$$

where

$$d_t^k = 0 \qquad \forall k \ge N_t + 1,$$

and $1 \le N_1 \le N_2 \le \cdots \le N_T = N$. Such an affine factor-based uncertainty model is a common assumption in robust optimization. See for instance, Ben-Tal and Nemirovski (1998). Under a factor-based demand model, the random factors, $\tilde{z}_k$, $k = 1, \dots, N$ are realized sequentially. At period t, the factors, $\tilde{z}_k$, $k = 1, \dots, N_t$ have already been unfolded. In progressing to period $t+1$, the new factors $\tilde{z}_k$, $k = N_t + 1, \dots, N_{t+1}$ are made available.

Demand that is affected by random noise or shocks can be represented by the factor-based demand model. For independently distributed demand, which is assumed in most inventory models, we have

$$d_t(\tilde{\boldsymbol{z}}) = d_t^0 + \tilde{z}_t, \qquad t = 1, \dots, T,$$

in which $\tilde{z}_t$ are independently distributed. However, in many industrial contexts, demands across periods may be correlated. In fact, many demand histories behave more like random walks over time, with frequent changes in directions and rate of growth or decline. See Johnson and Thompson (1975) and Graves (1999). In those settings, we may consider standard forecasting techniques such as an ARMA(p, q) demand process (see Box et al. (1994)) as follows:

$$d_t(\tilde{\boldsymbol{z}}) = \begin{cases} d_t^0 & \text{if } t \leq 0 \\ \displaystyle\sum_{j=1}^{p} \phi_j d_{t-j}(\tilde{\boldsymbol{z}}) + \tilde{z}_t + \sum_{j=1}^{\min\{q,t-1\}} \theta_j \tilde{z}_{t-j} & \text{otherwise,} \end{cases}$$

where $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ are known constants. Indeed, it is easy to show by induction that $d_t(\tilde{\boldsymbol{z}})$ can be expressed in the form of a factor-based demand model. Song and Zipkin (1993) presented a world-driven demand model where the demand is Poisson with rate controlled by finite Markov states representing different business environments. However, it may be difficult to determine exhaustively the business states and their state transition probabilities. On the other hand, factor-based models have been used extensively in finance for modeling returns as affine functions of external factors, in which the coefficients of the factors can be determined statistically. In the same way, we can apply the factor-based demand model to characterize the influence of demands with external factors such as market outlook, oil prices and so forth. Effects of trend, seasonality, cyclic variation, and randomness can also be incorporated.

3.2 Robust Inventory Model

The stochastic inventory control problem requires full information of the demand distributions, which is practically prohibitive. Furthermore, even if the probability distributions are known, due to computational complexity, we may not be able to obtain the optimum solution. Note that under the factor-based demand model, it is easy to evaluate the demand distribution when the factors are normally distributed. However, this is not necessarily the case for other distributions. Nemirovski and Shapiro (2006) noted that evaluating the distribution of a weighted sum of uniformly distributed independent random variables is already NP-hard. As such, it would generally be intractable to evaluate the cumulative distributions of the random demand with nonnormally distributed factors. Consequently, it would be technically intractable to compute the myopic critical fractile based on the seemingly benign factor-based demand model. The robust optimization approach we are proposing aims to address these issues collectively.

Table 3.1: Forward and backward deviation of some common probability distributions

Distribution	σ_f	σ_b
Normal with standard deviation, σ	σ	σ
Uniform with standard deviation, σ	σ	σ
Exponential with standard deviation, σ	∞	σ

Instead of assuming full distributions on the factors, which are practically prohibitive, we adopt a modest distributional assumption on the random factors, such as known means, supports, and some aspects of deviations. The factors may be partially characterized using the directional deviations, that were recently introduced by Chen et al. (2007).

Definition 1 *(Directional deviations) Given a random variable $\tilde{z}$, the forward deviation is defined as*

$$\sigma_f(\tilde{z}) \triangleq \sup_{\theta>0}\left\{\sqrt{2\ln(\mathrm{E}(\exp(\theta(\tilde{z}-E(\tilde{z})))))/\theta^2}\right\} \tag{3.5}$$

and backward deviation is defined as

$$\sigma_b(\tilde{z}) \triangleq \sup_{\theta>0}\left\{\sqrt{2\ln(\mathrm{E}(\exp(-\theta(\tilde{z}-E(\tilde{z})))))/\theta^2}\right\}. \tag{3.6}$$

Table 3.1 shows the forward and backward deviation of some common probability distributions. We also present in Table 3.2, the directional deviations of a truncated exponential random variable $\tilde{z}$ in $[0, \bar{z}]$ with the following density function:

$$f_{\tilde{z}}(z) = \frac{\exp(-z)}{1-\exp(-\bar{z})}.$$

Although the forward deviation of a pure exponential distributed random variable is infinite, the truncated exponential distribution has a reasonably small forward deviation compared to the support $\bar{z}$. Even when $\bar{z} = 10$, the forward deviation is only slightly more than twice its standard deviation.

Given a sequence of independent samples, we can essentially estimate the magnitude of the directional deviations from (3.5) and (3.6). Some of the properties of the directional deviations include:

Table 3.2: Directional deviations for truncated exponential variable with support $[0, \bar{z}]$.

$\bar{z}$	4	5	6	7	8	9	10	100
Standard deviation	0.834	0.911	0.954	0.977	0.989	0.995	0.998	1.000
σ_f	1.037	1.239	1.419	1.583	1.733	1.871	2.000	7.000
σ_b	0.834	0.911	0.954	0.977	0.989	0.995	0.998	1.000

Proposition 1 *(Chen et al. (2007))*
Let σ, p and q be, respectively, the standard, forward and backward deviations of a random variable $\tilde{z}$ with zero mean.
(a)

$$p \geq \sigma \qquad q \geq \sigma.$$

If $\tilde{z}$ is normally distributed, then $p = q = \sigma$.
(b) *For all $\theta \geq 0$,*

$$\mathrm{P}(\tilde{z} \geq \theta p) \leq \exp(-\theta^2/2);$$

$$\mathrm{P}(\tilde{z} \leq -\theta q) \leq \exp(-\theta^2/2).$$

Proposition 1(a) shows that the directional deviations are no less than the standard deviation of the underlying distribution, and under the normal distribution, these two values coincide with the standard deviation. As exemplified in Proposition 1(b), the directional deviations provide an easy bound on the distributional tails. The advantage of using the directional deviations is the ability to capture distributional asymmetry and stochastic independence, while keeping the resultant optimization model computationally amicable. We refer the reader to the paper by Natarajan et al. (2008) for the computational experience of using directional derivations derived from real-life data.

In this work, we adopt the random factor model introduced by Chen and Sim (2009), which encompasses most of the uncertainty models found in the literature of robust optimization.

Assumption U: We assume that the uncertainties $\{\tilde{z}_j\}_{j=1:N}$ are zero mean random variables, with positive definite covariance matrix, $\mathbf{\Sigma}$. We denote a subset,

$\mathcal{I} \subseteq \{1, \ldots, N\}$, which can be an empty set, such that $\tilde{z}_j$, $j \in \mathcal{I}$ are stochastically independent. Moreover, the corresponding forward and backward deviations are given by $p_j = \sigma_f(\tilde{z}_j)$ and $q_j = \sigma_b(\tilde{z}_j)$, respectively, for $j \in \mathcal{I}$ and that $p_j = q_j = \infty$ for $j \notin \mathcal{I}$[1]

Let $\mathcal{W}$ be a convex set containing the support of $\tilde{\boldsymbol{z}}$. The choice of the support set $\mathcal{W}$ can influence the computational tractability of the problem. Henceforth, we assume that the support set is a second order conic representable set (also known as conic quadratic representable set) proposed in Ben-Tal and Nemirovski (1998), which includes polyhedral and ellipsoidal sets. A common support set is the interval set, which is given by $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$, in which $\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}} > \mathbf{0}$.

For notational convenience, we define the following sets:

$$\begin{array}{ll} \mathcal{I}_1 \triangleq \{j \ : \ p_j < \infty\} & \bar{\mathcal{I}}_1 \triangleq \{j \ : \ p_j = \infty\} \\ \mathcal{I}_2 \triangleq \{j \ : \ q_j < \infty\} & \bar{\mathcal{I}}_2 \triangleq \{j \ : \ q_j = \infty\}. \end{array}$$

Furthermore, if $p_j = \infty$ (respectively, $q_j = \infty$), its product with zero remains zero, that is, $p_j \times 0 = 0$ (respectively, $q_j \times 0 = 0$).

3.2.1 Bound on $\mathbf{E}((\cdot)^+)$

In the absence of full distributional information, it would be meaningless to evaluate the optimum objective as depicted in Problem (3.3). Instead, we assume that the modeler is averse to distributional ambiguity and aims to minimize a good upper bound on the objective function. Such an approach of soliciting inventory decisions based on partial demand information is not new. In the 1950s, Scarf (1958) considered a min-max newsvendor problem with uncertain demand $\tilde{d}$ given by only its mean and standard deviations. Scarf was able to obtain solutions to the tight upper bound of the newsvendor problem. The central idea in addressing such a problem is to solicit a good upper bound on $\mathrm{E}((\cdot)^+)$, which appears at the objective of the newsvendor problem and also in Problem (3.3). The following result is well known:

[1] It will be shown subsequently that the bound on expectation using directional deviations is valid only when the factors are stochastically independent. For dependent uncertainties, we set $p_j = q_j = \infty$ for $j \notin \mathcal{I}$.

Proposition 2 *(Scarf's upper bound Scarf (1958)) Let $\tilde{z}$ be a random variable in $[-\mu, \infty)$ with mean μ and standard deviation σ, then, for all $a \geq -\mu$,*

$$E((\tilde{z}-a)^+) \leq \begin{cases} \dfrac{1}{2}\left(-a+\sqrt{\sigma^2+a^2}\right) & \textit{if } a \geq \frac{\sigma^2-\mu^2}{2\mu} \\ -a\dfrac{\mu^2}{\mu^2+\sigma^2}+\mu\dfrac{\sigma^2}{\mu^2+\sigma^2} & \textit{if } a < \frac{\sigma^2-\mu^2}{2\mu} \end{cases}.$$

Moreover, the bound is achievable.

Interestingly, Bertsimas and Thiele (2006) used the bound of Proposition 2 to calibrate the budget of uncertainty parameter in their robust inventory models. Unfortunately, it is generally computationally intractable to evaluate tight probability bounds involving multivariate random variables with known moments and support information (see Bertsimas and Popescu (2002)). We adopt the bound of Chen and Sim (2009) to evaluate the expected positive part of an affine sum of random variables under Assumption U. This bound is constructed from 5 different bounds to $\mathrm{E}((\cdot)^+)$, consisting of support, (second) moments, and deviation measures.

Definition 2 *We say a function $f(\boldsymbol{z})$ is* nonzero crossing *with respect to $\boldsymbol{z} \in \mathcal{W}$ if at least one of the following conditions holds:*

1. $f(\boldsymbol{z}) \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$
2. $f(\boldsymbol{z}) \leq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$.

Theorem 1 *(Chen and Sim (2009)) Let $\tilde{\boldsymbol{z}} \in \Re^N$ be a multivariate random variable under the Assumption U. Then*

$$E((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+) \leq \pi(y_0, \boldsymbol{y}),$$

where $\pi(y_0, \boldsymbol{y})$ *is given by*

$$
\begin{array}{rll}
\pi(y_0, \boldsymbol{y}) = \min & r_1 + r_2 + r_3 + r_4 + r_5 & \\
\text{s.t.} & y_{10} + \max\limits_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{z}'\boldsymbol{y}_1 \leq r_1 & \\
& 0 \leq r_1 & \\
& \max\limits_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{z}'(-\boldsymbol{y}_2) \leq r_2 & \\
& y_{20} \leq r_2 & \\
& \frac{1}{2} y_{30} + \frac{1}{2} \|(y_{30}, \boldsymbol{\Sigma}^{1/2} \boldsymbol{y}_3)\|_2 \leq r_3 & \\
& \inf_{\mu > 0} \frac{\mu}{e} \exp\left(\frac{y_{40}}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2}\right) \leq r_4 & \\
& u_j \geq p_j y_{4j} \ \forall j \in \mathcal{I}_1, \qquad y_{4j} \leq 0 \ \forall j \in \bar{\mathcal{I}}_1 & (3.7) \\
& u_j \geq -q_j y_{4j} \ \forall j \in \mathcal{I}_2, \qquad y_{4j} \geq 0 \ \forall j \in \bar{\mathcal{I}}_2 & \\
& y_{50} + \inf_{\mu > 0} \frac{\mu}{e} \exp\left(-\frac{y_{50}}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}\right) \leq r_5 & \\
& v_j \geq q_j y_{5j} \ \forall j \in \mathcal{I}_2, \qquad y_{5j} \leq 0 \ \forall j \in \bar{\mathcal{I}}_2 & \\
& v_j \geq -p_j y_{5j} \ \forall j \in \mathcal{I}_1, \qquad y_{5j} \geq 0 \ \forall j \in \bar{\mathcal{I}}_1 & \\
& y_{10} + y_{20} + y_{30} + y_{40} + y_{50} = y_0 & \\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5 = \boldsymbol{y}. & \\
& r_i, y_{i0} \in \Re, \boldsymbol{y}_i \in \Re^N, i = 1, \ldots, 5, \boldsymbol{u}, \boldsymbol{v} \in \Re^N. &
\end{array}
$$

Moreover, the bound is tight if $y_0 + \boldsymbol{y}'\boldsymbol{z}$ *is a nonzero crossing function with respect to* $\boldsymbol{z} \in \mathcal{W}$. *That is, if*

$$y_0 + \boldsymbol{y}'\boldsymbol{z} \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

we have $E\big((\, y_0 + \boldsymbol{y}'\boldsymbol{z}\,)^+\big) = \pi(y_0, \boldsymbol{y}) = y_0$. *Likewise, if*

$$y_0 + \boldsymbol{y}'\boldsymbol{z} \leq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W},$$

we have $E\big((\, y_0 + \boldsymbol{y}'\boldsymbol{z}\,)^+\big) = \pi(y_0, \boldsymbol{y}) = 0$.

Remark 1: The convexity of $\pi(y_0, \boldsymbol{y})$ depends on the convexity of the following function

$$f(u_0, \boldsymbol{u}) = \inf_{\mu > 0} \mu \exp\left(\frac{u_0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{\mu^2}\right).$$

It is easy to see that $g(u_0, \boldsymbol{u}) = \exp(u_0 + \|\boldsymbol{u}\|_2^2)$ is a convex function, and it is straightforward to check that $h(u_0, \boldsymbol{u}, \mu) = \mu g(u_0/\mu, \boldsymbol{u}/\mu)$ is a convex function on domain $\mu > 0$. Hence, $f(u_0, \boldsymbol{u}) = \inf\limits_{\mu > 0} h(u_0, \boldsymbol{u}, \mu)$ is a convex function. Due to the presence of such a function, the set of constraints in Problem (3.7) is not exactly second-order cone representable (see Ben-Tal and Nemirovski (2001)). Fortunately, using a few second-order cones, we can accurately approximate such constraints to a good level of numerical precision. The interested readers can refer to Chen and Sim (2009).

Remark 2: Note that the first and third constraints involving the support set $\mathcal{W}$, take the form of

$$\max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{v}'\boldsymbol{z} \leq r$$

or, equivalently, as

$$\boldsymbol{v}'\boldsymbol{z} \leq r \qquad \forall \boldsymbol{z} \in \mathcal{W}.$$

Such a constraint is known as the robust counterpart whose explicit formulation under different choices of tractable support set $\mathcal{W}$ is well discussed in Ben-Tal and Nemirovski (1998, 2001). Because $\mathcal{W}$ is a second order conic representable set, the robust counterpart is also second-order cone representable. For instance, if $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$, the corresponding robust counterpart is representable by the following linear inequalities:

$$\underline{\boldsymbol{z}}'\boldsymbol{t} + \bar{\boldsymbol{z}}'\boldsymbol{s} \leq r,$$

for some $\boldsymbol{s}, \boldsymbol{t} \geq \boldsymbol{0}$ satisfying $\boldsymbol{s} - \boldsymbol{t} = \boldsymbol{v}$.

Remark 3: Note that under the Assumption U, it is not necessary to provide all the information, such as the directional deviations. Therefore, whenever such information is unavailable, we can assign an infinite value to the corresponding parameter. For instance, supposing that factor $\tilde{z}_j$ has standard deviation σ and unknown directional deviations; we would set $p_j = q_j = \infty$. When the bounds on p_j and q_j are finite, the $\pi(\cdot)$ bound will be tighter.

Remark 4: In the absence of uncertainty, the nonzero crossing condition ensures that the bound is tight. That is, $y^+ = \mathrm{E}(y^+) = \pi(y, \boldsymbol{0})$.

Remark 5: The uncertainty model assumes that we have exact estimates of the covariance, means, and deviation measures from data. However, it is possible to consider a model of data uncertainty in which the covariance, means, and deviation measures are uncertain and belong to some uncertainty set. This can be done by modifying the bound $\pi(y_0, \boldsymbol{y})$ and applying standard robust optimization techniques such as those of Ben-Tal and Nemirovski (1998) and Bertsimas and Sim (2006).

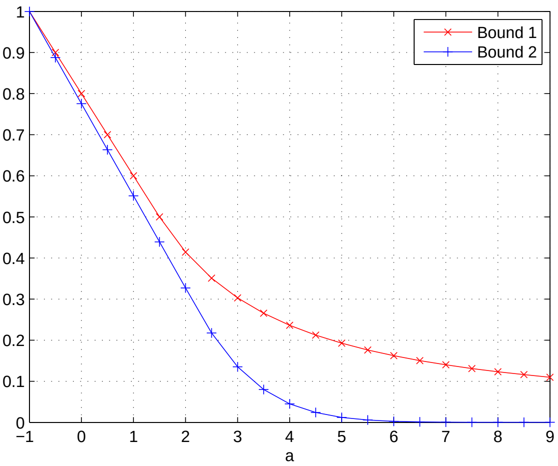


Figure 3.1: Comparing the bounds of $\mathrm{E}((\tilde{z}-a)^+)$

The robust model of Bertsimas and Thiele (2006) uses Proposition 2. Next, we show that for a univariate random variable with one-sided support, the bound of Theorem 1 is just as tight.

Proposition 3 *Let $\tilde{z}$ be a random variable in $[-\mu, \infty)$ with mean μ and standard deviation σ, then for all $a \geq -\mu$,*

$$E((\tilde{z}-a)^+) \leq \pi(-a,1) = \begin{cases} \frac{1}{2}\left(-a+\sqrt{\sigma^2+a^2}\right) & \text{if } a \geq \frac{\sigma^2-\mu^2}{2\mu} \\ -a\frac{\mu^2}{\mu^2+\sigma^2} + \mu\frac{\sigma^2}{\mu^2+\sigma^2} & \text{if } a < \frac{\sigma^2-\mu^2}{2\mu} \end{cases}.$$

Proof : See Appendix A.

We can further improve the bound if the distribution of the random variable $\tilde{z}$ is sufficiently light tailed such that the directional deviations are close to its standard deviation, such as those of normal and uniform distributions. Figure 3.1 compares the bounds of $\mathrm{E}((\tilde{z}-a)^+)$ in which $\mu = 1$ and $\sigma = \sigma_f(\tilde{z}) = \sigma_b(\tilde{z}) = 2$. Bound 1 corresponds to the bound of Proposition 2, whereas Bound 2 corresponds to the bound of Theorem 1. Clearly, despite the lack of tightness results, incorporating the directional deviations can potentially improve the bound on $\mathrm{E}((\tilde{z}-a)^+)$. We will further demonstrate the benefits in our computational experiments.

3.2.2 Tractable Replenishment Policies

Having introduced the demand uncertainty model, a suitable approximation of the replenishment policy $x_t(\tilde{\boldsymbol{d}}_{t-1})$ is needed to obtain a tractable formulation. That is, we seek a formulation in which the policy can be obtained by solving an optimization problem that runs in polynomial time and is scalable across time periods. We review two tractable replenishment policies, static as well as linear with respect to the random factors of demand, which are decision rules prevalent in the context of robust optimization. We also introduce a new replenishment policy known as the truncated linear replenishment policy, that improves over these policies.

Static replenishment policy

The static replenishment policy, also known as the open-loop policy, has order decisions not influenced by random factors of demand as follows:

$$x_t(\tilde{\boldsymbol{d}}_{t-1}) = x_t^0. \tag{3.8}$$

A tractable model under such a replenishment policy is as follows:

$$\begin{aligned} Z_{SRP} = \min \quad & \sum_{t=1}^{T} \left(c_t x_t^0 + h_t \pi \left(y_{t+1}^0, \boldsymbol{y}_{t+1} \right) + b_t \pi \left(-y_{t+1}^0, -\boldsymbol{y}_{t+1} \right) \right) \\ \text{s.t.} \quad & y_{t+1}^0 = y_t^0 + x_{t-L}^0 - d_t^0 \qquad t = 1, \dots, T \\ & y_{t+1}^k = y_t^k - d_t^k \qquad k = 1, \dots, N, t = 1, \dots, T \\ & y_{t+1}^k = 0 \qquad k \geq N_t + 1, t = 1, \dots, T \\ & 0 \leq x_t^0 \leq S_t \qquad t = 1, \dots, T - L, \end{aligned} \tag{3.9}$$

with y_1^0 being the initial inventory level and $y_1^k = 0$ for all $k = 1, \dots, N$. For $L \geq 1$, x_t^0 are the known committed orders made at time periods $t = 1 - L, \dots, 0$.

Under Equation (3.8), it is evident from Equation (3.1) that the inventory level also takes an affine structure,

$$y_{t+1}(\tilde{\boldsymbol{d}}_t) = y_{t+1}^0 + \sum_{k=1}^{N} y_{t+1}^k \tilde{z}_k. \tag{3.10}$$

Using Theorem 1, we can bound the excess inventory level at time period t, that is, $\mathrm{E}\left((y_{t+1}(\tilde{\boldsymbol{d}}_t))^+ \right) \leq \pi(y_{t+1}^0, \boldsymbol{y}_{t+1})$. Proceeding similarly for the backlog inventory gives the objective function of Problem (3.9). Equating the coefficients of the constant and $\tilde{z}_k$ term of Equation (3.1) gives the first two sets of constraints in Problem (3.9), respectively. The last set of constraints enforces the range on order quantity, that is, nonnegativity and upper limit.

Theorem 2 *The expected cost of the stochastic inventory problem under the static replenishment policy,*

$$x_t^{SRP}(\tilde{\boldsymbol{d}}_{t-1}) = x_t^{0*} \qquad t = 1, \dots, T-L$$

in which x_t^{0*}, $t = 1, \dots, T-L$ is the optimum solution of Problem (3.9), is at most Z_{SRP}.

Proof : See Appendix B.

Linear replenishment policy

A more refined replenishment policy introduced in Ben-Tal et al. (2005), and Chen et al. (2007) is the linear replenishment policy where the order decisions are affinely dependent on the random factors of demand, that is,

$$x_t^{LRP}(\tilde{\boldsymbol{d}}_{t-1}) = x_t^0 + \boldsymbol{x}_t'\tilde{\boldsymbol{z}}, \tag{3.11}$$

in which the vector $\boldsymbol{x}_t = (x_t^1, \dots, x_t^N)$ satisfies the following nonanticipative constraints:

$$x_t^k = 0 \qquad \forall k \geq N_{t-1} + 1. \tag{3.12}$$

Because the order decision is made at the beginning of the tth period, the nonanticipative constraints ensure that the linear replenishment policy is not influenced by demand factors that are unavailable up to the beginning of the tth period. The model for the linear replenishment policy is as follows:

$$\begin{array}{rlll} Z_{LRP} = \min & \sum_{t=1}^{T}\left(c_t x_t^0 + h_t \pi\left(y_{t+1}^0, \boldsymbol{y}_{t+1}\right) + b_t \pi\left(-y_{t+1}^0, -\boldsymbol{y}_{t+1}\right)\right) & & \\ \text{s.t.} & y_{t+1}^k = y_t^k + x_{t-L}^k - d_t^k & k = 0, \dots, N, t = 1, \dots, T & \\ & y_{t+1}^k = 0 & k \geq N_t + 1, t = 1, \dots, T & \\ & x_t^k = 0 & k \geq N_{t-1} + 1, t = 1, \dots, T-L & \\ & 0 \leq x_t^0 + \boldsymbol{x}_t'\boldsymbol{z} \leq S_t & \forall \boldsymbol{z} \in \mathcal{W} \qquad t = 1, \dots, T-L, & \end{array} \tag{3.13}$$

with y_1^0 being the initial inventory level and $y_1^k = 0$ for all $k = 1, \dots, N$. For $L \geq 1$, x_t^0 are the known committed orders made at time periods $t = 1-L, \dots, 0$.

Under Equation (3.11), the inventory level has a structure similar to Equation (3.10). The objective function and the first set of constraints are hence obtained in a similar manner as Problem (3.9). The last set of constraints ensures that the linear replenishment policy is confined within the ordering capacity for all possible states of random factors. Observe that under the assumption that $\mathcal{W}$ is a tractable conic representable uncertainty set, the robust counterpart

$$0 \leq x_t^0 + \boldsymbol{x}_t'\boldsymbol{z} \leq S_t \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

can be represented concisely as tractable conic constraints. Therefore, Problem (3.13) is essentially a tractable conic optimization problem.

Theorem 3 *The expected cost of the stochastic inventory problem under the linear replenishment policy,*

$$x_t^{LRP}(\tilde{\boldsymbol{d}}_{t-1}) = x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}} \qquad t = 1, \ldots, T-L,$$

in which x_t^{k*}, $k = 0, \ldots, N, t = 1, \ldots, T-L$ *is the optimum solution of Problem (3.13), is at most* Z_{LRP}. *Moreover,* $Z_{LRP} \leq Z_{SRP}$.

Proof : See Appendix C.

Truncated linear replenishment policy

Chen et al. (2008) studied the weakness of linear decision rules (or policy) and showed that carefully chosen piecewise-linear decision rules can strengthen the approximation of stochastic optimization problems. Indeed, a base-stock policy such as Equation (3.4) can be shown by induction to be piecewise-linear with respect to the historical demands. In the same spirit, we introduce a new piecewise-linear replenishment policy that we call the truncated linear replenishment policy. It takes the following form:

$$x_t^{TLRP}(\tilde{\boldsymbol{d}}_{t-1}) = \min\left\{\max\left\{x_t^0 + \boldsymbol{x}_t'\tilde{\boldsymbol{z}}, 0\right\}, S_t\right\}, \qquad (3.14)$$

where the vector $\boldsymbol{x}_t = (x_t^1, \ldots, x_t^N)$ satisfies the following nonanticipative constraints:

$$x_t^k = 0 \qquad \forall k \geq N_{t-1} + 1. \qquad (3.15)$$

Note that the truncated linear replenishment policy is piecewise-linear and directly satisfies the ordering range constraint as follows:

$$0 \leq x_t^{TLRP}(\tilde{\boldsymbol{d}}_{t-1}) \leq S_t.$$

Before introducing the model, we present the following bound on the expectation of a nested sum of positive values of random variables:

Theorem 4 *Let* $\tilde{\boldsymbol{z}} \in \Re^N$ *be a multivariate random variable under Assumption U. Then*

$$E\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i=1}^{p}\left(x_i^0 + \boldsymbol{x}_i'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \leq \eta((y^0, \boldsymbol{y}), (x_1^0, \boldsymbol{x}_1), \ldots, (x_p^0, \boldsymbol{x}_p)), \quad (3.16)$$

where

$$\begin{aligned} &\eta((y^0, \boldsymbol{y}), (x_1^0, \boldsymbol{x}_1), \ldots, (x_p^0, \boldsymbol{x}_p)) \\ = &\min_{w_i^0, \boldsymbol{w}_i, i=1,\ldots,p} \left\{ \pi\left(y^0 + \sum_{i=1}^{p} w_i^0, \boldsymbol{y} + \sum_{i=1}^{p} \boldsymbol{w}_i\right) + \right. \\ &\left. \sum_{i=1}^{p}\left(\pi(-w_i^0, -\boldsymbol{w}_i) + \pi(x_i^0 - w_i^0, \boldsymbol{x}_i - \boldsymbol{w}_i)\right) \right\}. \end{aligned}$$

Moreover, the bound is tight if $y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i=1}^{p}\left(x_i^0 + \boldsymbol{x}_i'\boldsymbol{z}\right)^+$ *and* $x_i^0 + \boldsymbol{x}_i'\boldsymbol{z}$*,* $i = 1, \ldots, p$ *are nonzero crossing functions with respect to* $\boldsymbol{z} \in \mathcal{W}$*.*

Proof : See Appendix D.

Remark : It is easy to establish that

$$\begin{aligned} &\mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i=1}^{p}\left(x_i^0 + \boldsymbol{x}_i'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \\ \leq\ &\mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}\right)^+\right) + \sum_{i=1}^{p}\mathrm{E}\left(\left(x_i^0 + \boldsymbol{x}_i'\tilde{\boldsymbol{z}}\right)^+\right) \\ \leq\ &\pi\left(y^0, \boldsymbol{y}\right) + \sum_{i=1}^{p}\pi(x_i^0, \boldsymbol{x}_i). \end{aligned}$$

However, this is a weaker bound, considering the fact that

$$\begin{aligned} &\eta((y^0, \boldsymbol{y}), (x_1^0, \boldsymbol{x}_1), \ldots, (x_p^0, \boldsymbol{x}_p)) \\ = &\min_{w_i^0, \boldsymbol{w}_i, i=1,\ldots,p} \left\{ \pi\left(y^0 + \sum_{i=1}^{p} w_i^0, \boldsymbol{y} + \sum_{i=1}^{p} \boldsymbol{w}_i\right) + \right. \\ &\left. \sum_{i=1}^{p}\left(\pi(-w_i^0, -\boldsymbol{w}_i) + \pi(x_i^0 - w_i^0, \boldsymbol{x}_i - \boldsymbol{w}_i)\right) \right\} \\ \leq\ &\pi\left(y^0, \boldsymbol{y}\right) + \sum_{i=1}^{p}\pi(x_i^0, \boldsymbol{x}_i). \end{aligned}$$

The model for the truncated linear replenishment policy can be formulated as follows:

$$
\begin{aligned}
& Z_{TLRP} = \\
\min \quad & \sum_{t=1}^{T} c_t \pi(x_t^0, \boldsymbol{x}_t) + \sum_{t=1}^{L} \Big(h_t \pi(y_{t+1}^0, \boldsymbol{y}_{t+1}) + b_t \pi(-y_{t+1}^0, -\boldsymbol{y}_{t+1}) \Big) + \\
& \quad \sum_{t=L+1}^{T} \Big(h_t \eta \left((y_{t+1}^0, \boldsymbol{y}_{t+1}), (-x_1^0, -\boldsymbol{x}_1), \ldots, (-x_{t-L}^0, -\boldsymbol{x}_{t-L}) \right) + \\
& \qquad b_t \eta \left((-y_{t+1}^0, -\boldsymbol{y}_{t+1}), (x_1^0 - S_t, \boldsymbol{x}_1), \ldots, (x_{t-L}^0 - S_t, \boldsymbol{x}_{t-L}) \right) \Big) \\
\text{s.t.} \quad & y_{t+1}^k = y_t^k + x_{t-L}^k - d_t^k \qquad k = 0, \ldots, N, t = 1, \ldots, T \\
& y_{t+1}^k = 0 \qquad k \geq N_t + 1, t = 1, \ldots, T \\
& x_t^k = 0 \qquad k \geq N_{t-1} + 1, t = 1, \ldots, T - L
\end{aligned}
\tag{3.17}
$$

with y_1^0 being the initial inventory level and $y_1^k = 0$ for all $k = 1, \ldots, N$. For $L \geq 1$, x_t^0 are the known committed orders made at time periods $t = 1 - L, \ldots, 0$.

Under Equation (3.14) the inventory level, $y_{t+1}(\tilde{\boldsymbol{d}}_t)$ is no longer affinely dependent on $\tilde{\boldsymbol{z}}$. The terms in the objective function account for the costs associated with excess inventory level and backlog, taking into consideration the piecewise-linear policy. It can be shown that the truncated linear replenishment policy dominates over the linear replenishment policy as follows.

Theorem 5 *The expected cost of the stochastic inventory problem under the truncated linear replenishment policy,*

$$x_t^{TLRP}(\tilde{\boldsymbol{d}}_{t-1}) = \min\left\{\max\left\{x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}}, 0\right\}, S_t\right\} \qquad t = 1, \ldots, T - L$$

in which x_t^{k*}, $k = 0, \ldots, N$, $t = 1, \ldots, T - L$ is the optimum solution of Problem (3.17), is at most Z_{TLRP}. Moreover, $Z_{TLRP} \leq Z_{LRP}$.

Proof : See Appendix E.

We have shown that $Z_{STOC} \leq Z_{TLRP} \leq Z_{LRP} \leq Z_{SRP}$. The linear replenishment policy improves over the static replenishment policy because it is able to adapt to demand history. Because setting the coefficient of the random factors $\boldsymbol{x}_t$ to be zero in Problem (3.13) gives Problem (3.9), it is evident from Equation (3.11) that the linear replenishment policy subsumes the static replenishment policy. Observe that in Problem (3.13), from which the solution of the linear replenishment policy is derived, the set of constraints restricting the ordering quantity

$$0 \leq x_t^0 + \boldsymbol{x}_t'\boldsymbol{z} \leq S_t \qquad \forall \boldsymbol{z} \in \mathcal{W} \qquad t = 1, \ldots, T - L$$

can be overly constraining on the replenishment policy. For the case when the uncertainty set $\mathcal{W}$ is unbounded, such as $\mathcal{W} = \{\boldsymbol{z} \ : \ \boldsymbol{z} \geq -\underline{\boldsymbol{z}}\}$, the decision variables $\boldsymbol{x}_t$ will be driven to zeroes. This means that the ordering decision of Problem (3.13) degenerates to a static replenishment policy, losing the ability to adapt to the history of random factors. The truncated linear replenishment policy, on the other hand, avoids this issue. Moreover, we also note that in Problem (3.13), information of mean, variance, and directional deviations are not utilized for the set of constraints restricting the ordering quantity. In contrast, the truncated linear replenishment policy is defined to satisfy the ordering constraint. Hence, the robust model of Problem (3.17) does not have the explicit constraints on ordering levels and is able to utilize the additional information via the π and η functions for improving the bound.

It should be noted that establishing the bounds does not necessarily imply the superiority of truncated linear replenishment policy over static and linear ones. Nevertheless, this behavior is observed throughout our computational studies.

3.3 Extensions

In this section, we discuss some extensions to the basic model.

3.3.1 Fixed Ordering Cost

Unfortunately, with fixed ordering cost the inventory replenishment problem becomes nonconvex and is much harder to address. Using the idea of Bertsimas and Thiele (2006), we can formulate a restricted problem where the time period in which the orders that can be placed is determined at the start of the planning horizon as follows:

$$
\begin{aligned}
& Z_{STOCF} = \\
\min \quad & \sum_{t=1}^{T}\left(\mathrm{E}\left(c_t x_t(\tilde{\boldsymbol{d}}_{t-1}) + K_t\delta_t + h_t(y_{t+1}(\tilde{\boldsymbol{d}}_t))^+\right) + \mathrm{E}\left(b_t(y_{t+1}(\tilde{\boldsymbol{d}}_t))^-\right)\right) \\
\text{s.t.} \quad & y_{t+1}(\tilde{\boldsymbol{d}}_t) = y_t(\tilde{\boldsymbol{d}}_{t-1}) + x_{t-L}(\tilde{\boldsymbol{d}}_{t-L-1}) - \tilde{d}_t \qquad t = 1, \ldots, T \\
& 0 \le x_t(\tilde{\boldsymbol{d}}_{t-1}) \le S_t\delta_t \qquad t = 1, \ldots, T-L \\
& \delta_t \in \{0, 1\} \qquad\qquad\qquad t = 1, \ldots, T-L.
\end{aligned}
\tag{3.18}
$$

In Problem (3.18), inventory can only be replenished at a period where the corresponding binary variable δ_t takes the value of one. We can then incorporate the tractable replenishment policies developed in the previous section, and exploit $\pi(\cdot)$ and $\eta(\cdot)$ to bound the terms at the objective function. The resulting optimization model is a conic integer program since $\pi(\cdot)$ and $\eta(\cdot)$ are SOC functions. Conic integer program is already addressed in commercial solvers such as CPLEX 11.2. Admittedly, algorithms for solving conic integer programs are still in their infancy. On the theoretical front, Atamtürk and Narayanan (2009) recently developed general-purpose conic mixed-integer rounding cuts based on polyhedral conic substructures of second-order conic sets, which can be readily incorporated in branch-and-bound algorithms that solve continuous conic optimization problems at the nodes of the search tree. Their preliminary computational experiments suggest that the new cuts are quite effective in reducing the integrality gap of continuous relaxations of conic mixed-integer programs.

3.3.2 Supply Chain Networks

The models we have presented in the preceding section can also be extended to more complex supply chain networks such as the serial system, or, more generally tree networks. These are multistage systems where goods transit from one stage to the next stage, each time moving closer to their final destination. In many supply chains, the main storage hubs, or the sources of the network, receive their supplies from outside manufacturing plants in a treelike hierarchical structure and send items throughout the network until they finally reach the stores, or the sinks of the network. The extension to tree structure uses the concept of echelon inventory and closely follows Bertsimas and Thiele (2006). We refer interested readers to their paper.

We have completed the theoretical discussions on the models. In the next chapter, we will discuss via computational studies, the effectiveness of our robust approach. In particular, we compare it against existing approaches in the literature.

Chapter 4

Computational Studies of Multiperiod Inventory Model

In this chapter we discuss the computational performance of the static, linear and truncated linear replenishment policies against the optimum history dependent policy and two dynamic programming based heuristics, namely, the myopic policy and a history-independent base-stock policy. Specifically, we examined the quality of truncated linear replenishment policies over many realistic scenarios of planing horizon, cost parameters and demand correlations. To benchmark the performance, we have to assume knowledge of the underlying distribution of the demand. We did not conduct experiments to test robustness of policies against distributional ambiguity such as those studied in Bertsimas and Thiele (2006) and Chen and Sim (2009). Instead, we have focused on how well or poorly the tractable replenishment policies perform against the optimum policy obtained by dynamic programming, as well as against heuristics in inventory control.

We are aware of the folding horizon implementation, where the replenishment policy can be enhanced by solving repeatedly with updated demand information. For instance, the static replenishment policy proposed by Bertsimas and Thiele (2006) has a base-stock structure under the folding horizon implementation. Since more accurate information is used each time the model is solved, the results will only improve. Unfortunately, due to the computational intensiveness of the evaluation, we have excluded folding horizon implementations from our computational studies. For instance, under the folding horizon implementation, it would typically take about four minutes to evaluate the sample path of a ten period model based on the truncated linear replenishment policy. Through sizing experiments, we envisaged that it would require about 100,000 sample paths to reduce the standard error of the estimated objective value to less than 1%, which amounts to about 280 days of computational time.

4.1 Experimental Setup

The demand process we considered is motivated by Graves (1999) as follows:

$$d_t(\tilde{\boldsymbol{z}}) = \tilde{z}_t + \alpha\tilde{z}_{t-1} + \alpha\tilde{z}_{t-2} + \cdots + \alpha\tilde{z}_1 + \mu, \tag{4.1}$$

where the shocks factors $\tilde{z}_t$ are independently uniformly distributed random variables in $[-\bar{z}, \bar{z}]$, and have standard deviations and directional deviations numerically close to $0.58\bar{z}$.

Observe that the demand process of Equation (4.1) for $t \geq 2$ can be expressed recursively as

$$d_t(\tilde{\boldsymbol{z}}) = d_{t-1}(\tilde{\boldsymbol{z}}) - (1-\alpha)\tilde{z}_{t-1} + \tilde{z}_t. \tag{4.2}$$

Hence, this demand process is an integrated moving average (IMA) process of order $(0, 1, 1)$. See also Box et al. (1994). Note that given $\bar{\mu} = d_{t-1}(\tilde{\boldsymbol{z}}) - (1-\alpha)\tilde{z}_{t-1}$ at time period t, the distribution of $d_t(\tilde{\boldsymbol{z}})$ is uniform in $[-\bar{z} + \bar{\mu}, \bar{z} + \bar{\mu}]$.

A range of demand processes can be modeled by varying α. With $\alpha = 0$, the demand process follows an i.i.d process of uniformly distributed random variables. As α grows, the demand process becomes nonstationary and less stable with increasing variance. When $\alpha = 1$, the demand process is a random walk on a continuous state space.

We considered problems with $T = 5, 10, 20$ and 30, and selected parameters so that the demand, $d_t(\tilde{\boldsymbol{z}})$ is nonnegative for all $\alpha \in [0, 1]$. The lead time L is zero, $S_t = 260$, unit ordering cost $c_t = 0.1$, and unit holding cost $h_t = 0.02$ for all periods $t = 1, \ldots, T$. In view of the long computational time for dynamic programming, especially for $T = 20$ and 30, we have used more manageable parameters for the demand process as follows.

- For $T = 5$, we used $\mu = 200$, and $\bar{z} = 40$.

- For $T = 10$, we used $\mu = 200$, and $\bar{z} = 20$.
- For $T = 20$, we used $\mu = 240$, and $\bar{z} = 12$.
- For $T = 30$, we used $\mu = 240$, and $\bar{z} = 8$.

Since unfulfilled demands are lost at the end of T, we set a relatively high backlog cost, $b_T = 10b_1$, to heavily penalize unmet demand at the last period throughout our experiments. For notational convenience, we use b and h to denote the backlog and holding cost from $t = 1 \dots T-1$. In our study, we varied α from 0 to 1 in steps of 0.25 and set b/h to range from 10 to 50.

We benchmarked our solutions against solution based on dynamic programming, where the optimum replenishment policy for the case of zero lead time can be characterized by the following backward recursion:

$$\begin{aligned} & J_t(y_t, d_{t-1}, z_{t-1}) \\ = & \min_{0 \le x \le S_t} \mathrm{E}\Big(c_t x + G_t(y_t + x - d_{t-1} + (1-\alpha) z_{t-1} - \tilde{z}_t, \tilde{z}_t)) + \\ & \quad J_{t+1}(y_t + x - d_{t-1} + (1-\alpha) z_{t-1} - \tilde{z}_t, d_{t-1} - (1-\alpha) z_{t-1} + \tilde{z}_t)\Big) \end{aligned}$$

where $G_t(u) = h_t \max(u, 0) + b_t \max(-u, 0)$. By letting $v_t = d_t - (1-\alpha) z_t$, we have equivalently

$$J_t(y_t, v_{t-1}) = \min_{0 \le x \le S_t} \mathrm{E}\Big(c_t x + G_t(y_t + x - v_{t-1} - \tilde{z}_t) + J_{t+1}(y_t + x - v_{t-1} - \tilde{z}_t, v_{t-1} + \alpha \tilde{z}_t)\Big),$$

which reduces the state space by one dimension. The optimum replenishment policy at time t is a function of the current inventory level y_t and v_{t-1} as follows:

$$\begin{aligned} & x_t^{OPT}(y_t, v_{t-1}) = \\ & \arg \min_{0 \le x \le S_t} \mathrm{E}\Big(c_t x + G_t(y_t + x - v_{t-1} - \tilde{z}_t) + J_{t+1}(y_t + x - v_{t-1} - \tilde{z}_t, v_{t-1} + \alpha \tilde{z}_t)\Big). \end{aligned}$$

In our implementation, we discretized the value functions uniformly and used linear interpolations for evaluating the intermediate points. The underlying expectations were computed using the well-known Simpson's rule of numerical integration. To obtain a near optimum policy within reasonable time, we adjusted the level of discretization such that when the discretization is increased by two, the improvement in objective value is less than 1%.

We also considered two heuristics. The first is a history-independent base-stock policy (BSP), where we computed the replenishment policy recursively by ignoring the dependency of previous demands as follows:

$$J_t^{BSP}(y_t) = \min_{0\le x\le S_t} \mathrm{E}_{\tilde{d}_t}\Big(c_t x + G_t(y_t + x - \tilde{d}_t) + J_{t+1}^{BSP}(y_t + x - \tilde{d}_t)\Big),$$

where $\tilde{d}_t = \tilde{z}_t + \alpha\tilde{z}_{t-1} + \alpha\tilde{z}_{t-2} + \cdots + \alpha\tilde{z}_1 + \mu$. The replenishment policy is given by

$$x_t^{BSP}(y_t) = \arg\min_{0\le x\le S_t} \mathrm{E}_{\tilde{d}_t}\Big(c_t x + G_t(y_t + x - \tilde{d}_t) + J_{t+1}^{BSP}(y_t + x - \tilde{d}_t)\Big).$$

Under capacity limit on order quantities, the modified history-independent base-stock policy is optimum if the demands are independently distributed, which occurs only when $\alpha = 0$ (see Federgruen and Zipkin (1986)). Note that when $\alpha > 0$, evaluating the expectation exactly involves multi-dimensional integration, which can be computationally prohibitive. Therefore, at every dynamic programming recursion, we computed the value functions approximately using sampling approximations from 500 instances of demand realizations instead.

The other heuristic we considered is an adaptive myopic policy (MP), where the replenishment level for the case of zero lead time is derived by minimizing the following one-period expected cost as described below:

$$x_t^{MP}(y_t, v_{t-1}) = \arg\min_{0\le x\le S_t} \mathrm{E}\Big(c_t x + G_t(y_t + x - v_{t-1} - \tilde{z}_t)\Big).$$

Under the uniform distribution, the myopic policy can be obtained using the critical fractile as follows:

$$x_t^{MP}(y_t, v_{t-1}) = \min\left\{\left(v_{t-1} - \bar{z} + 2\bar{z}\left(1 - \frac{c_t + h_t}{b_t + h_t}\right) - y_t\right)^+, S_t\right\}.$$

In contrast with the optimum dynamic programming recursion, the adaptive myopic policy optimizes only the current period expected cost, and ignores all subsequent costs.

After obtaining the policies, we compared them using 100,000 simulated inventory runs and reported the sample means over all the runs. The results for the $T = 5, 10, 20$ and 30 problems are given in Table 4.1, Table 4.2, Table 4.3 and Table 4.4, respectively. The robust policies were obtained using the bounds of Theorem 1 and Theorem 4 where the support, covariance, directional deviations associated with random factors are specified. In the tables, we have used TLRP, LRP, SRP, BSP, MP to denote the sample mean of the expected cost under the simulated runs when the replenishment policies are the truncated linear replenishment policy, linear replenishment policy, static replenishment policy, history-independent base-stock policy and adaptive myopic policy, respectively. Correspondingly, we used OPT to denote the values derived from the optimum policy. For convenience, we used these abbreviations to denote the respective policies throughout this chapter. We also provided in parentheses, the performance of the corresponding policy with respect to the optimum value. For example, the performance of TLRP given in parentheses shows the value of TLRP/OPT. A value of 1.05 hence shows that the deviation from OPT is 5%. We also reported the model objective values for the robust models as Z_{TLRP}, Z_{LRP} and Z_{SRP} to four significant places. Throughout the tables, the sample errors of the mean are less than 1%, and the sample means are shown to three significant places.

4.2 Comparison of Policies

In all the cases tested, TLRP deviates from the optimum answer by not more than 7%, whereas LRP is observed to deviate by as much as 29%; SRP by as much as 48%, MP by as much as 26%, and BSP by as much as 20% from OPT.

For $\alpha = 0$, TLRP and LRP perform well, coming within 1% from OPT. We observed that when α is small, the model objective values of TLRP and LRP, Z_{TLRP} and Z_{LRP}, come near to the simulated inventory runs, indicating the closeness of the bound. MP and BSP perform reasonably well for $\alpha \leq 0.5$ with deviation of not more than 10%. However, for large α, the deviation can exceed 20%. We observed that TLRP is never worse off against LRP, SRP, and outperforms BSP and MP in most of the cases. Moreover, TLRP has the sharpest lead against LRP, SRP and MP when the α is high. It is also interesting to note that when $\alpha = 1$, the bounds of LRP and SRP are rather close, while TLRP has much better performance.

Overall, the out-performance of TLRP over the rest of the non-optimum policies can be as high as 14%. In relatively few cases, BSP and MP may outperform TLRP. However, the margins do not exceed 1%. The results suggest that TLRP has the best overall performance.

Table 4.1: Performance of truncated linear replenishment policy $T = 5$

	Simulated Inventory Runs						Objective Value		
b/h	TLRP	LRP	SRP	MP	BSP	OPT	Z_{TLRP}	Z_{LRP}	Z_{SRP}
				$\alpha = 0$					
10	$108_{(1)}$	$108_{(1)}$	$121_{(1.1)}$	$115_{(1.07)}$	$107_{(1)}$	108	108.0	108.0	120.8
30	$108_{(1)}$	$108_{(1)}$	$124_{(1.13)}$	$110_{(1.02)}$	$108_{(1)}$	108	108.0	108.0	124.4
50	$108_{(1)}$	$108_{(1)}$	$126_{(1.14)}$	$109_{(1.01)}$	$108_{(1)}$	108	108.0	108.0	125.8
				$\alpha = 0.25$					
10	$108_{(1.01)}$	$109_{(1.01)}$	$130_{(1.18)}$	$116_{(1.08)}$	$109_{(1.01)}$	107	108.3	109.1	130.3
30	$108_{(1)}$	$109_{(1.01)}$	$136_{(1.22)}$	$111_{(1.03)}$	$110_{(1.02)}$	108	108.6	109.2	135.5
50	$108_{(1)}$	$109_{(1.01)}$	$138_{(1.24)}$	$110_{(1.02)}$	$110_{(1.02)}$	108	108.8	109.2	137.6
				$\alpha = 0.50$					
10	$110_{(1.02)}$	$118_{(1.06)}$	$141_{(1.25)}$	$119_{(1.1)}$	$112_{(1.04)}$	108	111.2	117.7	140.5
30	$111_{(1.02)}$	$125_{(1.1)}$	$148_{(1.31)}$	$114_{(1.05)}$	$115_{(1.06)}$	109	114.3	125.0	147.5
50	$112_{(1.03)}$	$130_{(1.12)}$	$150_{(1.33)}$	$113_{(1.04)}$	$117_{(1.07)}$	109	116.7	129.6	150.5
				$\alpha = 0.75$					
10	$113_{(1.03)}$	$133_{(1.14)}$	$151_{(1.31)}$	$126_{(1.15)}$	$117_{(1.07)}$	110	119.0	133.3	151.1
30	$118_{(1.05)}$	$153_{(1.22)}$	$163_{(1.35)}$	$125_{(1.12)}$	$124_{(1.1)}$	112	131.9	152.5	162.9
50	$122_{(1.06)}$	$166_{(1.25)}$	$173_{(1.34)}$	$130_{(1.14)}$	$130_{(1.14)}$	114	142.7	166.2	172.7
				$\alpha = 1$					
10	$118_{(1.04)}$	$152_{(1.21)}$	$163_{(1.35)}$	$137_{(1.21)}$	$126_{(1.12)}$	113	132.3	152.3	163.3
30	$131_{(1.06)}$	$191_{(1.28)}$	$193_{(1.31)}$	$151_{(1.22)}$	$145_{(1.18)}$	123	164.8	191.0	193.3
50	$140_{(1.06)}$	$223_{(1.28)}$	$223_{(1.29)}$	$168_{(1.28)}$	$158_{(1.2)}$	132	195.2	222.9	223.3

Table 4.2: Performance of truncated linear replenishment policy $T = 10$

	Simulated Inventory Runs						Objective Value		
b/h	TLRP	LRP	SRP	MP	BSP	OPT	Z_{TLRP}	Z_{LRP}	Z_{SRP}
				$\alpha = 0$					
10	$206_{(1)}$	$206_{(1)}$	$220_{(1.06)}$	$214_{(1.04)}$	$206_{(1)}$	206	206.0	206.0	220.2
30	$206_{(1)}$	$206_{(1)}$	$224_{(1.08)}$	$209_{(1.01)}$	$206_{(1)}$	206	206.0	206.0	223.8
50	$206_{(1)}$	$206_{(1)}$	$225_{(1.08)}$	$208_{(1.01)}$	$206_{(1)}$	206	206.0	206.0	225.3
				$\alpha = 0.25$					
10	$206_{(1)}$	$206_{(1)}$	$240_{(1.14)}$	$214_{(1.04)}$	$207_{(1.01)}$	206	206.0	206.1	239.5
30	$206_{(1)}$	$206_{(1)}$	$247_{(1.18)}$	$209_{(1.01)}$	$208_{(1.01)}$	206	206.0	206.1	246.7
50	$206_{(1)}$	$206_{(1)}$	$250_{(1.19)}$	$208_{(1.01)}$	$209_{(1.02)}$	206	206.0	206.1	249.7
				$\alpha = 0.50$					
10	$206_{(1)}$	$213_{(1.03)}$	$260_{(1.23)}$	$214_{(1.04)}$	$210_{(1.02)}$	206	206.3	213.0	260.0
30	$206_{(1)}$	$215_{(1.04)}$	$271_{(1.28)}$	$209_{(1.01)}$	$212_{(1.03)}$	206	207.0	215.1	270.9
50	$206_{(1)}$	$216_{(1.04)}$	$275_{(1.3)}$	$208_{(1.01)}$	$214_{(1.04)}$	206	207.5	216.0	275.5
				$\alpha = 0.75$					
10	$207_{(1.01)}$	$232_{(1.1)}$	$281_{(1.31)}$	$215_{(1.04)}$	$214_{(1.04)}$	206	210.5	231.6	280.8
30	$211_{(1.02)}$	$242_{(1.14)}$	$296_{(1.38)}$	$211_{(1.02)}$	$218_{(1.05)}$	207	215.4	241.9	295.6
50	$213_{(1.03)}$	$247_{(1.16)}$	$302_{(1.41)}$	$211_{(1.02)}$	$221_{(1.07)}$	207	218.2	247.4	301.8
				$\alpha = 1$					
10	$213_{(1.02)}$	$257_{(1.18)}$	$302_{(1.39)}$	$220_{(1.06)}$	$221_{(1.06)}$	208	220.6	257.4	301.8
30	$222_{(1.05)}$	$281_{(1.25)}$	$322_{(1.46)}$	$222_{(1.05)}$	$231_{(1.1)}$	210	235.5	281.1	321.7
50	$228_{(1.07)}$	$296_{(1.29)}$	$331_{(1.48)}$	$229_{(1.08)}$	$240_{(1.13)}$	212	245	296.0	331.5

Table 4.3: Performance of truncated linear replenishment policy $T = 20$

	Simulated Inventory Runs						Objective Value		
b/h	TLRP	LRP	SRP	MP	BSP	OPT	Z_{TLRP}	Z_{LRP}	Z_{SRP}
				$\alpha = 0$					
10	$486_{(1)}$	$486_{(1)}$	$506_{(1.04)}$	$496_{(1.02)}$	$486_{(1)}$	486	486.0	486.0	506.3
30	$486_{(1)}$	$486_{(1)}$	$511_{(1.05)}$	$489_{(1.01)}$	$486_{(1)}$	486	486.0	486.0	511.2
50	$486_{(1)}$	$486_{(1)}$	$513_{(1.05)}$	$488_{(1)}$	$486_{(1)}$	486	486.0	486.0	513.2
				$\alpha = 0.25$					
10	$488_{(1)}$	$520_{(1.06)}$	$556_{(1.13)}$	$497_{(1.02)}$	$489_{(1.01)}$	486	490.7	520.0	556.1
30	$490_{(1.01)}$	$532_{(1.08)}$	$570_{(1.15)}$	$491_{(1.01)}$	$491_{(1.01)}$	487	495.7	532.0	570.3
50	$492_{(1.01)}$	$538_{(1.09)}$	$576_{(1.17)}$	$490_{(1.01)}$	$493_{(1.01)}$	487	499.0	537.9	576.4
				$\alpha = 0.50$					
10	$507_{(1.02)}$	$588_{(1.14)}$	$609_{(1.19)}$	$515_{(1.04)}$	$507_{(1.02)}$	496	528.0	587.7	609.3
30	$534_{(1.05)}$	$636_{(1.17)}$	$643_{(1.21)}$	$536_{(1.05)}$	$536_{(1.05)}$	511	569.4	635.9	642.5
50	$550_{(1.05)}$	$667_{(1.19)}$	$668_{(1.2)}$	$564_{(1.08)}$	$562_{(1.08)}$	522	600.0	667.3	667.9
				$\alpha = 0.75$					
10	$549_{(1.04)}$	$674_{(1.18)}$	$677_{(1.2)}$	$562_{(1.07)}$	$552_{(1.05)}$	527	601.4	673.7	677.1
30	$620_{(1.05)}$	$818_{(1.17)}$	$818_{(1.17)}$	$670_{(1.14)}$	$654_{(1.11)}$	590	754.2	817.8	817.8
50	$686_{(1.05)}$	$959_{(1.15)}$	$959_{(1.15)}$	$788_{(1.21)}$	$756_{(1.16)}$	652	898.2	958.5	958.5
				$\alpha = 1$					
10	$604_{(1.04)}$	$780_{(1.19)}$	$780_{(1.19)}$	$631_{(1.09)}$	$614_{(1.06)}$	578	708.0	780.1	780.1
30	$773_{(1.05)}$	$1120_{(1.14)}$	$1120_{(1.14)}$	$876_{(1.19)}$	$828_{(1.12)}$	739	1057	1118	1119
50	$935_{(1.04)}$	$1460_{(1.11)}$	$1460_{(1.11)}$	$1130_{(1.25)}$	$1040_{(1.15)}$	899	1398	1457	1457

Table 4.4: Performance of truncated linear replenishment policy $T = 30$

	Simulated Inventory Runs						Objective Value		
b/h	TLRP	LRP	SRP	MP	BSP	OPT	Z_{TLRP}	Z_{LRP}	Z_{SRP}
				$\alpha = 0$					
10	$726_{(1)}$	$726_{(1)}$	$749_{(1.03)}$	$736_{(1.01)}$	$726_{(1)}$	725	725.6	725.6	748.9
30	$726_{(1)}$	$726_{(1)}$	$754_{(1.03)}$	$729_{(1)}$	$727_{(1)}$	726	725.6	725.6	754.4
50	$726_{(1)}$	$726_{(1)}$	$757_{(1.04)}$	$728_{(1)}$	$729_{(1)}$	726	725.6	725.6	756.7
				$\alpha = 0.25$					
10	$726_{(1)}$	$766_{(1.05)}$	$830_{(1.12)}$	$736_{(1.01)}$	$729_{(1)}$	725	726.8	765.6	829.6
30	$727_{(1)}$	$778_{(1.06)}$	$850_{(1.15)}$	$729_{(1)}$	$731_{(1.01)}$	726	728.5	777.7	850.4
50	$727_{(1)}$	$783_{(1.07)}$	$860_{(1.17)}$	$728_{(1)}$	$732_{(1.01)}$	726	729.7	783.3	859.2
				$\alpha = 0.50$					
10	$738_{(1.01)}$	$862_{(1.14)}$	$913_{(1.21)}$	$746_{(1.02)}$	$742_{(1.01)}$	732	755.7	861.6	913.4
30	$762_{(1.03)}$	$909_{(1.18)}$	$953_{(1.25)}$	$757_{(1.02)}$	$763_{(1.03)}$	743	792.5	908.6	952.7
50	$778_{(1.04)}$	$936_{(1.19)}$	$972_{(1.26)}$	$767_{(1.03)}$	$778_{(1.04)}$	750	815.6	935.6	972.0
				$\alpha = 0.75$					
10	$787_{(1.03)}$	$976_{(1.21)}$	$1000_{(1.26)}$	$789_{(1.03)}$	$786_{(1.03)}$	763	840.3	976.1	1004
30	$862_{(1.06)}$	$1100_{(1.24)}$	$1100_{(1.25)}$	$886_{(1.09)}$	$888_{(1.09)}$	816	963.8	1102	1103
50	$902_{(1.06)}$	$1190_{(1.23)}$	$1190_{(1.23)}$	$974_{(1.15)}$	$970_{(1.14)}$	849	1064	1194	1194
				$\alpha = 1$					
10	$857_{(1.05)}$	$1110_{(1.24)}$	$1120_{(1.26)}$	$868_{(1.06)}$	$863_{(1.06)}$	818	965.4	1115	1119
30	$1020_{(1.06)}$	$1412_{(1.21)}$	$1412_{(1.21)}$	$1119_{(1.17)}$	$1100_{(1.15)}$	957	1286	1412	1412
50	$1150_{(1.06)}$	$1700_{(1.18)}$	$1700_{(1.18)}$	$1370_{(1.26)}$	$1310_{(1.2)}$	1090	1587	1704	1704

Table 4.5: Performance of truncated linear replenishment policy $T = 5$ with and without directional deviations

		Simulated Inventory Runs			Objective Value	
α	b/h	TLRP directional deviations	TLRP no directional deviations	OPT	Z_{TLRP} directional deviations	Z_{TLRP} no directional deviations
0	10	$108_{(1.01)}$	$108_{(1.01)}$	107	108.0	108.0
	50	$108_{(1)}$	$108_{(1)}$	108	108.0	180.0
0.5	10	$110_{(1.02)}$	$110_{(1.02)}$	108	111.2	122.0
	50	$112_{(1.02)}$	$113_{(1.03)}$	109	116.7	175.0
1	10	$118_{(1.04)}$	$122_{(1.07)}$	113	132.3	159.4
	50	$140_{(1.06)}$	$163_{(1.24)}$	132	195.2	347.1

4.3 Influence of Directional Deviations

Table 4.5 shows a comparison of the TLRP with and without information on the directional deviations. In the latter case, the robust policies were obtained using the bound of Theorem 1 with information only on the support and covariance associated with the random factors. When $\alpha = 0$, information on directional deviations has little impact on the model objective. It is observed that TLRP gives an improvement when $\alpha = 1$ and $b/h = 50$. The additional computational burden posed by the directional deviations varies with the size of the model. For $T = 30$, the computational time of TLRP with and without the directional deviations are 197 seconds and 143 seconds, respectively. For $T = 20$, the computational time are 37.6 seconds and 30.5 seconds, respectively. For the $T = 5$ and $T = 10$ models, the computational time with and without the directional deviations are practically the same.

4.4 Effects of Demand Variability

We also investigated the influence of demand variability on the performance of the best robust policy, namely, TLRP. Shown in Table 4.6 are results of TLRP for the $T = 5$ model, with $\mu = 200$, $b/h = 50$, for $\alpha = 0$, $\alpha = 0.25$, $\alpha = 0.5$, and various degrees of variability, as reported by $\bar{z}$. When $\alpha = 0$, we are able to perform the experiments for larger coefficient of variations. The case of $\bar{z} = 200$ corresponds to coefficient of variation being 0.58. We observe that the bound of Z_{TLRP} degrades significantly as demand variability increases. However, the impact on the performance against the optimum policy is marginal, which is rather surprising given the fact that we use significantly less distributional information in our demand model.

One may find in industry demands with coefficients of variation of four and even higher. The coefficient of variation in our computational studies is limited by the random factors being uniform distributed. To achieve larger values of σ/μ, we assumed that demands across periods are 2-point distributed i.i.d random variables. The demand at each period is zero with probability β and 200 with probability $1-\beta$. The parameter β controls the coefficient of variation and the other parameters used were $S_t = 260, \alpha = 0, b/h = 50$. We compare the performance of TLDR with and without directional deviations and present the results in Table 4.7. For both cases, the results are similar with the bound of Z_{TLRP} coming close to OPT. TLRP performs very well when demands have very high coefficient of variations. The phenomenon that the robust optimization performs well when uncertainties have very high coefficient of variations has also been observed in the computational studies of Chen et al. (2008).

4.5 Analysis of Policies

Although the robust models appear to be complex, implementing the policy derived from the model is extremely easy. The truncated linear replenishment policy is computed simply by taking an affine sum of random factors using weights given by the TLRP model solution and then restricting the range of the order quantity. For

Table 4.6: Performance of truncated linear replenishment policy $T = 5$ with various demand range

		Simulated Inventory Runs		Objective Value
α	$\bar{z}$	TLRP	OPT	Z_{TLRP}
0	20	$104_{(1)}$	104	104.0
	40	$108_{(1)}$	108	108.0
	60	$112_{(1)}$	112	112.0
	80	$125_{(1.03)}$	118	131.8
	100	$136_{(1.04)}$	131	161.3
	120	$158_{(1.03)}$	153	212.8
	160	$244_{(1.02)}$	239	385.8
	200	$380_{(1.02)}$	372	629.2
0.25	20	$102_{(1)}$	102	102.0
	40	$108_{(1)}$	108	108.8
	60	$119_{(1.04)}$	114	128.9
	80	$142_{(1.05)}$	135	190.7
	100	$195_{(1.04)}$	187	306.4
0.5	20	$104_{(1)}$	104	104.0
	40	$112_{(1.03)}$	109	116.7
	60	$139_{(1.06)}$	131	187.9

Table 4.7: Performance of truncated linear replenishment policy under a 2-point demand distribution

		Simulated Inventory Runs			Objective Value	
β	σ/μ	TLRP directional deviations	TLRP no directional deviations	OPT	Z_{TLRP} directional deviations	Z_{TLRP} no directional deviations
				$T = 5$		
0.80	2.00	$52_{(1)}$	$52_{(1)}$	52	52.0	52.0
0.85	2.38	$49_{(1)}$	$49_{(1)}$	49	49.0	49.0
0.90	3.00	$46_{(1)}$	$46_{(1)}$	46	46.0	46.0
0.95	4.36	$43_{(1)}$	$43_{(1)}$	43	43.0	43.0
0.98	7.00	$41_{(1)}$	$41_{(1)}$	41	41.2	41.2
				$T = 10$		
0.80	2.00	$88_{(1)}$	$88_{(1)}$	88	88.0	88.0
0.85	2.38	$81_{(1)}$	$81_{(1)}$	81	81.0	81.0
0.90	3.00	$74_{(1)}$	$74_{(1)}$	74	74.0	74.0
0.95	4.36	$67_{(1)}$	$67_{(1)}$	67	67.0	67.0
0.98	7.00	$63_{(1)}$	$63_{(1)}$	63	62.8	62.8

Table 4.8: A sample path of the truncated linear replenishment policy

t	z_t	d_t	x_t^{TRLP}	y_{t+1}
1	13.3	213.3	260.0	46.7
2	6.2	212.9	214.5	48.4
3	14.5	224.2	224.8	48.9
4	-24.1	192.9	249.6	105.6
5	-32.3	172.7	139.5	72.5

example, a sample problem where $\alpha = 0.5$ has the following model solution:

$$\begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \\ x_5^0 \end{bmatrix} = \begin{bmatrix} 260.00 \\ 191.93 \\ 218.29 \\ 243.57 \\ 126.31 \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \boldsymbol{x}_3' \\ \boldsymbol{x}_4' \\ \boldsymbol{x}_5' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & 0 \\ 1.7 & 0 & 0 & 0 & 0 \\ 0 & 1.04 & 0 & 0 & 0 \\ 1.29 & 1.44 & 1.6 & 1.5 & 0 \end{bmatrix}.$$

Table 4.8 shows the sample path, constructed using weights from the model solution, and then applying the relevant capacity constraints,

$$x_i^{TLRP}(\boldsymbol{z}) = \min\{(x_i^0 + \boldsymbol{x}_i'\boldsymbol{z})^+, 260\}.$$

In the above example, the inventory manager would order a quantity of 260, 215, 225, 250 and 140 for periods 1 to 5, respectively.

Ben-Tal et al. (2005) showed that the linear replenishment policy is equivalent to a history-independent base-stock if and only if it exhibits Markovian behavior and takes the form $x_t(\tilde{\boldsymbol{d}}_{t-1}) = x_t^0 + \tilde{z}_{t-1}$. The truncated linear replenishment policy has a different structure and in general, we are unable to show the connection with a base-stock structure. When the demands are independent, that is, $\alpha = 0$, it is observed that TLRP exhibits Markovian behavior for most input parameters. There are also instances that LRP is Markovian when the TLRP is not. For example, for $T = 10, \mu = 220, \alpha = 0, b/h = 40, \bar{z} = 40$, TLRP and LRP are the same and having a Markovian structure. See Table 4.9. However, when $\bar{z} = 80$, the TLRP and LRP policies presented in Table 4.10 and Table 4.11, respectively, show a difference in the structure. For the case of correlated demands, we did not observe any Markovian structure in our experiments.

Table 4.9: TLRP and LRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 40, S_t = 240$

$$\begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \\ x_5^0 \\ x_6^0 \\ x_7^0 \\ x_8^0 \\ x_9^0 \\ x_{10}^0 \end{bmatrix} = \begin{bmatrix} 240 \\ 220 \\ 220 \\ 220 \\ 220 \\ 220 \\ 220 \\ 220 \\ 220 \\ 220 \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \boldsymbol{x}_3' \\ \boldsymbol{x}_4' \\ \boldsymbol{x}_5' \\ \boldsymbol{x}_6' \\ \boldsymbol{x}_7' \\ \boldsymbol{x}_8' \\ \boldsymbol{x}_9' \\ \boldsymbol{x}_{10}' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

4.6 Computational Time

We formulated the robust models using an in-house developed software, *PROF* (Platform for Robust Optimization Formulation). The Matlab based software is essentially a algebraic modeling language for robust optimization that contains reusable functions for modeling multiperiod robust optimization using decision rules. After formulating the model, it calls upon a commercial SOCP solver, MOSEK 5.0 for solution. We have implemented bounds for $\pi(\cdot)$ of Theorem 1 and $\eta(\cdot)$ of Theorem 4. The sample formulation of Problem (3.17) provided in Appendix F shows the ease of formulating the TLRP model using the software. The size of the problem we considered is presented in Table 4.12. Our computation was carried out on a 2.4 GHz desktop with 2 Gb memory. The computational time depends on the number of periods. It typically takes less than 0.3 seconds to solve the TLRP model for $T = 5$. For $T = 10, 20$ and 30, the times taken were 3 seconds, 30 seconds and 3 minutes, respectively, suggesting that the computational time scales reasonably well with respect to the size of the problem. Moreover, the time needed for computation does not depend on the replenishment lead time, demand variability, and correlations. On the other hand, much of the computational effort lies in solving the optimum history dependent policy using dynamic programming. In the experiments, we have customized and optimized the dynamic programming

Table 4.10: TLRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 80, S_t = 240$

$$\begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \\ x_5^0 \\ x_6^0 \\ x_7^0 \\ x_8^0 \\ x_9^0 \\ x_{10}^0 \end{bmatrix} = \begin{bmatrix} 240.0 \\ 231.0 \\ 231.4 \\ 231.7 \\ 232.2 \\ 232.9 \\ 233.7 \\ 234.9 \\ 236.4 \\ 135.7 \end{bmatrix},$$

$$\begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \boldsymbol{x}_3' \\ \boldsymbol{x}_4' \\ \boldsymbol{x}_5' \\ \boldsymbol{x}_6' \\ \boldsymbol{x}_7' \\ \boldsymbol{x}_8' \\ \boldsymbol{x}_9' \\ \boldsymbol{x}_{10}' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.22 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.21 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.20 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.17 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.16 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.13 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.09 & 0 & 0 \\ 0.78 & 0.78 & 0.79 & 0.80 & 0.82 & 0.84 & 0.87 & 0.91 & 1.00 & 0 \end{bmatrix}.$$

Table 4.11: LRP for $\mu = 220, \alpha = 0, T = 10, b/h = 40, \bar{z} = 80, S_t = 240$

$$\begin{bmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \\ x_5^0 \\ x_6^0 \\ x_7^0 \\ x_8^0 \\ x_9^0 \\ x_{10}^0 \end{bmatrix} = \begin{bmatrix} 240.0 \\ 229.6 \\ 229.8 \\ 229.4 \\ 229.2 \\ 229.5 \\ 229.6 \\ 229.9 \\ 230.3 \\ 230.8 \end{bmatrix},$$

$$\begin{bmatrix} \boldsymbol{x}'_1 \\ \boldsymbol{x}'_2 \\ \boldsymbol{x}'_3 \\ \boldsymbol{x}'_4 \\ \boldsymbol{x}'_5 \\ \boldsymbol{x}'_6 \\ \boldsymbol{x}'_7 \\ \boldsymbol{x}'_8 \\ \boldsymbol{x}'_9 \\ \boldsymbol{x}'_{10} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.26 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.26 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.26 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.26 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.26 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.26 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.24 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 \end{bmatrix}.$$

algorithm so that we can reduce the computational time to less than three hours. For instance, we implemented the Golden-section search method and exploited the fact that $v_t = \mu + \alpha\tilde{z}_1 + \cdots + \alpha\tilde{z}_t \in [\mu - t\alpha\underline{z}, \mu + t\alpha\bar{z}]$ to reduce the size of the state space. Table 4.13 compares the computational times of the TLRP model against the optimum dynamic programming model.

4.7 Summary

In this chapter, we studied the computational performance of the static, linear and truncated linear replenishment policies against the optimum history dependent policy and two dynamic programming based heuristics, namely, the myopic policy and a history-independent base-stock policy. Our computational results suggest that the truncated linear replenishment policy, together with information on the directional deviations, yield reasonably good solutions against the optimum and give the best overall performance among linear and static policies and simple dynamic programming based heuristics. Specifically, the contribution of our approach over the related works of Bertsimas and Thiele (2006) and Ben-Tal et al. (2005) can be summarized as follows.

- Our proposed robust optimization approximation is based on a comprehensive factor-based demand model that captures correlations such as the autoregressive nature of demand, the effect of external factors, as well as trends and seasonality, among others. In addition, we cater for distributional ambiguity in the underlying factors by considering a family of distributions characterized by the mean, covariance, support and directional deviations. In contrast, the robust optimization model of Bertsimas and Thiele (2006) is restricted to independent demands with an identical mean and variance, while the model of Ben-Tal et al. (2005) is confined to completely distribution-free demand uncertainty.

- We propose a new policy called the truncated linear replenishment policy, which gives improved approximation to the multiperiod inventory control problem over static and linear decision rules used in the robust optimization proposals of Bertsimas and Thiele (2006) and Ben-Tal et al. (2005),

Table 4.12: Size of the TLRP model, where $\mathcal{L}^n = \{(x_0, \boldsymbol{x}) \in \Re \times \Re^{n-1} \;:\; \|\boldsymbol{x}\|_2 \leq x_0\}$.

T	5	10	20	30
Affine constraints	5911	24721	125041	348961
Free variables	3366	15806	89911	266316
nonnegative variables	1700	9000	56400	174200

T	5	10	20	30	T	5	10	20	30
$\mathcal{L}^2$ cones	12	12	12	12	$\mathcal{L}^{18}$ cones	-	-	206	206
$\mathcal{L}^3$ cones	1226	4026	14426	31226	$\mathcal{L}^{19}$ cones	-	-	218	218
$\mathcal{L}^4$ cones	38	38	38	38	$\mathcal{L}^{20}$ cones	-	-	230	230
$\mathcal{L}^5$ cones	50	50	50	50	$\mathcal{L}^{21}$ cones	-	-	282	242
$\mathcal{L}^6$ cones	72	62	62	62	$\mathcal{L}^{22}$ cones	-	-	102	254
$\mathcal{L}^7$ cones	27	74	74	74	$\mathcal{L}^{23}$ cones	-	-	-	255
$\mathcal{L}^8$ cones	-	86	86	86	$\mathcal{L}^{24}$ cones	-	-	-	278
$\mathcal{L}^9$ cones	-	98	98	98	$\mathcal{L}^{25}$ cones	-	-	-	290
$\mathcal{L}^{10}$ cones	-	110	110	110	$\mathcal{L}^{26}$ cones	-	-	-	302
$\mathcal{L}^{11}$ cones	-	142	122	122	$\mathcal{L}^{27}$ cones	-	-	-	314
$\mathcal{L}^{12}$ cones	-	52	134	134	$\mathcal{L}^{28}$ cones	-	-	-	326
$\mathcal{L}^{13}$ cones	-	-	146	146	$\mathcal{L}^{29}$ cones	-	-	-	338
$\mathcal{L}^{14}$ cones	-	-	158	158	$\mathcal{L}^{30}$ cones	-	-	-	350
$\mathcal{L}^{15}$ cones	-	-	170	170	$\mathcal{L}^{31}$ cones	-	-	-	422
$\mathcal{L}^{16}$ cones	-	-	182	182	$\mathcal{L}^{32}$ cones	-	-	-	152
$\mathcal{L}^{17}$ cones	-	-	194	194					

Table 4.13: Computational time

T	5	10	20	30
TLRP	0.3 sec	3 sec	30 sec	3 min
OPT, $\alpha = 0$	5 sec	18 sec	25 sec	85 sec
OPT, $\alpha = 1$	12 min	30 min	1.5 hr	2.5 hr

respectively. We *do not* restrict the policy structure to base-stock. We have developed a new bound on a nested sum of expected positive values of random variables and show that the parameters of the truncated linear replenishment policy can be obtained by solving a tractable deterministic mathematical optimization problem in the form of a SOCP, whose solution time is independent on replenishment lead time, demand variability, and correlations.

Chapter 5

Safeguarding Fill Rate Against Distributional Uncertainty

This chapter kicks off the second part of the thesis where we propose an approach to optimize fill rate using descriptive statistics so as to assure that a high fill rate is achieved even when there is distributional uncertainty. That is, the order quantity needs to achieve an expected fill rate target for a family of distributions with the same demand range, demand median and range of the probability density function. Whereas part one discusses a single-product multiperiod problem, the problem here is essentially a *single-period multiproduct* one. The goal of the chapter is to discuss what an appropriate uncertainty set would be, based on the information we could glean from historical data, and the quality of the robust model as seen in the price to pay for incorporating robustness. The latter is important because while robustness is desired, robust models should not be overly expensive.

In most practical settings, the distribution of demand is seldom known exactly but only approximately. In particular, given a set of empirical data, it is common to find not one but several possible fits to the distribution. Uncertainty in type of demand distribution may also arise with changing trends. For example, many consumer goods are known to exhibit seasonal variation with demand distributed differently over time. Motivated by the practical need to incorporate uncertainty in the type of demand distribution, this chapter proposes a model to optimize fill rate using descriptive statistics.

In this model, it is assumed that the demand is bounded in $[\underline{d}, \bar{d}]$, which is realistic in many practical settings. The demand probability density function, pdf, is denoted by $h(t)$ and we use $\mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})$ to denote the family of demand distributions with support $[\underline{d}, \bar{d}]$ and pdf bounded within $[\underline{h}, \bar{h}]$. We adopt the following notation where the subscript $j, j = 1 \dots p$ denotes the jth product; for the single-product case, the subscript may be omitted. Other notation are as follows

- x_j: order quantity
- $\tilde{d}_j$: stochastic exogenous demand
- m_j: median of the demand
- c_j: unit ordering cost
- τ_j: fill rate target

5.1 Multiproduct Fill Rate Model

5.1.1 Definition of Fill Rate

Two different notions of fill rate are commonly encountered in supply chain management - case fill rate and line fill rate. Case fill rate measures the quantity of cases filled as a proportion of cases ordered, usually as of the initial shipment in fulfillment of an order. On the other hand, line fill rate measures the number of line items that are completely (100%) filled, divided by the total number of line items ordered, see for instance `http://www.supplychainmetric.com/fillrate.thm`. Our model is concerned with case fill and not line fill. Throughout this thesis, we use the term *fill rate* to denote the notion of case fill rate. More formally, given a stock of x, the fill rate of inventory is the proportion of demand satisfied by on-hand stocks. Note that fulfillment can never exceed the demand, and hence fill rate takes a maximum value of one. Against uncertain demands, the instantaneous fill rate of inventory can be written as $\min\{x, \tilde{d}\}/\tilde{d} = \min\{x/\tilde{d}, 1\}$. Taking expectation gives the expected fill rate, $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$. Observe that this measures a different quantity from one minus the stockout probability. In particular, when demand is fully met, we attach a utility (or value) of one, else we take the proportion of fulfillment x/d

as the utility. In the case of one minus the stockout probability, the utility is one when demand is fully met, and zero otherwise. The expected fill rate is not strictly a probability and attaches some value to partially filled demands. While one minus the stockout probability does not distinguish 1% fulfillment from 99% fulfillment, fill rate does not have this issue and has been used by service critical industries to quantify the extent of service delivered. One other advantage of expected fill rate is that the function is concave which results in a convex optimization problem that can be solved efficiently.

In practice, it is not uncommon for different products sharing the same budget basket to have varying importance. This happens frequently in spare-part provisioning, where items are extremely wide-ranging. On one end are low-cost substitutable products such as screws and gaskets, while on the other extreme are expensive specialized parts. Given known weights that sum to unity representing the relative importance, we can define an aggregate expected fill rate as follows:

$$\sum_{j=1}^{p} \lambda_j \mathrm{E}\left(\min\{x_j/\tilde{d}_j, 1\}\right), \quad \text{with } \lambda_j \geq 0, \quad \sum_{j=1}^{p} \lambda_j = 1.$$

The aggregate fill rate gives an overall measure of performance of inventory across the basket of products. In the absence of concrete linkages between the products, λ is used as a proxy to combine the item fill rate into an aggregate system level fill rate. The weighting can be based on the importance of the items, or heuristically derived from quantities such as the mean demand. For example, if the demand of item i is on average twice that of item j, one could assert that $\lambda_i = 2\lambda_j$. When the items are equally valued, the multiproduct problem aims to minimize inventory purchase costs across all product types against fill rate constraints at the item level, τ_j and the system level, τ_0. In some contexts λ is not used and the system level fill rate is taken as the minimum of the item fill rates, $\tau_0 = \min(\tau_1, \ldots, \tau_p)$. This is common in military inventory management where the items may represent essentials like food, water, fuel, and so forth. Only when all the items are available will the military unit be operational, which explains the rationale of using the minimum. An illustration of λ and system fill rate is provided in the multiproduct model of Problem (5.5), as well as the example in Section 6.3. We now introduce the notion of distributional uncertainty, before discussing the model.

5.1.2 Distributional Uncertainty

The distribution function of the demand is an input required for computation of expected fill rate. Much literature assumes the function to be known and forecast from historical data. The data requirements for estimating the distribution can be inferred from the Dvoretzky-Kiefer-Wolfowitz inequality (see Dvoretzky et al. (1956), Massart (1990)) which states that

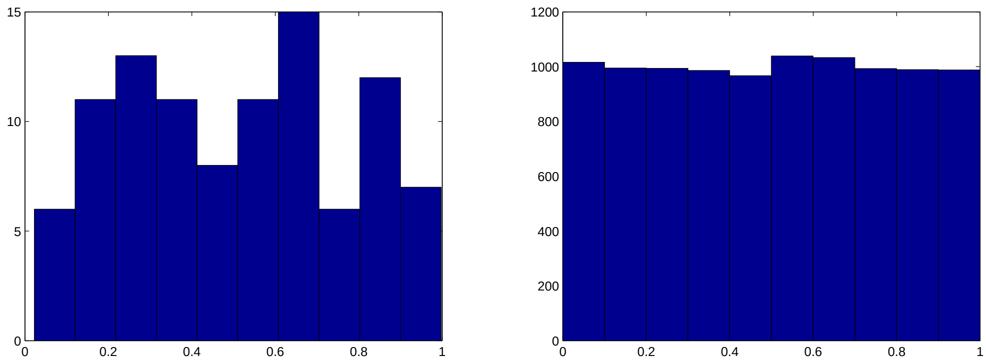
$$\Pr(\sup_t |\hat{F}_n(t) - F(t)| > \epsilon) \leq 2e^{-2n\epsilon^2}. \tag{5.1}$$

In the above, $\hat{F}_n$ is the associated empirical distribution function computed using n samples and F is the true cdf. This inequality is a classical result in statistical and probability literature used to compute the sample size needed to guarantee that the estimation of the cumulative distribution function is accurate. It should be highlighted that the bound is tight, see Massart (1990). Application of the inequality shows that in order to obtain accuracy of 0.01 with 0.99 confidence, a sample size of $n \geq 26491$ is needed.

To evaluate the accuracy of the estimation, simulations were carried out using a uniform distribution in $(0, 1)$ with sample size of n using Matlab. We denote the ith sample of the data by y_i, and the ith *ordered sample* by $y_{(i)}$. The empirical distribution function is given by $\hat{F}_n(t) = 1/n \sum_{i=1}^n 1_{(y_i \leq t)}$, where $1_{(y_i \leq t)}$ is an indicator variable. Observe that $\hat{F}_n$ is a step function with discrete jumps. Instead of deriving the worst-case error as per Equation (5.1), we derived the error at the jump points. Observe that at the ith jump point, F is the ith ordered statistics of the sample and $\hat{F}_n = i/n$, so the error is $\zeta_i = |\hat{F}_n(i/n) - F(i/n)| = |i/n - y_{(i)}|$. The quantity $\max_i \zeta_i$ approximately corresponds to but is less than ϵ of Equation (5.1), and it is clear that $\max_i \zeta_i \to \epsilon$ as $n \to \infty$. The quantity, $\sum_i \zeta_i/n$, gives an indication of the mean error. The simulation was repeated N times. For ease of exposition, we use superscript ζ^j to denote the outcome of the jth simulation. Table 5.1 shows the result of $N = 100000$ simulations for a range of sample sizes. The second column shows the maximum error while the third column shows the mean error, averaged over N simulations. The variance of the error (averaged over N simulations) is given in the last column, and we have used n rather than $n - 1$ in the computation for simplicity.

Table 5.1: Errors in the estimation of distribution (obtained from simulation)

	Max error	Mean error	Variance of error
Sample size (n)	$\frac{1}{N}\sum_j \max_i \zeta_i^j$	$\frac{1}{N}\sum_j \sum_i \zeta_i^j/n$	$\frac{1}{N}\sum_j \left(\sum_i (\zeta_i^j)^2/n - \sum_i \zeta_i^j/n\right)$
50	0.11	0.045	10×10^{-4}
100	0.081	0.032	5.0×10^{-4}
1000	0.026	0.010	5.0×10^{-5}
5000	0.012	0.0044	1.0×10^{-5}
10000	0.0057	0.0018	1.6×10^{-6}



a) 100 samples b) 10000 samples

Figure 5.1: Uncertainty in distribution type

To visualize the distributions, we provided in Figure (5.1) the histograms of $n = 100$ and $n = 10000$. For the $n = 100$ case, $\max_i \zeta_i$ and $\sum_i \zeta_i/n$ are 0.073 and 0.024, respectively. Comparing these with Table 5.1, it is evident that Figure (5.1) is a typical scenario one would encounter in practice.

Given the difficulty of knowing the density function precisely, we aim to derive order quantities that will safeguard against variation in the distribution type, which could be due to limited sample size used in forecasting the demand. That is, rather than computing the fill rate using the (limited) demand data and assuming it to represent the true distribution well, we want some assurance that the solution obtained from the model is valid, even when using limited samples of demand data. A common approach in robust inventory literature is to allow $\tilde{d}$ to assume some possible distributions with the same descriptive statistics. As discussed earlier in

Chapter 3, the use of descriptive statistics is not new in the inventory literature. It was used by Scarf (1958), where he considered all possible demand distributions with the same first and second moments. We next show that it is not practical to use the moment approach for expected fill rate. To assure high fill rate, we need to obtain a lower bound using the first two moments of $\tilde{d}$. We can use the fact that for any random variables $\tilde{a}$, $\tilde{b}$,

$$\begin{aligned} \mathrm{E}\left(\min\{\tilde{a},\tilde{b}\}\right) &= \tfrac{1}{2}\mathrm{E}\left(\tilde{a}\right) + \tfrac{1}{2}\mathrm{E}\left(\tilde{b}\right) - \tfrac{1}{2}\mathrm{E}\left(|\tilde{b}-\tilde{a}|\right) \\ &\geq \tfrac{1}{2}\mathrm{E}\left(\tilde{a}\right) + \tfrac{1}{2}\mathrm{E}\left(\tilde{b}\right) - \tfrac{1}{2}\sqrt{\mathrm{E}\left((\tilde{b}-\tilde{a})^2\right)}, \end{aligned} \tag{5.2}$$

where the last inequality is due to $\mathrm{E}\left(|\tilde{r}|\right) \leq \sqrt{\mathrm{E}\left(\tilde{r}^2\right)}$ (Jensen's inequality). So,

$$\begin{aligned} \mathrm{E}\left(\min\{\tfrac{x}{\tilde{d}},1\}\right) &\geq \tfrac{1}{2}x\mathrm{E}\left(\tfrac{1}{\tilde{d}}\right) + \tfrac{1}{2} - \tfrac{1}{2}\sqrt{\mathrm{E}\left((1-\tfrac{x}{\tilde{d}})^2\right)} \\ &= \tfrac{1}{2}x\mathrm{E}\left(\tfrac{1}{\tilde{d}}\right) + \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1-2x\mathrm{E}\left(1/\tilde{d}\right)+x^2\mathrm{E}\left(1/\tilde{d}^2\right)}. \end{aligned} \tag{5.3}$$

To use Inequality (5.3), we would need the quantities $\mathrm{E}\left(1/\tilde{d}\right)$ and $\mathrm{E}\left(1/\tilde{d}^2\right)$. Only in special distributions, such as those with reciprocal symmetry, see Seshadri (1964), would we be able to derive $\mathrm{E}\left(1/\tilde{d}\right)$ from $\mathrm{E}\left(\tilde{d}\right)$, and $\mathrm{E}\left(1/\tilde{d}^2\right)$ from $\mathrm{E}\left(\tilde{d}^2\right)$. Distributions with reciprocal symmetry are often too limited to realistically model real-life demand distributions. They are often unbounded from the right and truncating them as a remedy would destroy the reciprocal property. An alternative is to use the fact that for any nonnegative random variable $\tilde{d}$, Jensen's inequality gives $\mathrm{E}\left(1/\tilde{d}\right) \geq 1/\mathrm{E}\left(\tilde{d}\right)$. Applying to Inequality (5.3), we have

$$\mathrm{E}\left(\min\left\{\frac{x}{\tilde{d}},1\right\}\right) \geq \frac{1}{2}x/\mathrm{E}\left(\tilde{d}\right) + \frac{1}{2} - \frac{1}{2}\sqrt{1-2x/\mathrm{E}\left(\tilde{d}\right)+x^2\mathrm{E}\left(1/\tilde{d}^2\right)}.$$

A further bound on $\mathrm{E}\left(1/\tilde{d}^2\right)$ is still required, which means that the result will be very much weakened and not likely to be useful in practice.

5.1.3 Descriptive Statistics & Uncertainty Sets

Given this difficulty, there is a need to explore the use of other descriptive statistics to construct the lower bound. Our approach is to deliberately incorporate safeguards into the models by using more amenable descriptive statistics. For instance, given that the demand may assume one of the two possible distributions shown in Figure (5.2), the following descriptive statistics can be used as input parameters for optimization:

- $\underline{d} = 1, \bar{d} = 100$,
- $\underline{h} = 0, \bar{h} = 0.03$,
- $71.0 \leq m(\tilde{d}) \leq 79.6$, since the median of case 1 demand is 71.0, and the median of case 2 demand is 79.6.

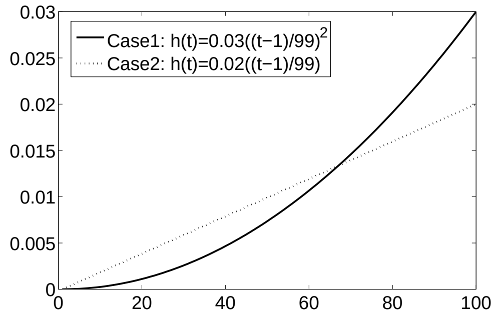


Figure 5.2: Two possible distributions

The key idea is to provision for stocks to achieve a high fill rate over a family of distributions wide enough to contain the true distribution which is uncertain. Our approach is to develop bounds using the range of the demand and the range of the pdf. That is, bounds of $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right) \forall \tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})$. The quantity $\underline{h}, \bar{h}$ are bounds on the pdf and can be estimated from empirical data by way of a histogram.

In the case of a perfectly uniform distribution, $\underline{h} = \bar{h} = \hat{h} \triangleq 1/(\bar{d} - \underline{d})$. So for evenly distributed data, an approach is to estimate $\underline{h}, \bar{h}$ with respect to some deviation δ from the uniform distribution. Specifically, we add a buffer to the range of pdf as a means to safeguard against distributional uncertainty:

$$\underline{h} = (1 - \delta)\hat{h}, \qquad \bar{h} = (1 + \delta)\hat{h}, \qquad 0 \leq \delta \leq 1. \tag{5.4}$$

Clearly, large δ implies a larger family of distributions and more conservative results. This approach closely resembles the uncertainty set approach in robust optimization which has gained substantial acceptance as a tool to manage uncertainty.

Besides $\bar{h}$ and $\underline{h}$, it will be seen shortly that another useful statistics is the median, the discussion of which we postpone to a later section. Given these statistics, the family of distribution could consist of all $\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})$, such that $\underline{m} \leq m(\tilde{d}) \leq \bar{m}$, where $\underline{m}$ and lower and upper bounds to the median, respectively.

We are interested in an order quantity of $x \in (0, \bar{d})$. Outside this range, the optimal value for the fill rate becomes trivial. The multiproduct fill rate optimization problem with demand information given by descriptive statistics can be described as follows:

$$
\begin{aligned}
\min \quad & \sum_{j=1}^{p} c_j x_j \\
\text{s.t.} \quad & \inf_{\substack{\tilde{d}_j \in \mathcal{D}(\underline{d}_j, \bar{d}_j, \underline{h}_j, \bar{h}_j) \\ \underline{m}_j \leq m(\tilde{d}_j) \leq \bar{m}_j}} \mathrm{E}\left(\min\left\{\frac{x_j}{\tilde{d}_j}, 1\right\}\right) \geq \tau_j \qquad\qquad j = 1 \ldots p \\
& \sum_{j=1}^{p} \lambda_j \inf_{\substack{\tilde{d}_j \in \mathcal{D}(\underline{d}_j, \bar{d}_j, \underline{h}_j, \bar{h}_j) \\ \underline{m}_j \leq m(\tilde{d}_j) \leq \bar{m}_j}} \mathrm{E}\left(\min\left\{\frac{x_j}{\tilde{d}_j}, 1\right\}\right) \geq \tau_0 \\
& 0 \leq x_j \leq \bar{d}_j \qquad\qquad j = 1 \ldots p.
\end{aligned} \tag{5.5}
$$

The first set of constraints stipulate targets on the item fill rates, while the second constraint stipulates a target on the aggregate fill rate. When there is no requirement for the aggregate fill rate to be higher than some specified target, we can set $\tau_0 = 0$ to render the system level fill rate constraint inactive. The same applies to the item level fill rate. The above is a single-period model. To extend it to multiple periods, one needs to insert fill rate constraints and inventory balance equations at the end of each period, see for instance Zipkin (2000).

5.2 Fill Rate Bounds

This section discusses the fill rate bounds. We start off with a basic bound, after which we show how it can be improved.

5.2.1 Bounds using the Pdf Range

The main issue is the expected fill rate expression $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$. Observe that

$$\min\{x/\tilde{d}, 1\} = 1 + \min\{x/\tilde{d} - 1, 0\} = 1 - \max\{1 - x/\tilde{d}, 0\} = 1 - (1 - x/\tilde{d})^+.$$

When the pdf is constant in (a, b), that is $h(t) = h\ \forall t \in (a, b)$, we have the following for $x \leq b$:

$$\begin{aligned}
\int_a^b (\min\{1, x/t\}) h \, \mathrm{d}t &= \int_a^b \left(1 - (1 - x/t)^+\right) h \, \mathrm{d}t \\
&= h(b-a) - h \int_{\max\{x,a\}}^b 1 - x/t \, \mathrm{d}t \\
&= h(b-a) - h\Big(b - \max\{x, a\} - x\ln(b) + x\ln\left(\max\{x, a\}\right)\Big) \\
&= h\max\{x - a, 0\} + h\min\Big\{-x\ln(x), -x\ln(a)\Big\} + hx\ln(b).
\end{aligned} \tag{5.6}$$

Notice that $(1 - x/t)^+ = 0$ when $x > b$, and therefore

$$\int_a^b (\min\{1, x/t\}) h \, \mathrm{d}t = \kappa(x, a, b, h) \triangleq \begin{cases} h(x-a)^+ + h\min\Big\{-x\ln(x), -x\ln(a)\Big\} + hx\ln(b) & x \leq b \\ h(b-a) & x > b. \end{cases} \tag{5.7}$$

With the kappa function, we now present the basic bounds on the expected fill rate function.

Theorem 6 *For demand with support in $(\underline{d}, \bar{d})$, and pdf bounded in $[\underline{h}, \bar{h}]$, the expected fill rate is bounded by the quantities below.*

1. *Lower bound,*

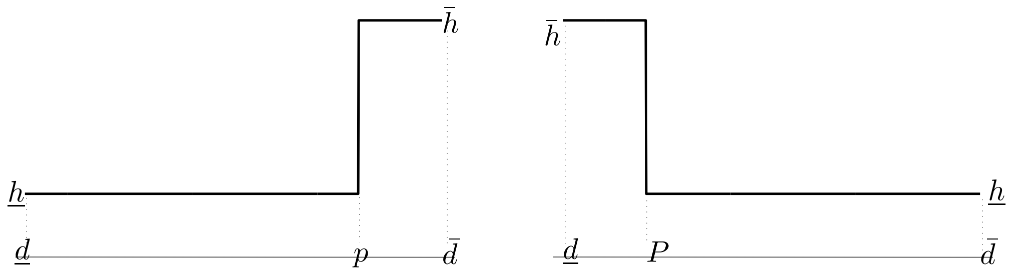
$$\inf_{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})} E\Big(\min\{x/\tilde{d}, 1\}\Big) = \kappa(x, \underline{d}, p, \underline{h}) + \kappa(x, p, \bar{d}, \bar{h}), \quad \text{where} \quad p = \frac{\bar{h}\bar{d} - \underline{h}\underline{d} - 1}{\bar{h} - \underline{h}}. \tag{5.8}$$

2. *Upper bound,*

$$\sup_{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})} E\Big(\min\{x/\tilde{d}, 1\}\Big) = \kappa(x, \underline{d}, P, \underline{h}) + \kappa(x, P, \bar{d}, \bar{h}), \quad \text{where} \quad P = \frac{1 - \underline{h}\bar{d} + \bar{h}\underline{d}}{\bar{h} - \underline{h}}. \tag{5.9}$$

Proof :

1. For $\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})$, the worst-case fill rate for any order quantity $x \in (0, \bar{d})$, corresponds to the case when the demand (amongst the family of distributions in $\mathcal{D}$) is the largest. This happens for a pdf with the maximum mass packed to the right. When this happens, the worst-case fill rate is achieved, which means that the corresponding upper bound constructed from the pdf is tight.



a) Worst-case Fill Rate b) Best-case Fill Rate

Figure 5.3: Fill rate cases

Figure (5.3a) shows such a distribution. Observe that it consists of two piecewise uniform distributions. Let p be the intersection point of the two uniform distributions. From the geometry, using $\int h(t)\mathrm{d}t = 1$ we obtain $p = \frac{\bar{h}\bar{d} - \underline{h}\underline{d} - 1}{\bar{h} - \underline{h}}$, and the result follows by Equation (5.7).

2. Similarly, the best-case fill rate is achieved by a distribution with the maximum mass packed to the left. See Figure (5.3b). From the geometry, $P = \frac{1 - \underline{h}\bar{d} + \bar{h}\underline{d}}{\bar{h} - \underline{h}}$ and the result follows. ■

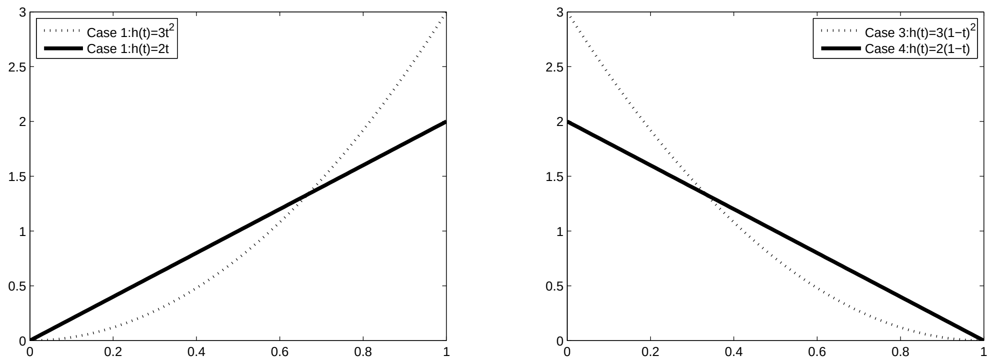
The quantity

$$\sup_{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})} \mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right) - \inf_{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h})} \mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$$

bounds the variation of fill rate within the family of distributions, and provides insights on the price incurred when incorporating distributional uncertainty in the model.

5.2.2 Bounds using the Median

The bounds of Theorem 6 are achieved by the distributions shown in Figure 5.3, which means that the bounds are tight. However, the price to pay for incorporating the uncertainty in the distribution type may be high, especially when $\bar{h} - \underline{h}$ is large. Means to reduce the price of distributional uncertainty are hence necessary to obtain a practical model. Given that $\tilde{d}$ may take one of the two possible distributions of Figure (5.4a), it would make sense to exclude other oppositely skewed distributions, such as those of Figure (5.4b) even though they have the same demand range and pdf range. Observe that the mean demands of the distributions of



a) Left-skewed Demand   b) Right-skewed Demand

Figure 5.4: Demand cases

Figure (5.4a) are substantially larger than those of Figure (5.4b). This suggests the approach of constraining the family of distributions such that the mean demand lies within a restricted range $u \leq \mathrm{E}\left(\tilde{d}\right) \leq U$.

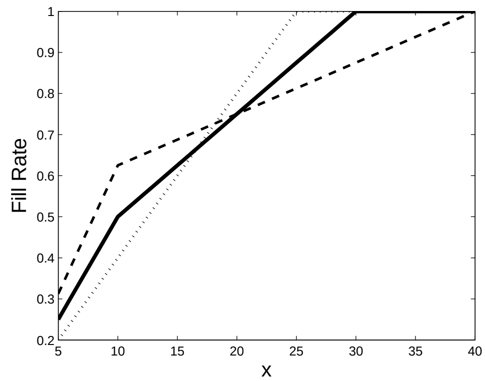


Figure 5.5: Fill rate for three demand distributions with mean of 25

Figure (5.5) shows the expected fill rate for three distributions all with the same mean of 25, as follows.

1. $\tilde{d} = 40$ with probability 0.5, and 10 with probability 0.5,

2. $\tilde{d} = 30$ with probability 0.75, and 10 with probability 0.25,

3. $\tilde{d} = 25$ with probability 1.

As is evident from the figure, $\inf\limits_{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h}), \mathrm{E}(\tilde{d}) = U} \mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$ is not achieved by a single demand distribution, which makes derivation of tight bounds difficult. We therefore utilize another measure of central tendency - the median. The use of the median has two key advantages. Firstly, it allows tight bounds to be developed easily, as will be shown very shortly. Secondly, the median is well known in robust statistics literature to be more resilient to data contamination than the mean, see for instance, Ricardo et al. (2006).

Theorem 7 *For demand with support in* $(\underline{d}, \bar{d})$*, pdf bounded in* $[\underline{h}, \bar{h}]$*, and median* $\underline{m} \leq m(\tilde{d}) \leq \bar{m}$ *the expected fill rate is bounded by the quantities below.*

1. *Lower bound,*

$$\inf_{\substack{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h}) \\ \underline{m} \leq m(\tilde{d}) \leq \bar{m}}} E\left(\min\{x/\tilde{d}, 1\}\right) = \kappa(x, \underline{d}, p_1, \underline{h}) + \kappa(x, p_1, \bar{m}, \bar{h}) + \kappa(x, \bar{m}, p_2, \underline{h}) + \kappa(x, p_2, \bar{d}, \bar{h}), \tag{5.10}$$

$$\text{where} \qquad p_1 = \frac{\bar{h}\bar{m} - \underline{h}\underline{d} - 0.5}{\bar{h} - \underline{h}}, \qquad p_2 = \frac{\bar{h}\bar{d} - \underline{h}\bar{m} - 0.5}{\bar{h} - \underline{h}}.$$

2. *Upper bound,*

$$\sup_{\substack{\tilde{d} \in \mathcal{D}(\underline{d}, \bar{d}, \underline{h}, \bar{h}) \\ \underline{m} \leq m(\tilde{d}) \leq \bar{m}}} E\left(\min\{x/\tilde{d}, 1\}\right) = \kappa(x, \underline{d}, P_1, \bar{h}) + \kappa(x, P_1, \underline{m}, \underline{h}) + \kappa(x, \underline{m}, P_2, \bar{h}) + \kappa(x, P_2, \bar{d}, \underline{h}), \tag{5.11}$$

$$\text{where} \qquad P_1 = \frac{0.5 - \underline{h}\underline{m} + \bar{h}\underline{d}}{\bar{h} - \underline{h}}, \qquad P_2 = \frac{0.5 + \bar{h}\underline{m} - \underline{h}\bar{d}}{\bar{h} - \underline{h}}.$$

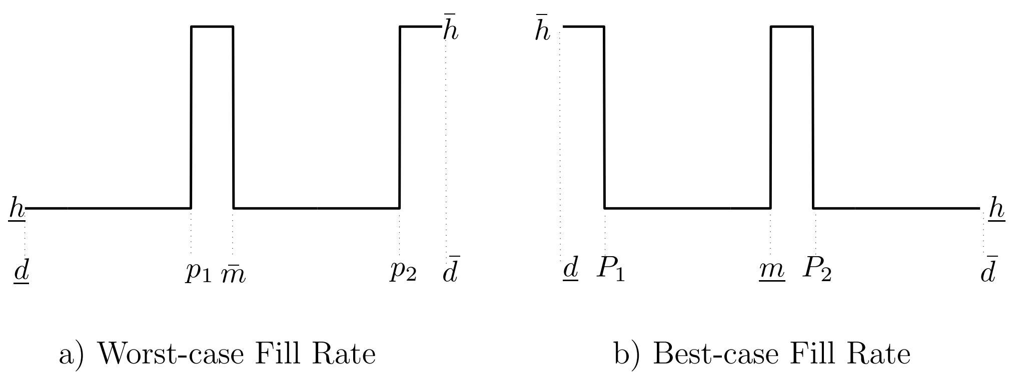


Figure 5.6: Fill rate cases with median constraint

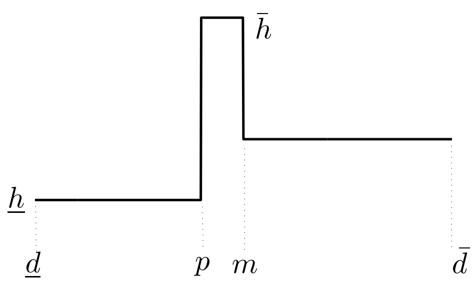
Proof :

1. The worst-case fill rate is achieved by the distribution with median $= \bar{m}$ and the maximum mass packed to the right. See Figure (5.6a). Observe that it consists of four piecewise uniform distributions. Let p_1, p_2 be intersection points as indicated in the figure. Since $\int_{\underline{d}}^{\bar{m}} h(t)\mathrm{d}t = \int_{\bar{m}}^{\bar{d}} h(t)\mathrm{d}t = 0.5$, from the geometry we obtain $p_1 = \frac{\bar{h}\bar{m}-\underline{hd}-0.5}{\bar{h}-\underline{h}}$, $p_2 = \frac{\bar{h}\bar{d}-\underline{h}\bar{m}-0.5}{\bar{h}-\underline{h}}$ and the result follows by Equation (5.7).

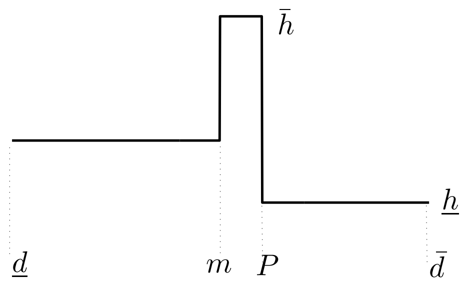
2. The best-case fill rate is achieved by a distribution with median $= \underline{m}$ and maximum mass packed to the left. From the geometry, $P_1 = \frac{0.5-\underline{hm}+\bar{h}\underline{d}}{\bar{h}-\underline{h}}$, $P_2 = \frac{0.5+\bar{h}\underline{m}-\underline{h}\bar{d}}{\bar{h}-\underline{h}}$, giving the result. ■

5.2.3 Bounds using Modal Information

For the case of unimodal distributions, when the pdf increases to some peak and then decreases, the price of distributional uncertainty can be reduced using modal information. The idea is to restrict the family of distributions to unimodal ones.



a) Worst-case Fill Rate



b) Best-case Fill Rate

Figure 5.7: Unimodal demand

Theorem 8 *Let $\mathcal{G}$ be the family of unimodal demand distributions with support in $(\underline{d}, \bar{d})$, pdf bounded in $[\underline{h}, \bar{h}]$ with median m more than the peak. A lower bound on the expected fill rate for $\tilde{d} \in \mathcal{G}$ is*

$$\inf_{\tilde{d}\in\mathcal{G}} E\left(\min\{x/\tilde{d}, 1\}\right) = \kappa(x, \underline{d}, p, \underline{h}) + \kappa(x, p, m, \bar{h}) + \kappa\left(x, m, \bar{d}, \frac{0.5}{\bar{d}-m}\right), \quad \textit{where} \quad p = \frac{\bar{h}m - \underline{hd} - 0.5}{\bar{h} - \underline{h}}. \tag{5.12}$$

Proof : Observe that the pdf to the right of the peak is non-increasing. For demand $> m$, we need to pack half of the entire mass as much as possible to the right such that the pdf is non-increasing, which is achieved by spreading the mass out evenly. That is, $h(t) = \frac{0.5}{\bar{d}-m}, t > m$. The worst-case fill rate hence takes the form of Figure (5.7a). Similar to Theorem 7, $p = \frac{\bar{h}\bar{m}-\underline{h}\underline{d}-0.5}{\bar{h}-\underline{h}}$ by the geometry and the result follows. ■

When the peak of the pdf falls to the right of the median, we can modify Theorem 8 to use the q-percentile rather than the median, where q is selected such that the percentile lies to the right of the peak. The percentile can be estimated using empirical data, though the statistical robustness of the estimate may not be as favorable as the median. Clearly, for any two percentiles that lie to the right of the peak, the smaller percentile will yield a better result. Notice that with $q = 100\%$, the result is identical to the lower bound of Theorem 7. Theorem 8 can hence be considered as a refinement for the case of an unimodal distribution. Using similar concepts, we obtain the following counterpart for the upper bound.

Theorem 9 *Let $\mathcal{H}$ be the the family of unimodal demand distributions with support in $(\underline{d}, \bar{d})$, pdf bounded in $[\underline{h}, \bar{h}]$ with median m less than the peak. An upper bound on the expected fill rate for $\tilde{d} \in \mathcal{H}$ is*

$$\sup_{\tilde{d}\in\mathcal{H}} E\left(\min\{x/\tilde{d}, 1\}\right) = \kappa\left(x, \underline{d}, m, \frac{0.5}{m-\underline{d}}\right) + \kappa(x, m, P, \bar{h}) + \kappa(x, P, \bar{d}, \underline{h}), \quad \textit{where} \quad P = \frac{0.5 + \bar{h}m - \underline{h}\bar{d}}{\bar{h} - \underline{h}}. \tag{5.13}$$

Proof : For demand $< m$, we need to pack half of the entire the mass as much as possible to the left such that the pdf is non-decreasing, which is achieved by spreading the mass out evenly. That is, $h(t) = \frac{0.5}{\underline{d}-m}, t < m$. The best-case fill rate hence takes the form of Figure (5.7b). Similar to Theorem 7, $P = \frac{0.5+\bar{h}\underline{m}-\underline{h}\bar{d}}{\bar{h}-\underline{h}}$ by the geometry and the result follows. ■

With slightly more information, we can reduce the price of distributional uncertainty further. If $\underline{q}, \bar{q}$ are percentiles such that $\underline{q}$-percentile $<$ peak $<$ $\bar{q}$-percentile, modifying Theorem 8 with the $\bar{q}$-percentile and Theorem 9 with the $\underline{q}$-percentile gives the lower and upper bounds, respectively. Our framework can be further generalized. Suppose the distributions are multimodal, the above result is valid when $\underline{q}$-percentile $<$ left-most peak $<$ right-most peak $<$ $\bar{q}$-percentile.

We have completed the theoretical discussions on the models. In the next chapter, we will discuss via computational studies, the effectiveness of our robust approach.

Chapter 6

Computational Studies of Fill Rate Model

In our computational study, we first tested whether our approach has the ability of obtaining meaningful solutions even in the absence of complete demand information. We compared the bounds of Theorem 6 with Theorem 7 to examine the effect of incorporating median information, and to provide insights on the bounds. This is followed by a second set of tests to investigate the effectiveness of our robust approach against the traditional approach of enforcing fill rate constraint using data samples. Here, the parameters are not known but estimated from data. Thereafter, we provided an example constructed using real-life demand to illustrate the model. The technical tests cover bounds at the single-product level. In the example with real-life demand, bounds at the multiproduct level are constructed. Throughout this chapter, we use price of distributional uncertainty to refer to the following

$$\sup_{\tilde{d}\in\mathcal{D}} \mathrm{E}\left(\min\{x/\tilde{d},1\}\right) - \inf_{\tilde{d}\in\mathcal{D}} \mathrm{E}\left(\min\{x/\tilde{d},1\}\right),$$

which is the price incurred for incorporating robustness when the family of distribution is $\mathcal{D}$. In short, it refers to the price of robustness.

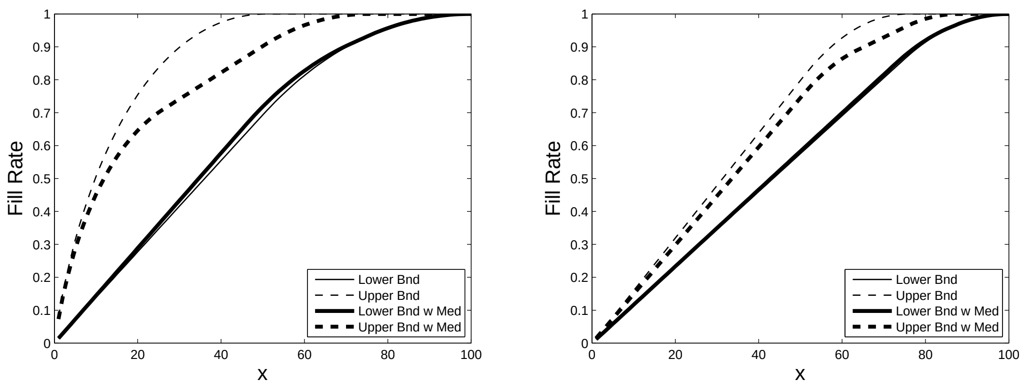
6.1 Technical Tests

For the purposes of technical testing, we used two reference distributions with pdf $h_1(t)$ and $h_2(t), t \in (0,1)$ to construct the descriptive statistics, which means that the family of distributions needs to contain $h_1(t)$ and $h_2(t)$. The test cases are as follows.

1. $h_1(t) = 1$, $h_2(t) = 2t$, scaled to $[\underline{d}, \bar{d}] = [1, 100]$ and $[51, 100]$.

2. Left-skewed distributions, as illustrated in Figure (5.4a), $h_1(t) = 2t$, $h_2(t) = 3t^2$, scaled to $[\underline{d}, \bar{d}] = [1, 100]$ and $[51, 100]$.

3. Right-skewed distributions, as illustrated in Figure (5.4b), $h_1(t) = 2(1 - t)$, $h_2(t) = 3(1 - t)^2$, scaled to $[\underline{d}, \bar{d}] = [1, 100]$ and $[51, 100]$.



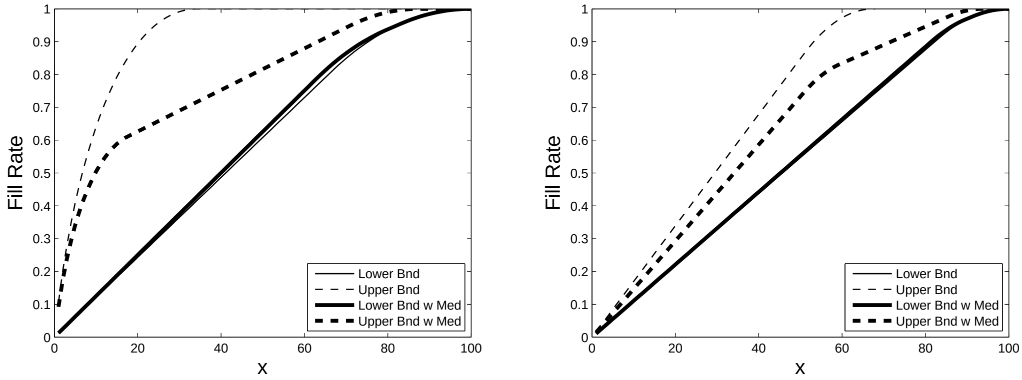
a) $50.5 \le m(\tilde{d}) \le 71.0$ $(\underline{h}, \bar{h}) = [0, 0.02], [\underline{d}, \bar{d}] = [1, 100]$

b) $75.5 \le m(\tilde{d}) \le 85.6$ $(\underline{h}, \bar{h}) = [0, 0.04], [\underline{d}, \bar{d}] = [51, 100]$

Figure 6.1: Case 1 - $h_1(t) = 1$, $h_2(t) = 2t$

The results are presented in Figures (6.1), (6.2) and (6.3), respectively. Note that the y-axis represents fill rate while the x-axis represents order quantity. The results can be taken to be representative for more general demand scaling, as discussed below.

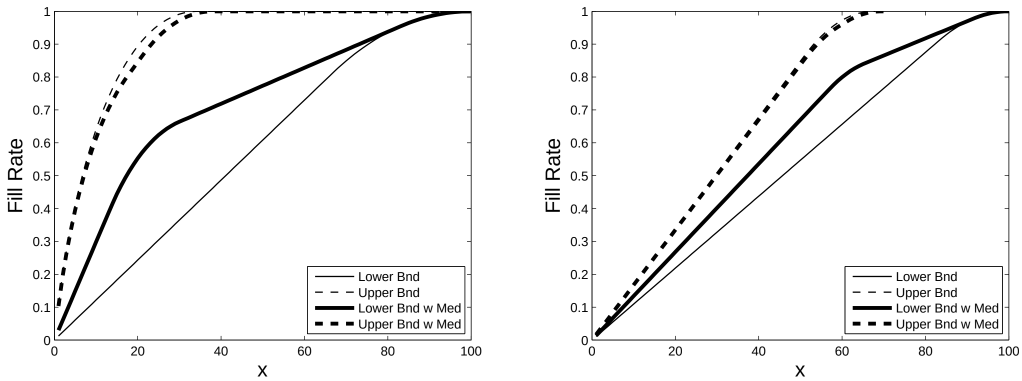
- Recall that fill rate is scale-invariant in the sense that multiplying the demand by any positive constant can be compensated by multiplying the order quantity by the same constant. That is, for any $\beta > 0$, $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right) = \mathrm{E}\left(\min\{\beta x/\beta\tilde{d}, 1\}\right)$. For instance, demand scaled to $[1, 50]$ is approximately half the demand scaled to $[1, 100]$, so the trend for $[1, 100]$ will apply to $[1, 50]$ approximately.



a) $71.0 \leq m(\tilde{d}) \leq 79.6$ $[\underline{h}, \bar{h}] = [0, 0.03], [\underline{d}, \bar{d}] = [1, 100]$

b) $85.6 \leq m(\tilde{d}) \leq 89.9$ $[\underline{h}, \bar{h}] = [0, 0.06], [\underline{d}, \bar{d}] = [51, 100]$

Figure 6.2: Case 2 - $h_1(t) = 2t$, $h_2(t) = 3t^2$



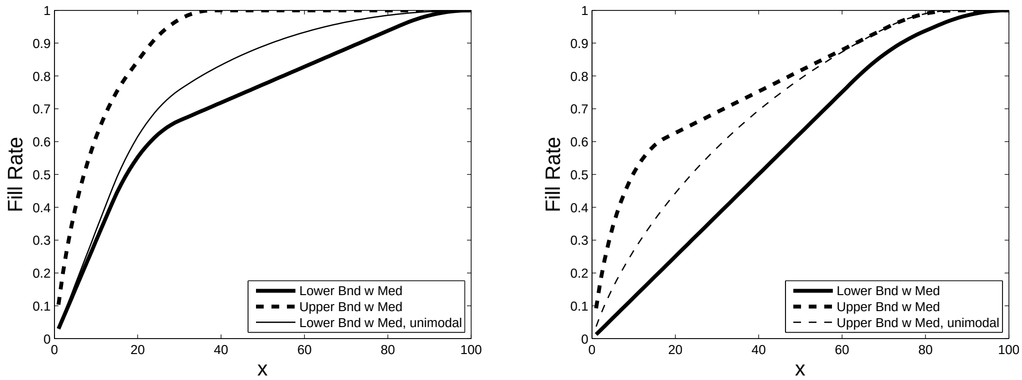
a) $21.4 \leq m(\tilde{d}) \leq 30.0$ $[\underline{h}, \bar{h}] = [0, 0.03], [\underline{d}, \bar{d}] = [1, 100]$

b) $61.1 \leq m(\tilde{d}) \leq 65.4$ $[\underline{h}, \bar{h}] = [0, 0.06], [\underline{d}, \bar{d}] = [51, 100]$

Figure 6.3: Case 3 - $h_1(t) = 2(1-t)$, $h_2(t) = 3(1-t)^2$

- For the same upper support, fill rate decreases when we increase the lower support. This is evident from the figures, when we compare fill rates *without median information* for $[1, 100]$ with $[51, 100]$.

In all the test cases, incorporating median information results in stronger bounds. The price of incorporating distributional uncertainty varies with the order quantity. For left-skewed distributions, Figure (6.2), the prices are much smaller at expected fill rate of 0.9 or more, where the order quantities are high. For right-skewed distributions, Figure (6.3), the bounds of Theorem 7 are comparatively much weaker. However, using Theorem 8 to exploit the fact that the distributions are unimodal results in an improved lower bound, as shown in Figure (6.4a). Similarly, using Theorem 9, we can improve the upper bound of left-skewed distributions, as shown in Figure (6.4b). Theoretically, the price of distributional uncertainty can still be reduced by using a more "optimal" percentile instead of the median, but this was not pursued in this set of experiments.



a) $h_1(t) = 2(1-t)$, $h_2(t) = 3(1-t)^2$

$21.4 \leq m(\tilde{d}) \leq 30.0$

$[\underline{h}, \bar{h}] = [0, 0.03], [\underline{d}, \bar{d}] = [1, 100]$

b) $h_1(t) = 2t$, $h_2(t) = 3t^2$

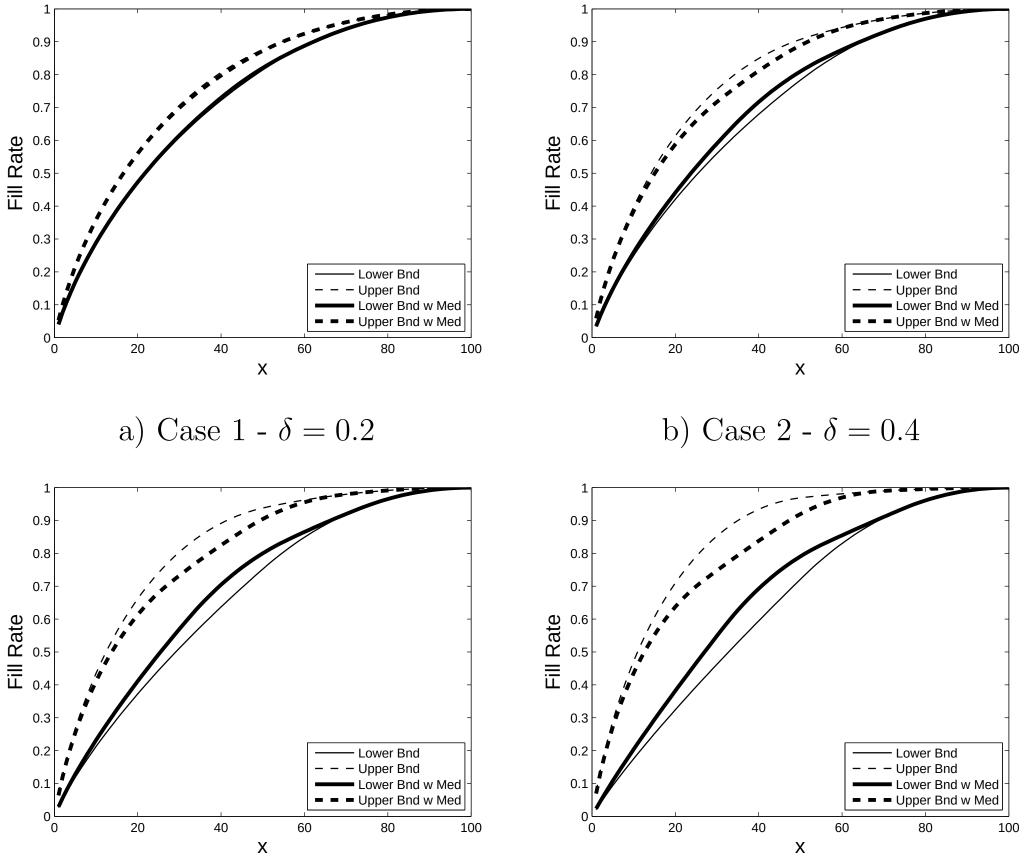
$71.0 \leq m(\tilde{d}) \leq 79.6$

$[\underline{h}, \bar{h}] = [0, 0.03], [\underline{d}, \bar{d}] = [1, 100]$

Figure 6.4: Results with unimodal Information

In another set of experiments, we tested the approach of adding a buffer to the pdf ranges with some factor δ with respect to the uniform distribution. See Equation (5.4). The objective of this set of experiments is to investigate the suitability of our model in safeguarding against deviations from evenly distributed demands. We used $\delta = 0.2, 0.4, 0.6, 0.8$, $[\underline{d}, \bar{d}] = [1, 100]$ with $\pm 10\%$ margin added to the median, $45.5 \leq m(\tilde{d}) \leq 55.6$. The results are presented in Figure (6.5).

As is evident from Figure (6.5), the price of safeguarding the uniform distribution is inexpensive for $\delta \leq 0.6$. For small δ, 0.2 to 0.4, the effect of median information is not significant. At an order quantity of 60, the lower bounds with median information for the four cases are close: $0.89, 0.87, 0.85, 0.83$ for $\delta = 0.2, 0.4, 0.6, 0.8$, respectively. It should be highlighted that Corsten and Gruen (2004) shows that average fill rates for U.S. supermarkets are around 0.9. For skewed distributions, Figure (6.4) shows with the use of modal information, the price to pay for distributional uncertainty is not overly high for this level of fill rate.



a) Case 1 - $\delta = 0.2$

b) Case 2 - $\delta = 0.4$

c) Case 3 - $\delta = 0.6$

d) Case 4 - $\delta = 0.8$

Figure 6.5: Fill rate with buffer to the pdf ranges

6.2 Effectiveness of Robust Fill Rate Model

In this set of tests, we investigated the effectiveness of our robust approach against the traditional approach of enforcing fill rate constraint using data samples. While we have established the lower bound on the expected fill rate function under the assumption that the true pdf and the true median are bounded in $(\underline{h}, \bar{h})$ and $(\underline{m}, \bar{m})$, respectively, we are interested to examine whether our model is still robust when these quantities are not known but estimated from data.

In the traditional approach, the order quantity is derived using available data samples of demands against a fill rate target τ. We denote the sample size used by n, the fill rate target by τ, and we use subscript indexing $x_{T,\,n,\,\tau}$ to denote the order quantity obtained by this approach. That is,

$$x_{T,\,n,\,\tau} = \min_{x} \left\{ x \mid \mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right) \geq \tau, \text{ where } \tilde{d} = d_1 \dots d_n \right\}.$$

For the robust approach, we use $x_{R,\,n,\,\tau}$ to denote the order quantity obtained using the lower bound with median information of Theorem 7. That is,

$$x_{R,\,n,\,\tau} = \min_x \left\{ x \;\middle|\; \begin{array}{r} \inf_{\tilde{d}\in\mathcal{D}} \mathrm{E}\left(\min\{x/\tilde{d},1\}\right) = \kappa(x,\underline{d},p_1,\underline{h}) \quad + \\ \kappa(x,p_1,\bar{m},\bar{h}) + \kappa(x,\bar{m},p_2,\underline{h}) + \kappa(x,p_2,\bar{d},\bar{h}) \geq \tau \end{array} \right\},$$

where $p_1 = \frac{\bar{h}\bar{m}-\underline{h}\underline{d}-0.5}{\bar{h}-\underline{h}}, p_2 = \frac{\bar{h}\bar{d}-\underline{h}\bar{m}-0.5}{\bar{h}-\underline{h}}$, and $\underline{d},\bar{d},\underline{h},\bar{h}$ are quantities estimated from the same set of sample data $\tilde{d} = d_1 \ldots d_n$ used in the traditional approach. It should be emphasized that we let $\bar{m}$, the upper bound on the true median, to equal to the sample median, rather than any valid upper bound. The quantities $\underline{h}$ and $\bar{h}$ were estimated using a histogram. We then evaluated the expected fill rate achieved by $x_{T,\,n,\,\tau}$ and $x_{R,\,n,\,\tau}$ under the true distribution, which we denote by $f_{T,\,n,\,\tau}$ and $f_{R,\,n,\,\tau}$, respectively. That is, $f_{T,\,n,\,\tau} = \mathrm{E}\left(\min\{x_{T,\,n,\,\tau}/\tilde{d},1\}\right)$, where $\tilde{d}$ is the true distribution, and the same applies to $f_{R,\,n,\,\tau}$. The experiment is repeated $N = 1,000$ times, and we use superscript to denote the ith outcome. For comparison purposes, we made use of the following quantities,

- $\mathrm{E}\left((\tau - \tilde{f}_{T,\,n,\,\tau})^+\right) = \dfrac{1}{N}\sum_i (\tau - f^i_{T,\,n,\,\tau})^+$, which measures the level of *under-achievement*, and
- $\mathrm{E}\left((\tilde{f}_{T,\,n,\,\tau} - \tau)^+\right) = \dfrac{1}{N}\sum_i (f^i_{T,\,n,\,\tau} - \tau)^+$, which measures the level of *over-achievement.*

We now discuss why these quantities were used for comparison purposes. Assuming that we know the true distribution or have available a very large number of sample demands, we can use the samples to derive $x_{R,\,n,\,\tau}$. If we repeat the experiments and compute the above quantities, it is evident that $\mathrm{E}\left((\tau - \tilde{f}_{T,\,n,\,\tau})^+\right) = \mathrm{E}\left((\tilde{f}_{T,\,n,\,\tau} - \tau)^+\right) = 0$. That is, the order quantity achieves the fill rate target *exactly*, with no underachievement and no overachievement. This would be ideal, but deriving it in practice is hard, because data samples are limited. Under practical scenarios, service-oriented industries would aim to reduce the risk of underachievement by provisioning for more stocks to safeguard against distributional uncertainty. A model with less underachievement would therefore be more robust and offers better safeguard. To compare the models, it is assumed that we have knowledge of the true distribution which we took to be Beta distributions scaled to $(1, 100)$. We used sample sizes of $n = 20$ and $n = 60$, and fill rate targets τ of 0.85 and 0.95. In the estimation of $\underline{h}$ and $\bar{h}$, we used histograms with 5 bins and 15 bins, for $n = 20$ and $n = 60$, respectively.

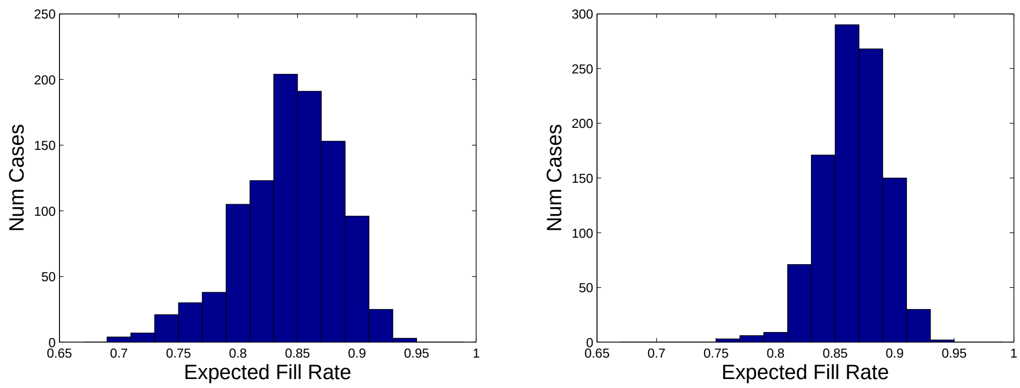
The results are presented in Table (6.1). In Theorem 7, we have established the lower bound on the expected fill rate function under the assumption that the true pdf and the true median are bounded in $(\underline{h}, \bar{h})$ and $(\underline{m}, \bar{m})$, respectively. As is evident from the results, when these quantities are not known but estimated from data, the lower bound is effective. In all cases but one, the robust model has lower underachievement and is able to safeguard fill rate better than the traditional method. For the case, $\beta(2, 1), n = 20$, and $\tau = 0.95$, the underachievement of the traditional approach is marginally better, 0.0085 versus 0.0090 achieved by the robust model. In this case the robustness of our approach is somewhat compromised, because of the shape of the distribution and the error in estimating the parameters. However, both models perform similarly for the case when we let the upper bound on the median, $\bar{m}$, to be 1.1 times the sample median rather than 1.0 times. (In practice, one can add a buffer to the estimated parameters or construct uncertainty sets around them.) The robust model generally has higher overachievement, which is as expected.

To visualize the result, we provided histograms of expected fill rate achievement for the case of $n = 20, \tau = 0.85$, $\beta(1, 1)$ in Figure (6.6). In the figure, the left tail of the robust model, $\tilde{f}_R$, is much shorter than that of the traditional model, $\tilde{f}_T$. We observe that the robust fill rate, $\tilde{f}_R$, is squeezed to the right, with more cases achieving 0.85 and higher fill rate.

It should be highlighted that the traditional approach provisions for order quantity against a single distribution, and hence $\mathrm{E}\left((\tau - \tilde{f}_{T,\, n,\, \tau})^+\right) = \mathrm{E}\left((\tilde{f}_{T,\, n,\, \tau} - \tau)^+\right)$ $= 0$ when we use large sample sizes to derive the order quantity. When sample size is large, we will have zero underachievement for the robust model, but the overachievement will be positive because we are provisioning for order quantity against a *family* of distribution $\mathcal{D}$. If we evaluate the achievement under the worst-case demand in the family $\mathcal{D}$, the robust model will give zero underachievement as well as zero overachievement when $(\underline{h}, \bar{h})$ bounds the pdf tightly. In statistical literature, it is well known that the histogram converges to the true pdf when large sample

Table 6.1: Robust fill rate model versus traditional approach

	Underachievement		Overachievement	
Distribution	Traditional	Robust	Traditional	Robust
		$n = 20, \tau = 0.85$		
$\beta(1,1)$	0.0207	0.0047	0.0131	0.0208
$\beta(2,2)$	0.0179	0.0039	0.0134	0.0315
$\beta(1,2)$	0.0233	5.91e-04	0.0147	0.0713
$\beta(2,1)$	0.0151	0.0126	0.0106	0.0113
		$n = 20, \tau = 0.95$		
$\beta(1,1)$	0.0132	0.0056	0.0057	0.0077
$\beta(2,2)$	0.0113	0.0041	0.0062	0.0152
$\beta(1,2)$	0.0141	0.0032	0.0063	0.0238
$\beta(2,1)$	0.0085	0.0090	0.0051	0.0024
		$n = 60, \tau = 0.85$		
$\beta(1,1)$	0.0104	1.08e-05	0.0083	0.0433
$\beta(2,2)$	0.0098	0	0.0075	0.0671
$\beta(1,2)$	0.0118	0	0.0095	0.1061
$\beta(2,1)$	0.0045	4.33e-04	0.0032	0.0095
		$n = 60, \tau = 0.95$		
$\beta(1,1)$	0.0055	4.69e-06	0.0039	0.0197
$\beta(2,2)$	0.0057	2.45e-06	0.0037	0.0305
$\beta(1,2)$	0.0067	1.17e-06	0.0047	0.0389
$\beta(2,1)$	0.0046	5.21e-04	0.0031	0.0092



a) $\tilde{f}_{T,\,20,\,0.85}$ (Traditional Model) b) $\tilde{f}_{R,\,20,\,0.85}$ (Robust Model)

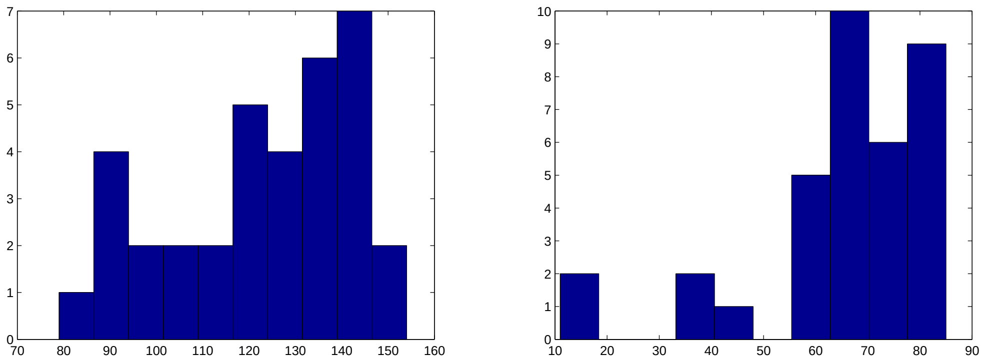
Figure 6.6: Histogram of Expected Fill Rate Achievements

size is large. So tight bounds are possible, but only in theory.

6.3 An Example with Real-life Demand

Following on from the technical testing, experiments using industrial demands were carried out. It should be highlighted that the purpose here is to derive the fill rate trend using real-life demands, and actual application of the model is likely to take a different form. Motivated by the popularity of demand satisfaction problem in the power generation industry, see for instance Williams (1999) the power generation, hydro power and refinery optimization examples, we used demands from the electricity industry to construct an example.

Consider the scenario of an electricity service provider serving two locations, location 1 and location 2 with demands $\tilde{d}_1$ and $\tilde{d}_2$, respectively. For presentation purposes we have reported the demands in units of 10000 MWh. That is, a figure of 50 means 50×10000 MWh. From a supply chain perspective, we can define a service level for each retailer locally and also globally as an aggregate. Denote x_1 and x_2, respectively, as the resources (say coal for instance) allocated to fulfilling the demand from the locations. The resources are measured in quantities of the demand. For instance, x_1 represents the quantity of coal to produce 10000 MWh of electricity for location 1. The problem is to decide x_1, x_2 to maximize the aggregate fill rate given by $\tau_0 = \lambda_1 \mathrm{E}\left(\min\{x_1/\tilde{d}_1, 1\}\right) + \lambda_2 \mathrm{E}\left(\min\{x_2/\tilde{d}_2, 1\}\right)$, subject to cost constraint $c_1x_1 + c_2x_2 \leq C$. The rationale of the objective function lies in the fact



a) Location 1 (Barry)   b) Location 2 (Gorgas)

Figure 6.7: Monthly electricity generation (histogram)

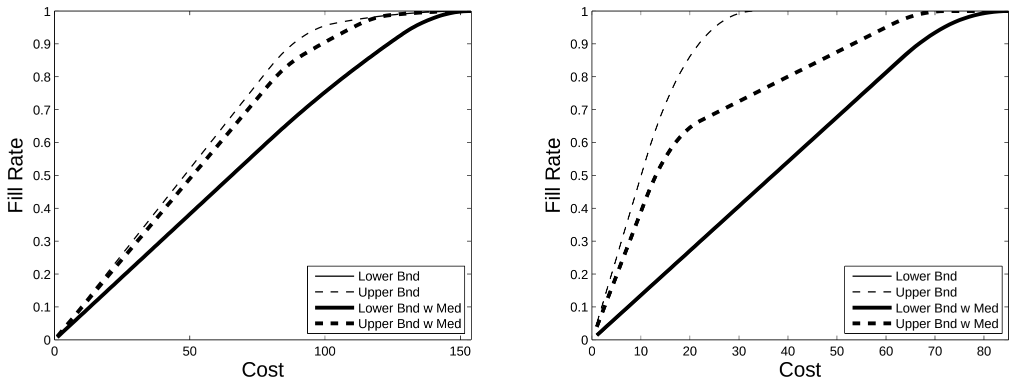
that the service provider is obliged to serve not one but both locations. The weight λ can be derived from the population size of the location, where a larger population implies more importance. Arguably, demand is proportional to population size so as a simplification, we may use the demand mean to derive λ. Observe that the problem has a concave objective function with a single (linear) constraint, which corresponds to a knapsack problem. It is more illustrative to derive an aggregate fill rate versus cost curve. To this end, the problem is optimized using marginal analysis, a popular algorithm for solving the knapsack problem. For simplicity, we have assumed that the resources are of the same type, which means that we can use $c_1 = c_2 = 1$. For location 1, we used the data from 2005 to 2007 from the facility called Barry; for location 2, data from the facility called Gorgas was used. See the Appendix G. Figure (6.7) shows the histogram of the demands, and the descriptive statistics are presented in Table 6.2.

The results are shown in Figure (6.8) and Figure (6.9), with the x-axis representing the cost. We did not use the exact sample median but have added a margin of about $\pm 10\%$ when constructing the bounds. For location 1, the bounds are reasonably close without median information whereas they are much wider for location 2, which means that the price of distributional uncertainty is expensive. Therefore, the median bounds were used during the marginal analysis process (the optimization stage). The aggregate fill rate versus cost curve is shown in Figure (6.9).

In practice, for more complicated problems we can exploit the fact that the

Table 6.2: Monthly electricity generation (descriptive statistics)

	Location 1	Location 2
demand mean	123	65
demand median	125	68
demand variance	396	313
$\underline{d}$	79	11
$\bar{d}$	154	85
$\underline{h}$	0.004	0
$\bar{h}$	0.033	0.045



a) Location 1, $112.5 \leq m(\tilde{d}) \leq 137.5$ b) Location 2, $61.2 \leq m(\tilde{d}) \leq 74.8$

Figure 6.8: Fill rate of electricity demand

expected fill rate function $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$ is concave in x and adopt a linear programming formulation. Recall that the minimum of affine functions $f(x) = \min\{ax + b, \alpha x + \beta\}$ is concave in x, which means that the function $g(x) = \min\{x/d, 1\}$ is concave. Also note that the expectation operator preserves convexity (concavity). Given discrete demands $d_1 \dots d_N$, we have $\mathrm{E}\left(\min\{x/\tilde{d}, 1)\right) = \sum_{i=1}^{N} \Pr(\tilde{d} = d_i) \cdot \min\{x/d_i, 1\}$, which is a weighted sum of the concave function g. This implies that $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right)$ is concave in x. For continuous demands, this argument holds with integration replacing summation, that is, $\mathrm{E}\left(\min\{x/\tilde{d}, 1\}\right) = \int h(t) \min\{x/\tilde{d}, 1\} \, \mathrm{d}t$, see Chapter 3 of Boyd and Vandenberghe (2004).

The functions of Theorems 6 to 9 are concave in x, since they are exact identities for expected fill rate. Therefore it is always possible to approximate Problem

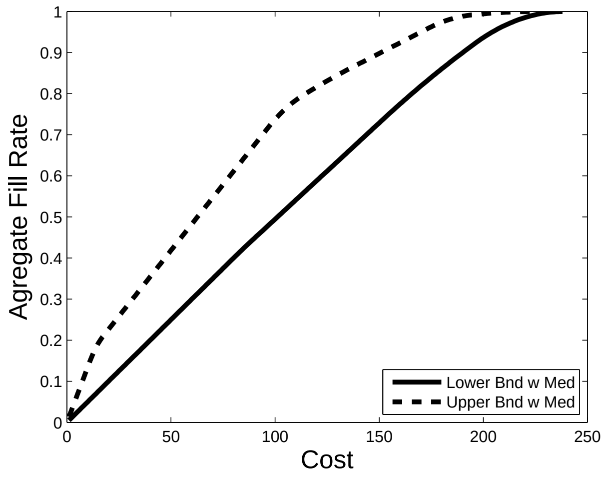


Figure 6.9: Aggregate fill rate of electricity demand

(5.5) using a concave piecewise-linear function, which results in a linear programming formulation for the multiproduct problem. We can solve a piecewise-linear approximation problem that underestimates

$$\inf_{\substack{\tilde{d}\in\mathcal{D}(\underline{d},\bar{d},\underline{h},\bar{h}) \\ \underline{m}\leq m(\tilde{d})\leq\bar{m}}} \mathrm{E}\left(\min\{x/\tilde{d},1\}\right)$$

and a piecewise-linear approximation problem that overestimates

$$\sup_{\substack{\tilde{d}\in\mathcal{D}(\underline{d},\bar{d},\underline{h},\bar{h}) \\ \underline{m}\leq m(\tilde{d})\leq\bar{m}}} \mathrm{E}\left(\min\{x/\tilde{d},1\}\right).$$

6.4 Summary

This part of the research proposes a new methodology to assure that high fill rate is achieved even with uncertain demand distributions. Using the moment approach to construct bounds for the expected fill rate is impractical, but it is possible to exploit other descriptive statistics and we propose a framework to optimize expected fill rate using information such as the range and median of the demand, and the range of the pdf. Using the median bounds, the price of distributional uncertainty does not appear to be too expensive, especially for high fill rates of 0.9 or more.

In practice, one would construct the bounds for the products individually and incorporate information gradually, incorporating median and then modal information. A final fill rate bound is then selected. The items may be ranked according to the price of distributional uncertainty to identify those requiring exceptionally high price. The inventory manager could then focus his/her effort to improve the demand forecasting of these items. He/she would need to collect more data on these items, so as to reduce the price of distributional uncertainty to a reasonable level. The final step is to solve the multiproduct model using the finalized bounds to obtain the price of distributional uncertainty at the aggregated level.

Chapter 7

Conclusions

We close the thesis with this concluding chapter, which highlights the contributions of our research. Recall that the thesis proposes two methodologies, a technique to optimize multiperiod inventory model robustly and an approach to safeguard fill rate against distributional uncertainty. We begin by highlighting the advantages of our models and the insights gained.

- The approach of the multiperiod inventory model has the advantage of being able to obtain the replenishment policy by solving a tractable polynomial-time solvable SOCP of modest size. The computational studies suggest that the truncated linear replenishment policy performs better than linear and static ones. Moreover, the robustness of the truncated linear replenishment policy is exemplified by its outperformance against optimal policies despite using significantly less information. Although the robust policy does not necessarily have a base-stock structure, our computational studies suggest that it can perform better than simple heuristics derived from dynamic programming.

- In the latter half of the thesis, we propose a new methodology to assure that high fill rate is achieved even with uncertain demand distributions. We have shown that using the moment approach to construct bounds for the expected fill rate is impractical, which explains the scarcity of distribution-free approaches in the fill rate practice. Nevertheless it is possible to exploit other descriptive statistics and we propose a framework to optimize expected fill rate using information such as the range and median of the demand, and the range of the pdf.

7.1 Contributions

7.1.1 Multiperiod Inventory Model

Instead of the "Budget of Uncertainty" demand model, we focus on uncertain demands being robustly characterized by their descriptive statistics. The former requires the size of the uncertainty set to be specified, which as exemplified in Bertsimas and Sim (2006), can be dependent on the types of stochastic optimization problem we are addressing. The "Budget of Uncertainty" approach to uncertainty, although it has its strengths, is less appealing when we compare it vis-à-vis with stochastic demand models. Specifically, the contribution of our approach over the related works of Bertsimas and Thiele (2006) and Ben-Tal et al. (2005) can be summarized as follows.

(a) Our proposed robust optimization approximation is based on a comprehensive factor-based demand model that captures correlations such as the autoregressive nature of demand, the effect of external factors, as well as trends and seasonality, among others. In addition, we cater for distributional ambiguity in the underlying factors by considering a family of distributions characterized by the mean, covariance, support and directional deviations. In contrast, the robust optimization model of Bertsimas and Thiele (2006) is restricted to independent demands with an identical mean and variance, whereas the model of Ben-Tal et al. (2005) is confined to completely distribution-free demand uncertainty.

(b) We propose a new policy called the truncated linear replenishment policy, which gives improved approximation to the multiperiod inventory control problem over static and linear decision rules used in the robust optimization proposals of Bertsimas and Thiele (2006) and Ben-Tal et al. (2005), respectively. We *do not* restrict the policy structure to base-stock. We have developed a new bound on a nested sum of expected positive values of random variables and show that the parameters of the truncated linear replenishment policy can be obtained by solving a tractable deterministic mathematical optimization problem in the form of a SOCP, whose solution time is independent on replenishment lead time, demand variability, and correlations.

(c) We studied the computational performance of the static, linear and truncated linear replenishment policies against the optimum history dependent policy and two dynamic programming based heuristics, namely, the myopic policy and a history-independent base-stock policy. We analyze the impact of the solutions over realistic ranges of planing horizon, cost parameters and demand correlations. In contrast, the computational experiments of Bertsimas and Thiele (2006) are confined to independent demands, while the experiment considered in Ben-Tal et al. (2005) does not benchmark against stochastic demand models. Our computational results suggest that the truncated linear replenishment policy, together with information on the directional deviations, yield reasonably good solutions against the optimum and give the best overall performance among linear and static policies and simple dynamic programming based heuristics.

7.1.2 Fill Rate Model

The second model, which is discussed in Chapter 5, proposes an approach to derive the order quantity that achieves an expected fill rate target using descriptive statistics, such as the demand range, the demand median, and the range of the probability density function, so as to safeguard against distributional uncertainty. Our model differs from many robust optimization models in that we do not make use of moments. We can derive the price of incorporating the uncertainty in distribution through solving a pair of approximate problems, by means of linear programming formulation. We have also extended the model to make use of additional information such as the location of the peak demand, and show that incorporating information on the location of the demand peak with respect to specific percentiles reduces the price of uncertainty. (The α-percentile is the value below which α percent of the demands can be found). Our work focuses primarily on single-period problems. However, by inserting fill rate constraints and inventory balance equations at the end of each period, it is possible to apply the methodology to multiperiod problems.

The model may be used to help practitioners estimate the price of distributional uncertainty and evaluate whether it is worth remedying or mitigating. We feel that while this may not eliminate the need for more precise demand forecasting, it is nevertheless a tool that accesses the need for more elaborate data collection efforts. As exemplified by the uniform distribution, the price of distributional uncertainty may be inexpensive. The inventory manager could then focus his/her effort to improve the demand forecasting of the items that have a high price of distributional uncertainty.

7.2 Concluding Remarks

As mentioned in the introductory chapter, even with the long and fruitful research history, the major issues and problems in the area of inventory optimization has not been fully resolved. It has been the author's motivation and desire that this thesis will serve the dual purpose of helping practitioners better manage inventory as well as motivating further research. Inventory management will surely evolve and so will the techniques and solutions. This research has provided a pleasant learning experience, and it is the author's hope that others may benefit from it as well.

Bibliography

Adida, E., Perakis, G. (2006): A Robust Optimization Approach to Dynamic Pricing and Inventory Control with no Backorders. *Mathematical Programming* 107 (1-2): 97-129.

Agrawal, V., Seshadri S. (2000): Distribution Free Bounds for Service Constrained (Q, r) Inventory Systems. *Naval Research Logistics*, 47, 635-656.

Atamtürk, A., Narayanan, V. (2010): Conic mixed-integer rounding cuts. *Mathematical Programming Series A*, 122 (1), 1-20.

Azoury, K. S. (1985): Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science*, 31 (9) 1150-1160.

Baldenius, T., Reichelstein S. (2005): Incentives for Efficient Inventory Management The Role of Historical Cost. *Management Science*, 51 (7).

Ben-Tal, A., A. Goryashko, E. Guslitzer and A. Nemirovski. (2004): Adjusting robust Solutions of uncertain linear programs. *Mathematical Programming*, 99, 351-376.

Ben-Tal, A., B. Golany, A Nemirovski., J. Vial. (2005): Supplier-Retailer Flexible Commitments Contracts: A Robust Optimization Approach. *Manufacturing & Service Operations Management*, 7 (3), 248-273.

Ben-Tal, A., Nemirovski, A. (1998): Robust convex optimization. *Mathematics of Operations Research*, 23, 769-805.

Ben-Tal, A., Nemirovski, A. (1999): Robust solutions to uncertain programs. *Operations Research Letters*, 25, 1-13.

Ben-Tal, A., Nemirovski, A. (2000): Robust solutions of Linear Programming problems contaminated with uncertain data. *Mathematical Programming*, 88, 411-424.

Ben-Tal, A., Nemirovski, A. (2001): Lectures on modern convex optimization: analysis, algorithms, and engineering applications. *MPR-SIAM Series on Optimization*, SIAM, Philadelphia.

Bertsimas, D., D. Pachamanova, and M. Sim. (2004): Robust Linear Optimization Under General Norms. *Operations Research Letters*, 32, 510-516.

Bertsimas, D., Iancu D., A., Parrilo, P. (2009): Optimality of Affine Policies in Multi-stage Robust Optimization. *forthcoming Mathematics of Operations Research.*

Bertsimas, D., Sim, M. (2003): Robust Discrete Optimization and Network Flows. *Mathematical Programming*, 98, 49-71.

Bertsimas, D., Sim, M. (2004): Price of Robustness. *Operations Research*, 52 (1), 35-53.

Bertsimas, D. and M. Sim. (2006): Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107 (1), 5 - 36.

Bertsimas, D., Thiele, A. (2006): A Robust Optimization Approach to Inventory Theory. *Operations Research*, 5 (1), 150-168.

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

Bienstock, D. and and Ozbay, N. (2008): Computing robust base-stock levels. *Discrete Optimization*, 5 (2), 389-414.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C. (1994): *Time Series Analysis Forecasting and Control*, 3rd Ed. Holden-Day, San Francisco, CA. 110-114.

Boyd, S., Vandenberghe, L., (2004): Convex Optimization. *Cambrige Univeristy Press.*

Chen, X. Sim, M., Sun, P. (2007): A Robust Optimization Perspective of Stochastic Programming. *Operations Research*, 55 (6),1058-1071.

Chen, X., Sim, M., Sun, P, Zhang, J. (2008): A linear-decision based approximation approach to stochastic programming. *Operations Research*, 56 (2), 344-357.

Chen, F. Y., Krass, D. (2001): Inventory models with minimal servical constraints. *European Journal Of Operatioonal Research*, 131, 120-140.

Chen, W. Q., Sim, M. (2009): Goal Driven Optimization. *Operations Research*, 57 (2), 342-357.

Corsten, D., Gruen, T. (2004) Stock-Outs Cause Walkouts. *Harvard Business Review*, 82 (5).

Ding, Q., Kouvelis, P., Milner, J. M. (2006): Dynamic Pricing Through Discounts for Optimizing Multiple-Class Demand Fulfillment. *Operations Research*, 51 (1) 169-183.

Dvoretzky, A., Kiefer, J., Wolfowitz, J. (1956), Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Annals of Mathematical Statistics*, 27 (3), 642-669.

El-Ghaoui, L., Lebret, H. (1997): Robust solutions to least-square problems to uncertain data matrices. *SIAM Journal on Matrix Analysis and Applications*, 18, 1035-1064.

El-Ghaoui, L., Oustry, F., Lebret, H. (1998): Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9, 33-52.

Federgruen, A., P. Zipkin. (1986): An inventory model with limited production capacity and uncertain demands II. The discounted-cost criterion. *Mathematics of Operations Research*, 11, 208-215.

Gallego, G., Moon I., (1993): The A min-max distribution newsboy problem: Review and extensions. *Journal of Operations Research Society*, 44, 825-834.

Gallego, G., Ryan, J., Simchi-Levi, D. (2001): Minimax analysis for finite horizon inventory models. *IIE Transactions*, 33, 861-874.

Glasserman, P., Tayur, S. (1995): Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Science*, 41, 263-282.

Godfrey, G. A., Powell, W. B. (2001): An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with applications to inventory and distribution. *Management Science*, 47, 1101-1112.

Graves, S., Rinnooy A., H., G., Zipkin P., (1993): Handbooks In Operations Research & Management Science, Volume 4, Logistics of Production & Inventory. *North Holland*.

Graves, S., (1999): A Single-Item Inventory Model for a Nonstationary Demand Process. *Manufacturing & Service Operations Management*, 1 (1), 50-61.

Harris, F., (1913): How many parts to make at once. *Factory, The Magazine of Management*, 10, 135-136, 152.

Iida, T., Zipkin, P. (2006): Approximate solutions of a dynamic forecasting inventory model. *Manufacturing Service Operations Management*, 8 (4) 401-425.

Johnson, G. D., Thompson, H. E., (1975): Optimality of myopic inventory policies for certain dependent demand processes. *Management Science*, 21 (11), 1303-1307.

Kasugai, H. and Kasegai T. (1960): Characteristics of dynamic maximin ordering policy. *Journal of Operations Research Society of Japan*, 3, 11-26.

Lee, H. L., Billington, C. (1992): Managing supply chain inventory pitfalls and opportunities. *Sloan Management Review*, 33 (3), 65.

Lee, H. L., Billington, C. (1995): The evolution of supply-chain management models and practice at Hewlett-Packard. *Interfaces*, 25 (5), 4264.

Lovejoy, W. S. (1990): Myopic policies for some inventory models with uncertain demand distributions. *Management Science*, 36 (6), 724-738.

Levi, R., P al, M., Roundy, R. O., Shmoys, D. B. (2007): Approximation Algorithms for Stochastic Inventory Control Models. *Mathematics of Operations Research*, 32 (2), 284-302.

Levi, R., Roundy, R. O., Shmoys, D. B. (2007): Provably Near-Optimal Sampling-Based Policies for Stochastic Inventory Control Models. *Mathematics of Operations Research*, 32 (4), 821-839.

Lu, X. W., Song, J. S., Regan, A. (2006): Inventory Planning with Forecast Updates: Approximate Solutions and Cost Error Bounds. *Operations Research*, 54 (6) 1079-1097.

Massart, P. (1990): The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *The Annals of Probability* 18 (3), 1269-1283.

Miller, B. (1986): Scarfs state reduction method, flexibility and a dependent demand inventory model. *Operations Research*, 34 (1) 83-90.

Natarajan, K., Pachamanova, D. and Sim, M. (2008): Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization. *Management Science*, 54 (3), 573-585.

Nemirovski, A. and Shapiro, A. (2006): Convex approximation of chance constrained programs. *SIAM Journal on Optimization*, 17 (4), 969-996.

Powell, W., Ruszczynski, A., Topaloglu, H. (2004): Learning algorithms for separable approximations of discrete stochastic optimization problems. *Mathematics of Operations Research*. 29 (4), 814-836.

Raymond, R. (1931): Quantity & Economy in Manufacture. *Mcgraw Hill.*

Ricardo, A. M., Douglas R. M., Victor J. Y. (2006): Robust Statistics: Theory and Methods. *Wiley Series in Probability and Statistics.*

Scarf, H. (1958): A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*, Stanford University Press, California.

Scarf, H. (1959): Bayes solutions of the statistical inventory problem. *Annals of Mathematical Statistics*, 30, 490-508.

Scarf, H. (1960): Some remarks on Bayes solution to the inventory problem. *Naval Research Logistics*, 7, 591-596.

Schwarz, L. B., Deuermeyer, B. L., Badinelli, R. D. (1985): Fill-rate optimiztion in a one warehouse n-retailer distribution system. *Management Science*, 31 (4) 488-489.

Seshadri, V. (1964): On random variables which have the same distribution as their reciprocals. *Canadian Mathematical Bulletin*, 9, 819-824.

Song, J. S. (1998): On the order fill rate in a multi-item, base-stock inventory system. *Operations Research*, 46 (6), 831-845.

Sherbrooke, C. C. (1992): Optimal Inventory Modeling of Systems. *Wiley Interscience.*

Song, M., Klabjan and Simchi-Levi (2007): Robust Stochastic Lot-Sizing by Means of Histograms. *Submitted manuscript.*

Song, J., Zipkin, P., (1993): Inventory control in a fluctuating demand environment. *Operations Research*, 41, 351-370.

Taylor, D. H. (1997): Global Cases in Logistics and Supply Chain Management. *Thomson, London, UK.*

Veinott, A. (1965): Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science*, 12, 206-222.

Williams, H. P., (1999): Model building in mathemetical programming, 4th edition. *John Wiley.*

Zipkin, P. (2000): Foundations of inventory management. *McGraw-Hill Higher Education*, New York.

Appendix A

Proof of Proposition 3

Proof : The bound $\mathrm{E}((\tilde{z}-a)^+) \leq \pi(-a,1)$ follows directly from Theorem 1. Since the bound of Proposition 2 is tight, it suffices to show

$$\pi(-a,1) \leq \begin{cases} \frac{1}{2}\left(-a+\sqrt{\sigma^2+a^2}\right) & \text{if } a \geq \frac{\sigma^2-\mu^2}{2\mu} \\ -a\frac{\mu^2}{\mu^2+\sigma^2}+\mu\frac{\sigma^2}{\mu^2+\sigma^2} & \text{if } a < \frac{\sigma^2-\mu^2}{2\mu} \end{cases}$$

With $\underline{z} = \mu$ and $p = q = \bar{z} = \infty$, we first simplify the bound as follows:

$$\begin{aligned}
\pi(y_0, \boldsymbol{y}) = \min \quad & r_1 + r_2 + r_3 \\
\text{s.t.} \quad & y_{10} + t_1\mu \leq r_1 \\
& 0 \leq r_1 \\
& -t_1 = y_{11} \\
& t_1 \geq 0 \\
& h_1\mu \leq r_2 \\
& y_{20} \leq r_2 \\
& h_1 = y_{21} \\
& h_1 \geq 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(A.1)} \\
& \tfrac{1}{2}y_{30} + \tfrac{1}{2}\sqrt{y_{30}^2 + \sigma^2 y_{31}^2} \leq r_3 \\
& y_{10} + y_{20} + y_{30} = -a \\
& y_{11} + y_{21} + y_{31} = 1 \\
= \min \quad & (y_{10} - y_{11}\mu)^+ + \max\{y_{21}\mu, y_{20}\} + \tfrac{1}{2}y_{30} + \tfrac{1}{2}\sqrt{y_{30}^2 + \sigma^2 y_{31}^2} \\
\text{s.t.} \quad & y_{11} \leq 0 \\
& y_{21} \geq 0 \\
& y_{10} + y_{20} + y_{30} = -a \\
& y_{11} + y_{21} + y_{31} = 1.
\end{aligned}$$

Clearly, with $y_{10} = y_{20} = 0$, $y_{30} = -a$, $y_{11} = y_{21} = 0$ and $y_{31} = 1$, we see that $\pi(y_0, \boldsymbol{y}) \leq -\frac{1}{2}a + \frac{1}{2}\sqrt{a^2 + \sigma^2}$. Now for $a < \frac{\sigma^2 - \mu^2}{2\mu}$, we let $y_{10} = y_{11} = 0$,

$$\begin{aligned}
y_{20} &= \mu\tfrac{\sigma^2 - \mu^2 - 2\mu a}{\mu^2 + \sigma^2}, \\
y_{21} &= \tfrac{\sigma^2 - \mu^2 - 2\mu a}{\mu^2 + \sigma^2} \geq 0, \\
y_{30} &= (\mu + a)\tfrac{\mu^2 - \sigma^2}{\mu^2 + \sigma^2}, \\
y_{31} &= 2\mu\tfrac{\mu + a}{\mu^2 + \sigma^2}.
\end{aligned}$$

which are feasible in Problem (A.1). Hence,

$$\begin{aligned}
\pi(-a, 1) &\leq (y_{10} - y_{11}\mu)^+ + \max\{y_{21}\mu, y_{20}\} + \tfrac{1}{2}y_{30} + \tfrac{1}{2}\sqrt{y_{30}^2 + \sigma^2 y_{31}^2} \\
&= -a - \frac{1}{2}(\mu + a)\frac{\mu^2 - \sigma^2}{\mu^2 + \sigma^2} + \frac{1}{2}\underbrace{\sqrt{(a + \mu)^2}}_{=a+\mu} \\
&= -a\frac{\mu^2}{\mu^2 + \sigma^2} + \mu\frac{\sigma^2}{\mu^2 + \sigma^2}. \blacksquare
\end{aligned}$$

Appendix B

Proof of Theorem 2

Proof : Under the static replenishment policy and using the factor-based demand model, the inventory level at the end of period t is given by

$$\begin{aligned} y_{t+1}^{SRP}(\tilde{\boldsymbol{d}}_t) &= y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{SRP}(\tilde{\boldsymbol{d}}_{\tau-L-1}) - \sum_{\tau=1}^{t} d_\tau(\tilde{\boldsymbol{z}}) \\ &= y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{0*} - \sum_{\tau=1}^{t} d_\tau^0 - \sum_{\tau=1}^{t}\sum_{k=1}^{N} d_\tau^k \tilde{z}_k \\ &= \underbrace{y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{0*} - \sum_{\tau=1}^{t} d_\tau^0}_{=y_{t+1}^{0*}} + \sum_{k=1}^{N} \underbrace{\left(\sum_{\tau=1}^{t}(-d_\tau^k)\right)}_{=y_{t+1}^{k*}} \tilde{z}_k \\ &= y_{t+1}^{0*} + \sum_{\tau=1}^{N} y_{t+1}^{k*}\tilde{z}_k \end{aligned}$$

where y_{t+1}^{k*} $k = 0, \ldots, N$, $t = 1, \ldots, T$ are the optimum solutions of Problem (3.9). Clearly, the static replenishment policy, $x_t^{SRP}(\tilde{\boldsymbol{d}}_{t-1})$ is feasible in Problem (3.3). Moreover, by Theorem 1, we have

$$\begin{aligned} & \mathrm{E}\left(c_t x_t^{SRP}(\tilde{\boldsymbol{d}}_{t-1}) + h_t\left(y_{t+1}^{SRP}(\tilde{\boldsymbol{d}}_t)\right)^+ + b_t\left(y_{t+1}^{SRP}(\tilde{\boldsymbol{d}}_t)\right)^-\right) \\ = \ & \mathrm{E}\left(c_t x_t^{0*} + h_t\left(y_{t+1}^{0*} + \sum_{k=1}^{N} y_{t+1}^{k*}\tilde{z}_k\right)^+ + b_t\left(-y_{t+1}^{0*} - \sum_{k=1}^{N} y_{t+1}^{k*}\tilde{z}_k\right)^+\right) \qquad \text{(B.1)} \\ \leq \ & c_t x_t^{0*} + h_t\pi\left(y_{t+1}^{0*}, \boldsymbol{y}_{t+1}^*\right) + b_t\pi\left(-y_{t+1}^{0*}, -\boldsymbol{y}_{t+1}^*\right). \end{aligned}$$

Hence, $Z_{STOC} \leq Z_{SRP}$. ■

Appendix C

Proof of Theorem 3

Proof : Observe that Problem (3.13) with additional constraints $x_t^k = 0,\ k = 1,\ldots,N,\ t = 1\ldots,T-L$ gives the same feasible constraint set as Problem (3.9). Moreover, the objective functions of both problems are the same. Hence, $Z_{LRP} \leq Z_{SRP}$. Under the linear replenishment policy, the inventory level at the end of period t is given by

$$
\begin{aligned}
& y_{t+1}^{LRP}(\tilde{\boldsymbol{d}}_t) \\
= \;& y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{LRP}(\tilde{\boldsymbol{d}}_{\tau-L-1}) - \sum_{\tau=1}^{t} d_\tau(\tilde{\boldsymbol{z}}) \\
= \;& y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} \left(x_{\tau-L}^{0*} + \sum_{k=1}^{N} x_{\tau-L}^{k*} \tilde{z}_k \right) - \sum_{\tau=1}^{t} d_\tau^0 - \sum_{\tau=1}^{t} \sum_{k=1}^{N} d_\tau^k \tilde{z}_k \\
= \;& \underbrace{y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{0*} - \sum_{\tau=1}^{t} d_\tau^0}_{=y_{t+1}^{0*}} + \sum_{k=1}^{N} \underbrace{\left(\sum_{\tau=1}^{t} (x_{\tau-L}^{k*} - d_\tau^k) \right)}_{=y_{t+1}^{k*}} \tilde{z}_k \\
= \;& y_{t+1}^{0*} + \sum_{\tau=1}^{N} y_{t+1}^{k*} \tilde{z}_k
\end{aligned}
$$

where $y_{t+1}^{k*}\ k = 0,\ldots,N,\ t = 1,\ldots,T$ are the optimum solutions of Problem (3.13). Clearly, the linear replenishment policy, $x_t^{LRP}(\tilde{\boldsymbol{d}}_{t-1})$ is feasible in Problem (3.3). Moreover, by Theorem 1 and by $\tilde{\boldsymbol{z}}$ being zero mean random variables, we have

$$\begin{aligned}
& \mathrm{E}\left(c_t x_t^{LRP}(\tilde{\boldsymbol{d}}_{t-1}) + h_t \left(y_{t+1}^{LRP}(\tilde{\boldsymbol{d}}_t)\right)^+ + b_t \left(y_{t+1}^{LRP}(\tilde{\boldsymbol{d}}_t)\right)^-\right) \\
= \; & \mathrm{E}\left(c_t \left(x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}}\right) + h_t \left(y_{t+1}^{0*} + \sum_{k=1}^{N} y_{t+1}^{k*}\tilde{z}_k\right)^+ + b_t \left(-y_{t+1}^{0*} - \sum_{k=1}^{N} y_{t+1}^{k*}\tilde{z}_k\right)^+\right) \\
\leq \; & c_t x_t^{0*} + h_t \pi\left(y_{t+1}^{0*}, \boldsymbol{y}_{t+1}^{*}\right) + b_t \pi\left(-y_{t+1}^{0*}, -\boldsymbol{y}_{t+1}^{*}\right).
\end{aligned} \tag{C.1}$$

Hence, $Z_{STOC} \leq Z_{LRP}$. ■

Appendix D

Proof of Theorem 4

Proof : We first show the following bound:

$$\left(y+\sum_{i=1}^{p} x_i^+\right)^+ \leq \left(y+\sum_{i=1}^{p} w_i\right)^+ + \sum_{i=1}^{p}\left((-w_i)^+ + (x_i - w_i)^+\right) \tag{D.1}$$

for all w_i, $i = 1, \ldots, p$. Note that for any scalars a, b

$$a^+ + b^+ \geq (a+b)^+ \tag{D.2}$$

$$a^+ + b^+ = a^+ + (b^+)^+ \geq (a+b^+)^+. \tag{D.3}$$

Therefore, we have

$$\begin{aligned}
& \left(y+\sum_{i=1}^{p} w_i\right)^+ + \sum_{i=1}^{p}\left((-w_i)^+ + (x_i - w_i)^+\right) \\
\geq\ & \left(y+\sum_{i=1}^{p}\left(w_i + (-w_i)^+ + (x_i - w_i)^+\right)\right)^+ && \text{from Inequality (D.3)} \\
=\ & \left(y+\sum_{i=1}^{p}\left(w_i^+ + (x_i - w_i)^+\right)\right)^+ \\
\geq\ & \left(y+\sum_{i=1}^{p} x_i^+\right)^+ && \text{from Inequality (D.2).}
\end{aligned}$$

For notational convenience, we denote $y(\tilde{\boldsymbol{z}}) = y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}$, $x_i(\tilde{\boldsymbol{z}}) = x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}$ and $w_i(\tilde{\boldsymbol{z}}) = w_i^0 + \boldsymbol{w}_i{}'\tilde{\boldsymbol{z}}$. To prove Inequality (3.16), it suffices to show that for any $w_i^0, \boldsymbol{w}_i, i = 1, \ldots, p$, we have

$$\begin{aligned}
& \pi\left(y^0 + \sum_{i=1}^{p} w_i^0, \boldsymbol{y} + \sum_{i=1}^{p} \boldsymbol{w}_i\right) + \sum_{i=1}^{p}\left(\pi(-w_i^0, -\boldsymbol{w}_i) + \pi(x_i^0 - w_i^0, \boldsymbol{x}_i - \boldsymbol{w}_i)\right) \\
\geq\ & \mathrm{E}\left(\left(y(\tilde{\boldsymbol{z}}) + \sum_{i=1}^{p} w_i(\tilde{\boldsymbol{z}})\right)^+\right) + \sum_{i=1}^{p}\left(\mathrm{E}\left((-w_i(\tilde{\boldsymbol{z}}))^+\right) + \mathrm{E}\left((x_i(\tilde{\boldsymbol{z}}) - w_i(\tilde{\boldsymbol{z}}))^+\right)\right) \\
\geq\ & \mathrm{E}\left(\left(y(\tilde{\boldsymbol{z}}) + \sum_{i=1}^{p} x_i(\tilde{\boldsymbol{z}})^+\right)^+\right),
\end{aligned}$$

where the first inequality follows from Theorem 1 and the last inequality follows from Inequality (D.1).

To prove the tightness of the bound, we consider the case when $x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}$, $i = 1, \ldots, p$ are nonzero crossing functions with respect to $\boldsymbol{z} \in \mathcal{W}$. Let

$$\mathcal{K} = \{k \ : \ x_k^0 + \boldsymbol{x}_k{}'\boldsymbol{z} \geq 0 \ \forall \boldsymbol{z} \in \mathcal{W}\}.$$

Hence,

$$y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right)^+ = y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i \in \mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right) \qquad \forall \boldsymbol{z} \in \mathcal{W}.$$

Therefore, if

$$y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right)^+ \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

or equivalently,

$$y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i \in \mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right) \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W},$$

we have

$$\begin{aligned} & \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \\ = \ & \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i \in \mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)\right)^+\right) \\ = \ & y^0 + \sum_{i \in \mathcal{K}} x_i^0. \end{aligned}$$

Likewise, if

$$y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right)^+ \leq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

or equivalently,

$$y^0 + \boldsymbol{y}'\boldsymbol{z} + \sum_{i \in \mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\boldsymbol{z}\right) \leq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W},$$

we have

$$\begin{aligned} & \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \textstyle\sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \\ = \ & \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \textstyle\sum_{i \in \mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)\right)^+\right) \\ = \ & 0. \end{aligned}$$

Indeed, for all $k \in \mathcal{K}$, let $(w_i^0, \boldsymbol{w}_i) = (x_i^0, \boldsymbol{x}_i)$ and for all $k \notin \mathcal{K}$, $(w_i^0, \boldsymbol{w}_i) = (0, \boldsymbol{0})$. Therefore, using the tightness result of Theorem 1, we have

$$
\begin{aligned}
& \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \\
\leq\ & \eta((y^0, \boldsymbol{y}), (x_1^0, \boldsymbol{x}_1), \ldots, (x_p^0, \boldsymbol{x}_p)) \\
=\ & \min_{w_i^0, \boldsymbol{w}_i, i=1,\ldots,p} \left\{ \pi\left(y^0 + \sum_{i=1}^{p} w_i^0, \boldsymbol{y} + \sum_{i=1}^{p} \boldsymbol{w}_i\right) + \right. \\
& \left. \sum_{i=1}^{p} \left(\pi(-w_i^0, -\boldsymbol{w}_i) + \pi(x_i^0 - w_i^0, \boldsymbol{x}_i - \boldsymbol{w}_i)\right) \right\} \\
\leq\ & \pi\left(y^0 + \sum_{i\in\mathcal{K}} x_i^0, \boldsymbol{y} + \sum_{i\in\mathcal{K}} \boldsymbol{x}_i\right) + \sum_{i\in\mathcal{K}} \left(\underbrace{\pi(-x_i^0, -\boldsymbol{x}_i)}_{=0} + \pi(0, \boldsymbol{0})\right) + \\
& \quad \sum_{i\notin\mathcal{K}} \left(\pi(-0, -\boldsymbol{0}) + \underbrace{\pi(x_i^0, \boldsymbol{x}_i)}_{=0}\right) \\
=\ & \pi\left(y^0 + \sum_{i\in\mathcal{K}} x_i^0, \boldsymbol{y} + \sum_{i\in\mathcal{K}} \boldsymbol{x}_i\right) \\
=\ & \begin{cases} y^0 + \sum_{i\in\mathcal{K}} x_i^0 & \text{if } y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i\in\mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right) \geq 0\ \forall \boldsymbol{z} \in \mathcal{W} \\ 0 & \text{if } y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i\in\mathcal{K}} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right) \leq 0\ \forall \boldsymbol{z} \in \mathcal{W} \end{cases} \\
=\ & \mathrm{E}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \sum_{i=1}^{p} \left(x_i^0 + \boldsymbol{x}_i{}'\tilde{\boldsymbol{z}}\right)^+\right)^+\right). \blacksquare
\end{aligned}
$$

Appendix E

Proof of Theorem 5

Proof : We first show that $Z_{TLRP} \leq Z_{LRP}$. Let $x_t^{k\dagger}$, $k = 0, \ldots, N$, $t = 1, \ldots, T-L$ and $y_{t+1}^{k\dagger}$, $k = 0, \ldots, N$, $t = 1, \ldots, T$ be the optimum solution to Problem (3.13), which is also feasible in Problem (3.17). Based on the following inequality,

$$\begin{aligned} & \eta((y^0, \boldsymbol{y}), (x_1^0, \boldsymbol{x}_1), \ldots, (x_p^0, \boldsymbol{x}_p)) \\ = & \min_{w_i^0, \boldsymbol{w}_i, i=1,\ldots,p} \left\{ \pi\left(y^0 + \sum_{i=1}^{p} w_i^0, \boldsymbol{y} + \sum_{i=1}^{p} \boldsymbol{w}_i \right) + \right. \\ & \left. \sum_{i=1}^{p} \left(\pi(-w_i^0, -\boldsymbol{w}_i) + \pi(x_i^0 - w_i^0, \boldsymbol{x}_i - \boldsymbol{w}_i) \right) \right\} \\ \leq & \pi\left(y^0, \boldsymbol{y}\right) + \sum_{i=1}^{p} \pi(x_i^0, \boldsymbol{x}_i), \end{aligned} \tag{E.1}$$

we have

$$\begin{aligned} Z_{TLRP} \leq & \sum_{t=1}^{T} c_t \pi(x_t^{0\dagger}, \boldsymbol{x}_t^\dagger) + \sum_{t=1}^{L} \Big(h_t \pi(y_{t+1}^{0\dagger}, \boldsymbol{y}_{t+1}^\dagger) + b_t \pi(-y_{t+1}^{0\dagger}, -\boldsymbol{y}_{t+1}^\dagger) \Big) + \\ & \sum_{t=L+1}^{T} \Big(h_t \eta \left((y_{t+1}^{0\dagger}, \boldsymbol{y}_{t+1}^\dagger), (-x_1^{0\dagger}, -\boldsymbol{x}_1), \ldots, (-x_{t-L}^{0\dagger}, -\boldsymbol{x}_{t-L}^\dagger) \right) + \\ & b_t \eta \left((-y_{t+1}^{0\dagger}, -\boldsymbol{y}_{t+1}^\dagger), (x_1^{0\dagger} - S_t, \boldsymbol{x}_1^\dagger), \ldots, (x_{t-L}^{0\dagger} - S_t, \boldsymbol{x}_{t-L}^\dagger) \right) \Big) \\ \leq & \sum_{t=1}^{T} c_t \pi(x_t^{0\dagger}, \boldsymbol{x}_t^\dagger) + \sum_{t=1}^{L} \Big(h_t \pi(y_{t+1}^{0\dagger}, \boldsymbol{y}_{t+1}^\dagger) + b_t \pi(-y_{t+1}^{0\dagger}, -\boldsymbol{y}_{t+1}^\dagger) \Big) + \\ & \sum_{t=L+1}^{T} \Big(h_t \pi\left(y_{t+1}^{0\dagger}, \boldsymbol{y}^\dagger \right) + h_t \sum_{i=1}^{t-L} \pi(-x_i^{0\dagger}, -\boldsymbol{x}_i^\dagger) \quad + \\ & b_t \pi\left(-y_{t+1}^{0\dagger}, -\boldsymbol{y}_{t+1}^\dagger \right) + b_t \sum_{i=1}^{t-L} \pi(-x_i^{0\dagger} - S_t, \boldsymbol{x}_i^\dagger) \Big). \end{aligned}$$

Observe that since $x_t^{0\dagger} + \boldsymbol{x}_t^{\dagger\prime} \boldsymbol{z} \geq 0$, $-x_t^{0\dagger} - \boldsymbol{x}_t^{\dagger\prime} \boldsymbol{z} \leq 0$ and $x_t^{0\dagger} - S_t + \boldsymbol{x}_t^{\dagger\prime} \boldsymbol{z} \leq 0$ for all $\boldsymbol{z} \in \mathcal{W}$, we have from Theorem 1, $\pi(x_i^{0\dagger}, \boldsymbol{x}_i^\dagger) = x_i^{0\dagger}$,$\pi(-x_i^{0\dagger}, -\boldsymbol{x}_i^\dagger) = 0$ and

$\pi(x_i^{0\dagger} - S_t, \boldsymbol{x}_i^\dagger) = 0$ for all $i = 1, \ldots, T-L$. Hence,

$$Z_{TLRP} \leq \sum_{t=1}^{T} \left(c_t x_t^{0\dagger} + h_t \pi\left(y_{t+1}^{0\dagger}, \boldsymbol{y}^\dagger\right) + b_t \pi\left(-y_{t+1}^{0\dagger}, -\boldsymbol{y}_{t+1}^\dagger\right)\right) = Z_{LRP}.$$

We next show that $Z_{STOC} \leq Z_{TLRP}$. Under the truncated linear replenishment policy, the inventory level at the end of period t is given by

$$y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t) = y_1^0 + \sum_{\tau=1}^{\min\{L,t\}} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{TLRP}(\tilde{\boldsymbol{d}}_{\tau-L-1}) - \sum_{\tau=1}^{t} d_\tau(\tilde{\boldsymbol{z}}).$$

Let x_t^{k*}, $k = 0, \ldots, N$, $t = 1, \ldots, T-L$ and y_{t+1}^{k*}, $k = 0, \ldots, N$, $t = 1, \ldots, T$ be the optimum solution to Problem (3.17). It suffices to show that the following bounds:

(a)

$$\mathrm{E}\left(\left(x_t^{TLRP}(\tilde{\boldsymbol{d}}_{t-1})\right)^+\right) \leq \pi(x_t^{0*}, \boldsymbol{x}_t^*).$$

(b) For $t \leq L$,

$$\mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^+\right) \leq \pi\left(y_{t+1}^{0*}, \boldsymbol{y}_{t+1}^*\right)$$

and

$$\mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^-\right) \leq \pi\left(-y_{t+1}^{0*}, -\boldsymbol{y}_{t+1}^*\right)$$

(c) For $t = L+1, \ldots, T$,

$$\mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^+\right) \leq \eta\left((y_{t+1}^{0*}, \boldsymbol{y}_{t+1}^*), (-x_1^{0*}, -\boldsymbol{x}_1^*), \ldots, (-x_{t-L}^{0*}, -\boldsymbol{x}_{t-L}^*)\right)$$

and

$$\mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^-\right) \leq$$
$$\eta\left((-y_{t+1}^{0*}, -\boldsymbol{y}_{t+1}^*), (x_1^{0*} - S_t, \boldsymbol{x}_1^*), \ldots, (x_{t-L}^{0*} - S_t, \boldsymbol{x}_{t-L}^*)\right).$$

For Bound (a), we note that

$$\begin{aligned}
\mathrm{E}\left(x_t^{TLRP}(\tilde{\boldsymbol{d}}_{t-1})\right) &= \mathrm{E}\left(\min\left\{\max\left\{x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}}, 0\right\}, S_t\right\}\right) \\
&\leq \mathrm{E}\left(\max\left\{x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}}, 0\right\}\right) \\
&= \mathrm{E}\left(\left(x_t^{0*} + \boldsymbol{x}_t^{*\prime}\tilde{\boldsymbol{z}}\right)^+\right) \\
&\leq \pi(x_t^{0*}, \boldsymbol{x}_t^*).
\end{aligned}$$

We focus on deriving Bound (c), as the exposition of Bound (b) is similar. Indeed, using the bound of Theorem 4, we have for $t \geq L+1$,

$$
\begin{aligned}
& \mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^+\right) \\
&= \mathrm{E}\left(\left(y_1^0 + \sum_{\tau=1}^{L} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{TLRP}(\tilde{\boldsymbol{d}}_{\tau-L-1}) - \sum_{\tau=1}^{t} d_\tau(\tilde{\boldsymbol{z}})\right)^+\right) \\
&= \mathrm{E}\Bigg(\Bigg(y_1^0 + \sum_{\tau=1}^{L} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} \min\left\{\max\left\{x_{\tau-L}^{0*} + {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}, 0\right\}, S_t\right\} - \\
&\qquad \sum_{\tau=1}^{t} d_\tau^0 - \sum_{\tau=1}^{t}\sum_{k=1}^{N} d_\tau^k \tilde{z}_k\Bigg)^+\Bigg) \\
&\leq \mathrm{E}\Bigg(\Bigg(y_1^0 + \sum_{\tau=1}^{L} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} \max\left\{x_{\tau-L}^{0*} + {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}, 0\right\} - \\
&\qquad \sum_{\tau=1}^{t} d_\tau^0 - \sum_{k=1}^{N}\sum_{\tau=1}^{t} d_\tau^k \tilde{z}_k\Bigg)^+\Bigg) \\
&= \mathrm{E}\Bigg(\Bigg(y_1^0 + \sum_{\tau=1}^{L} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} \left(x_{\tau-L}^{0*} + {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}\right) + \\
&\qquad \sum_{\tau=L+1}^{t} \max\left\{-x_{\tau-L}^{0*} - {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}, 0\right\} - \sum_{\tau=1}^{t} d_\tau^0 - \sum_{k=1}^{N}\sum_{\tau=1}^{t} d_\tau^k \tilde{z}_k\Bigg)^+\Bigg) \\
&= \mathrm{E}\Bigg(\Bigg(\underbrace{y_1^0 + \sum_{\tau=1}^{L} x_{\tau-L}^0 + \sum_{\tau=L+1}^{t} x_{\tau-L}^{0*} - \sum_{\tau=1}^{t} d_\tau^0}_{=y_{t+1}^{0*}} + \\
&\qquad \sum_{\tau=L+1}^{t} \left(-x_{\tau-L}^{0*} - {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}\right)^+ + \sum_{k=1}^{N} \underbrace{\left(\sum_{\tau=1}^{t}(x_{\tau-L}^{k*} - d_\tau^k)\right)}_{=y_{t+1}^{k*}} \tilde{z}_k\Bigg)^+\Bigg) \\
&= \mathrm{E}\left(\left(y_{t+1}^{0*} + {\boldsymbol{y}_{t+1}^*}'\tilde{\boldsymbol{z}} + \sum_{\tau=L+1}^{t} \left(-x_{\tau-L}^{0*} - {\boldsymbol{x}_{\tau-L}^*}'\tilde{\boldsymbol{z}}\right)^+\right)^+\right) \\
&\leq \eta\left((y_{t+1}^{0*}, \boldsymbol{y}_{t+1}^*), (-x_1^{0*}, -\boldsymbol{x}_1^*), \ldots, (-x_{t-L}^{0*}, -\boldsymbol{x}_{t-L}^*)\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\mathrm{E}\left(\left(y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^{-}\right)\\
&=\mathrm{E}\left(\left(-y_{t+1}^{TLRP}(\tilde{\boldsymbol{d}}_t)\right)^{+}\right)\\
&=\mathrm{E}\left(\left(-y_1^0-\sum_{\tau=1}^{L}x_{\tau-L}^0-\sum_{\tau=L+1}^{t}x_{\tau-L}^{TLRP}(\tilde{\boldsymbol{d}}_{\tau-L-1})+\sum_{\tau=1}^{t}d_\tau(\tilde{\boldsymbol{z}})\right)^{+}\right)\\
&=\mathrm{E}\Bigg(\Bigg(-y_1^0-\sum_{\tau=1}^{L}x_{\tau-L}^0-\sum_{\tau=L+1}^{t}\min\left\{\max\left\{x_{\tau-L}^{0*}+\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}},0\right\},S_t\right\}-\\
&\qquad\sum_{\tau=1}^{t}d_\tau^0+\sum_{\tau=1}^{t}\sum_{k=1}^{N}d_\tau^k\tilde{z}_k\Bigg)^{+}\Bigg)\\
&\le\mathrm{E}\Bigg(\Bigg(-y_1^0-\sum_{\tau=1}^{L}x_{\tau-L}^0-\sum_{\tau=L+1}^{t}\min\left\{x_{\tau-L}^{0*}+\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}},S_t\right\}+\\
&\qquad\sum_{\tau=1}^{t}d_\tau^0-\sum_{k=1}^{N}\sum_{\tau=1}^{t}d_\tau^k\tilde{z}_k\Bigg)^{+}\Bigg)\\
&=\mathrm{E}\Bigg(\Bigg(-y_1^0-\sum_{\tau=1}^{L}x_{\tau-L}^0-\sum_{\tau=L+1}^{t}\left(x_{\tau-L}^{0*}-\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}}\right)+\\
&\qquad\sum_{\tau=L+1}^{t}\left(-\min\left\{S_t-x_{\tau-L}^{0*}-\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}},0\right\}\right)+\sum_{\tau=1}^{t}d_\tau^0+\sum_{k=1}^{N}\sum_{\tau=1}^{t}d_\tau^k\tilde{z}_k\Bigg)^{+}\Bigg)\\
&=\mathrm{E}\Bigg(\Bigg(\underbrace{-y_1^0-\sum_{\tau=1}^{L}x_{\tau-L}^0-\sum_{\tau=L+1}^{t}x_{\tau-L}^{0*}+\sum_{\tau=1}^{t}d_\tau^0}_{=-y_{t+1}^{0*}}+\\
&\qquad\sum_{\tau=L+1}^{t}\left(x_{\tau-L}^{0*}-S_t+\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}}\right)^{+}+\sum_{k=1}^{N}\underbrace{\left(\sum_{\tau=1}^{t}(-x_{\tau-L}^{k*}+d_\tau^k)\right)}_{=-y_{t+1}^{k*}}\tilde{z}_k\Bigg)^{+}\Bigg)\\
&=\mathrm{E}\left(\left(-y_{t+1}^{0*}-\boldsymbol{y}_{t+1}^{*}{}'\tilde{\boldsymbol{z}}+\sum_{\tau=L+1}^{t}\left(x_{\tau-L}^{0*}-S_t+\boldsymbol{x}_{\tau-L}^{*}{}'\tilde{\boldsymbol{z}}\right)^{+}\right)^{+}\right)\\
&\le\eta\left((-y_{t+1}^{0*},-\boldsymbol{y}_{t+1}^{*}),(x_1^{0*}-S_t,\boldsymbol{x}_1^{*}),\ldots,(x_{t-L}^{0*}-S_t,\boldsymbol{x}_{t-L}^{*})\right).\blacksquare
\end{aligned}
$$

Appendix F

Sample Formulation in *PROF*

The following is a sample formulation of Problem (3.17) in *PROF* is presented in Table F.1. Note that the function `meanpositivebound()` implements $\pi(\cdot)$ of Equation (3.7), and `meannestedposbound()` implements $\eta(\cdot)$ of Theorem 4.

Table F.1: Formulation of Problem (3.17) using *PROF*

```
Zmax = 20;     T = 10;
Z.zlow = Zmax*ones(N,1);     Z.zupp = Zmax*ones(N,1);
Z.p = .58*Zmax*ones(N,1);     Z.q = .58*Zmax*ones(N,1);
Z.sigma =.58*Zmax*ones(N,1);
Ny = [0 1:T];
Nx = [zeros(1,L) 0:T-L-1];     Nxms = [zeros(1,L) 0:T-L-1];
zcoef = eye(T,T);     MeanD = mu*ones(T,1);
for n = 2:T
  zcoef(1:n-1,n)= alpha;
end

startmodel % Start PROF
x = linearrule(T,N,Nx);     xms = linearrule(T,N,Nxms);
y = linearrule(T+1,N,Ny);
for i=1:T
  addconst(xms(i,:)  == x(i,:)-S*ldrdata([0 1],N));
end

hbound=0;     sbound=0;
for t=1:T
  if L+1 ≤ t
    hbound = hbound +
          h*meannestedposbound(Z,y(t+1,0:t),-x(L+1:t,0:t),t);
    sbound = sbound +
          b(t)*meannestedposbound(Z,-y(t+1,0:t),xms(L+1:t,0:t),t);
  else
     hbound = hbound+h*meanpositivebound(Z,y(t+1,:),1,N);
     sbound = sbound + b(t)*meanpositivebound(Z,-y(t+1,:),1,N);
   end
end

minimize (sbound+hbound +
          c*sum(meanpositivebound(Z,x(L+1:T,:),T-L,N)))
addconst(x(1:L,0)==initx); addconst(y(1,0)==inity);
for i=1:T
  addconst(y(i+1,:)  ==
          y(i,:)+x(i,:)-ldrdata([0 MeanD(i);(1:N)' zcoef(:,i)],N));
end
m=endmodel;     s = m.solve('MOSEK')
```

Appendix G

Raw Data of Example

Table G.1 shows the monthly net electricity generation extracted from Electric Power Monthly found in the web-site `http://www.eia.doe.gov` for the period of 2005 to 2007. The Electric Power Monthly presents monthly electricity statistics for a wide audience including the general public. The purpose of the publication is to provide energy decision makers with accurate and timely information that may be used in forming various perspectives on electricity issues that lie ahead. In the web-site data, the figures from Jul 2007 is missing. The data is presented in units of 10000 MWh, that is, a figure of 50 in the table means 50×10000 MWh. We use B to denote the electricity generated by the facility called Barry, and G to denote the electricity generated by the facility Gorgas. It was stated in the web-site data that both facilities are operated by the Alabama Power Co in the U.S.

Table G.1: Monthly electricity generation data

	2007		2006		2005	
	B	G	B	G	B	G
Dec	111	69	93	60	119	71
Nov	107	74	99	70	95	61
Oct	93	81	88	68	141	73
Sep	117	61	92	57	79	68
Aug	143	74	125	66	134	72
Jul	-	-	146	78	137	72
Jun	144	78	146	82	113	80
May	136	83	154	85	128	83
Apr	134	48	142	65	122	78
Mar	147	11	132	68	130	36
Feb	107	12	125	66	122	35
Jan	137	67	121	66	140	62