STOCHASTIC OPTIMIZATION WITH DECISIONS TRUNCATED BY RANDOM VARIABLES AND ITS APPLICATIONS IN OPERATIONS

BY

XIANGYU GAO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Xin Chen, Chair
Associate Professor Michael Lim
Associate Professor Qiong Wang
Assistant Professor Linwei Xin

Abstract

We study stochastic optimization problems with decisions truncated by random variables and its applications in operations management. The technical difficulty of these problems is that the optimization problem is not convex due to the truncation. We develop a transformation technique to convert the original non-convex optimization problems to convex ones while preservation some desired structural properties, which are useful for characterizing optimal decision policies and conducting comparative statics. Our transformation technique provides a unified approach to analyze a broad class of models in inventory control and revenue management. In additional, we develop efficient algorithms to solve the transformed stochastic optimization problem.

Chapter 2 introduces the transformation technique and the preservation of structural properties. Chapter 3 applies this approach to analyze several important models in operations management, which includes inventory control models with random capacities and network revenue management using booking limits. Chapter 4 generalizes the transformation technique by allowing dependent random variables, a more general objective function and incorporating risk measures. Chapter 5 studies the computational issues and propose a heuristic algorithm based on the transformation technique.

To my parents, for their endless love.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Xin Chen. The wonderful guidance and financial supports he provided for me are essential for completing my study and this dissertation. Thanks to his sharp thoughts and observations, meetings with him can often bring me new ideas and breakthroughs. In addition, his patience allows me to explore related literature about a specific problem to get a big picture. Last but not least, his enthusiasm about research encourages me to stay hungry and enjoy the research I am doing.

In terms of funding, I acknowledge the support by National Science Foundation (NSF) Grants CMMI-1363261, CMMI-1538451, CMMI-1635160.

I would like to thank my thesis committee members Professor Xin Chen, Professor Michael Lim, Professor Qiong Wang, Professor Linwei Xin for their efforts and valuable suggestions. I would also like to thank Prof. Zhan Pang and Zhenyu Hu for their efforts and contributions to this thesis.

I am very fortunate to have great and dedicated professors at University of Illinois at Urbana-Champaign so that I have learned a great deal from the classes and discussions with them. I want to thank Prof. Qiong Wang, Prof. Enlu Zhou, Prof. Angelia Nedich, Prof. Uday Shanbhad and many other members of the IESE faculty for their teachings. I also want to thank Prof. Dan Bernhardt from the Economics department and Prof. Nolan Miller from the College of Business for their teachings as well as discussions of economics models that helped my study and research. During my study in the Department of Industrial and Enterprise Systems Engineering, I was blessed with the support from the staff members, especially Holly Kizer and Barbara Bohlen. Their dedicated work helped me to concentrate on my study and research.

I am grateful to the countless hours of discussions as well as the time spent with my friends and colleagues in our research group, including Zhenyu Hu, Shuanglong Wang, Limeng Pan, Wenbo Chen, Juan Xu and Menglong Li. I am also very fortunate to have friends during my study at University of Illinois such as Hao Jiang, Fan Ye, Jingnan Chen, Tao Zhu, Haohua Wan, Jeehee Lee, Eric and Joy Seward, Bob and Debbie Whatoff, David Friday, Peter Sempsrott, and many others for making my PhD life enjoyable and encouraging me during difficult times.

Finally, I am immensely indebted to my father Haixia Gao and my mother Ping Zhang, for their unconditional and tireless supports and encouragements.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivations

The central theme of supply chain management is matching supply with demands. The main driving force behind a mismatch is the uncertainties involved. Uncertainties on the demand side have been studied for decades, starting with the classic newsvendor models. But in recent years, uncertainties on the supply side has attracted considerable attentions from both industry and academia. One example is the 2011 Japan Tsunami, which had a huge impact on supplies in the automobile industry. Many everyday operations can result in supply uncertainties as well. For instance, machines may break down unexpectedly and need maintenance to resume working. Some workers may not show up causing labor variance. There may be quality issues, power breakouts, or shortage of raw material during the production process. In the delivery process, there may be transportation issues that lead to partial delivery.

In the literature, there are several papers studying inventory control problem with random supply capacities. In these problems, the firm firstly makes an ordering decision or production decision. Since the supply capacity is uncertain, the firm's actual received amount is the minimum of the ordering quantity and the realization of random capacity.

Interestingly, similar structures occur in revenue management problems as well. For instance, in an airline revenue management using booking limits, the airline company firstly set the booking limits for each fare classes before the random demands are realized. The actual amount of tickets sold is the minimum of the booking limits and demands.

All of the above problems share a common technical challenge, i.e., the optimization problem formulated by the decision maker is not convex. Existing methods to deal with this technical challenge in the literature are all problem specific, which usually involve very lengthy proofs depending on analyses of derivatives.

This objective of this thesis is to provide a unified approach to tackling this technical challenge. We developed a transformation technique that can convert the original non-convex optimization problems to equivalent convex ones while preserving some desired structural properties, which are essential for structural analysis. Our transformation technique can not only greatly simplify the analysis of some existing problems, but also help analyze and solve new problems as well. In additional, the transformation technique also enables us to develop efficient algorithms to solve these challenging problems numerically.

1.2 Organization of the Thesis

Chapter 2 develops the transformation technique as well as derives some structural preservation results. Section 2.1 gives a brief introduction. Section 2.2 starts with developing the transformation technique for unconstrained optimization problems. It then proceeds to constrained problems with more general objective functions. In section 2.3, we establish some structural preserving properties based our transformation technique. These properties are useful for structural analysis and conducting comparative statics, which will be shown in later chapters.

Chapter 3 demonstrate the applications of our methodology with several interesting models in operations management. Section 3.1 reviews some related literature. 3.2 studies the inventory transshipment model with random capacities. This model has been studied in the literature. Nevertheless, we manage to show that applying the transformation technique can greatly simplify the analysis. Section 3.3 considers a dual sourcing problem with supply capacities and arbitrary lead times. Section 3.4 investigate an assemble-to-order system with random capacities. In section 3.5, different from the aforementioned inventory models, we look into a network revenue management problem using booking limits.

Chapter 4 generalize the transformation technique developed in Chapter 2 in a variety of ways. Section 4.1 introduces the motivation of these generalizations. Section 4.2 consider an optimization problem with dependent random variables and a more general objective function form. Section 4.3 further incorporates decision maker's risk attitude into the model.

Chapter 5 studies the computational issues and propose a heuristic algorithm based on the transformation technique. Section 5.1 introduce a single-period inventory substitution problem which will be used for our computational studies. Section 5.2 provides two solution procedures that can be used to solve the transformed formulation of the inventory substitution problem. The first one is a linear programming formulation based on Monte Carlo sampling. The second one is a heuristic approach utilizing piecewise linear decision rules. Section 5.3 shows the computational studies, which compare the performance of the aforementioned two methods in terms of running time and accuracy.

Finally, the last chapter concludes this thesis by summarizing the directions for future research.

In this thesis, we use decreasing, increasing and monotonicity in a weak sense. We use $\Re$ and $\Re_+$ to denote the real space and the set with nonnegative reals, $\mathcal{Z}$ and $\mathcal{Z}_+$ to denote the set of integers and the set of nonnegative integers, respectively. For convenience, let $\mathcal{F}$ be either $\Re$ or $\mathcal{Z}$. Define $\bar{\Re} = \Re \cup \{\infty\}$, $e \in \mathcal{F}^n$ a vector whose components are all ones, e_j a unit vector whose jth component is one, and for $x, y \in \mathcal{F}^n$, $x \leq y$ if and only if $x_i \leq y_i$ for any $i = 1, ..., n$, $x^+ = \max(x, 0)$, $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$ (the component-wise minimum and maximum operations). The indicator function of any set $\mathcal{V} \subseteq \mathcal{F}^n$, denoted by $\delta_{\mathcal{V}}$, is defined as $\delta_{\mathcal{V}}(x) = 0$ for $x \in \mathcal{V}$ and $+\infty$ otherwise. We use the superscript T to denote the transpose of a vector or a matrix. We use uppercase letters (e.g. Ξ) to denote random vectors and lowercase letters (e.g., ξ) for their realizations. Given a random vector $\Xi = (\Xi_1, ..., \Xi_n)^T$, we use $\mathcal{X} = Supp(\Xi)$ to denote the support of this random vector. In addition, we define $\bar{\xi}_j = \operatorname{ess\,sup}\{\xi_j | \xi_j \in \mathcal{X}_j\}$, $\underline{\xi}_j = \operatorname{ess\,inf}\{\xi_j | \xi_j \in \mathcal{X}_j\}$ for $j = 1, ..., n$, where $\mathcal{X}_j$ is $\mathcal{X}$'s projection into the j-th coordinate. Let $\bar{\xi} = (\bar{\xi}_1, ..., \bar{\xi}_n)^T$, $\underline{\xi} = (\underline{\xi}_1, ..., \underline{\xi}_n)^T$, and almost surely is abbreviated as a.s..

Chapter 2

Transformation Technique

2.1 Introduction

In operations management literature, a common technical challenge encountered in many models is that decision variables are truncated by some random variables and the decisions are made before the values of these random variables are realized. A notable example is inventory control problems with supply capacity uncertainty in which the replenishment decision is truncated by the random supply capacity (see, e.g., Ciarallo et al. 1994, Wang & Gerchak 1996, Bollapragada et al. 2004, Hu et al. 2008, Feng 2010 and Feng & Shi 2012). Another example is capacity allocation problems in revenue management where the booking limit of each demand class is truncated by the random demand (see, e.g., Brumelle & Mcgill 1993, Robinson 1995 and Chen & Homem-de Mello 2010). This type of variable truncation often leads to stochastic optimization problems in the following form:

$$g(x, z) = \inf_{u:(x,z,u)\in\mathcal{A}} E[f(x, u \wedge (z + \Xi))],$$

where f is a function in decision variables u and state variables (x, z), $\mathcal{A}$ is the constraint set, Ξ is a random vector, and $\wedge$ denotes componentwise minimum.

For these applications, it is natural to ask how to solve the above optimization problem efficiently and whether the optimization operation can preserve some desired structural properties of f such as convexity or submodularity. However, solving and analyzing such a problem can be very difficult. An intrinsic challenge arises from the fact that the truncation by random variables may destroy convexity: the objective function may not be convex in the decision variables even if the function f is convex. Without the regular properties such as convexity, the problem could be both analytically and computationally intractable, in particular when facing multidimensional state and decision variables.

This chapter aims at addressing this challenge when the random variables are independently distributed by developing a novel transformation technique which converts the non-convex minimization problem (2.1) to an equivalent convex minimization problem. As we mentioned earlier, the original problem formulation may be non-convex for a convex function f because in the objective function there are terms involving the minimum of decision variables and random variables. We prove that the optimal objective values of the original and transformed problems are the same when f is convex and certain regularity conditions are imposed on $\mathcal{A}$. Furthermore, our transformation technique allows us to show that the optimization operation in problem (2.1) can preserve convexity, submodularity or $L^\natural$-convexity, which then enables us to perform comparative statics analysis in multi-dimensional state and decision spaces and characterize the monotone structure of optimal policies.

2.2 Transformation Technique

To study the optimization problem (2.1), we start with an unconstrained optimization problem without state variables. Given a function $f : \mathcal{F}^n \to \bar{\Re}$ and a random vector Ξ with $Supp(\Xi) = \mathcal{X} \subseteq \mathcal{F}^n$, consider the following optimization problem

$$\tau^* = \inf_{u \in \mathcal{F}^n} E[f(u \wedge \Xi)]. \tag{2.1}$$

In general, the above problem may not be a convex minimization problem even if the function f is convex. For instance, let $f(u) = u^2$ and Ξ can take values $0, 1$ or 2 with equal probabilities. In Figure 2.1, we plot the objective function $E[f(u \wedge \Xi)]$ with respect to the decision variable u. It is easy to see that it is not a convex function. Notice that in this example where $u \in \Re$, the objective function is quasi-convex. However, this is no longer true when the decision variable is multi-dimensional. For example, let $f(u) = \frac{1}{2}u^T Hu + c^T u$ where $H = [4, 2; 2, 3]$, $c = [-8, -2]$. Ξ_1 can take values 0 and 2 with equal probabilities. Ξ_2 can take values 1 and 3 with equal probabilities. Ξ_1 and Ξ_2 are independent of each other. Figure 2.2 demonstrates the contour map of the objective function with respect to (u_1, u_2). We can see that the objective function is not quasi-convex since the lower contour map is not always a convex set. Interestingly, we will show that under certain conditions, we can convert problem (2.1) into an equivalent convex minimization problem.

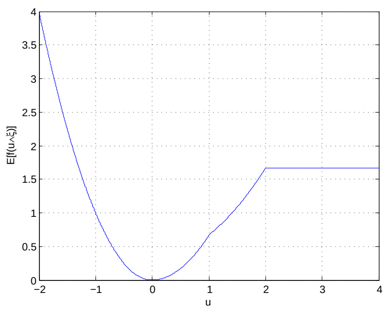


Figure 2.1: A one-dimensional example

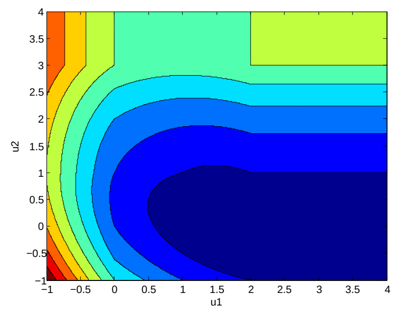


Figure 2.2: A two-dimensional example

For this purpose, note that the optimization problem (2.1) can be rewritten as follows.

$$\begin{aligned} \inf \quad & E[f(v(\Xi))] \\ \text{s.t.} \quad & v(\xi) = u \wedge \xi \quad \forall \xi \in \mathcal{X}, \\ & u \in \mathcal{F}^n, v(\cdot) \in \mathcal{M}, \end{aligned} \tag{2.2}$$

where $\mathcal{M}$ is the set of measurable functions. The feasible region of (2.2) is $\mathcal{F}^n \times (\mathcal{F}^n)^{\mathcal{X}}$ while the feasible region of (2.1) is $\mathcal{F}^n$. In the following theorem, we show that the equality constraint $v(\xi) = u \wedge \xi$ can be relaxed by the inequality constraint $v(\xi) \le \xi \ \forall \xi \in \mathcal{X}$ with $v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{F}^n$. For the rest of this chapter, we require that $v(\cdot)$ is measurable in all of our formulations and therefore omit $v(\cdot) \in \mathcal{M}$ for brevity. The following lemma will be useful for the proof of the theorem.

Lemma 2.1. *Suppose that the function $f : \mathcal{F} \to \bar{\Re}$ is quasi-convex. If x^* is a minimizer of $f(x)$ over $\mathcal{F}$, we have $f(x^* \wedge b) \le f(a)$ for any $a, b \in \mathcal{F}$ with $a \le b$.*

The equivalent transformation of problem (2.1) is given by the following theorem.

Theorem 2.1. *Suppose that (a) the function $f : \mathcal{F}^n \to \bar{\Re}$ is lower semi-continuous with $f(x) \to +\infty$ for $|x| \to \infty$; (b) f is componentwise convex (componentwise discrete convex if $\mathcal{F} = \mathcal{Z}$); (c) the random vector Ξ has independent components and it has realizations $\xi \in \mathcal{X} = Supp(\Xi)$. Then, τ^* defined in* (2.1) *is also the optimal objective value of the following optimization problem.*

$$\begin{aligned} \inf \quad & E[f(v(\Xi))] \\ s.t. \quad & v(\xi) \le \xi \quad \forall \xi \in \mathcal{X}, \\ & v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{F}^n \quad \forall \xi \in \mathcal{X}. \end{aligned} \tag{2.3}$$

Remark 2.1. *In the proof of the above theorem (see Appendix A), we illustrate that when $n = 1$,*

$$\min_{u \in \mathcal{F}} E[f(u \wedge \Xi)] = E[f(\hat{u} \wedge \Xi)],$$

where $\hat{u}$ is any minimizer of the function f. In fact, this observation is still valid when f is quasi-convex.

However, when $n > 1$, such a result no longer holds, i.e., $\min_{u \in \mathcal{F}} E[f(u \wedge \Xi)] \neq E[f(\hat{u} \wedge \Xi)]$, even if f is jointly convex. We now present an example. Specifically, let

$$n = 2, \mathcal{F} = \Re, f(u_1, u_2) = (u_1 + u_2 - 2)^2 + (u_1 - 1)^2 + (u_2 - 1)^2,$$

and Ξ_1 and Ξ_2 be independent and identically distributed and take values 0 *and* 2 *with equal probabilities. In this case, $\hat{u} = (1, 1)$. However, one can easily verify that $\arg\min_{u \in \mathcal{F}^n} E[f(u \wedge \Xi)] = (1.2, 1.2) \neq \hat{u}$.*

Remark 2.2. *In the above theorem, we require that*

$$v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n)).$$

This cannot be relaxed to allow

$$v(\xi) = (v_1(\xi), \ldots, v_n(\xi)).$$

To illustrate this, we use the above example again. Note that for problem

(2.3), the optimal objective value is 2.4 *and an optimal solution is given by*

$$v_1^*(0) = v_2^*(0) = 0, v_1^*(2) = v_2^*(2) = 1.2.$$

However, if one replaces $v(\xi) = (v_1(\xi_1), \dots, v_n(\xi_n))$ *by* $v(\xi) = (v_1(\xi), \dots, v_n(\xi))$ *in problem (2.3), the optimal objective value becomes* 2.25 *and an optimal solution is given by*

$$v^*(0,0) = (0,0), v^*(0,2) = (0,1.5), v^*(2,0) = (1.5,0), v^*(2,2) = (1,1).$$

Remark 2.3. *It is interesting to observe that* u *does not appear in problem* (2.3)*. Our proof implies that given an optimal solution* u^* *of problem* (2.1)*,* $v^* = (v^*(\xi) = u^* \wedge \xi | \xi \in \mathcal{X})$ *is optimal for problem* (2.3)*. On the other hand, given an optimal solution* v^* *of problem* (2.3)*, we can directly construct an optimal solution of problem* (2.1) *without solving any additional optimization problem. To see this, we start with* $n = 1$ *and define* $S = \{\xi | v^*(\xi) < \xi, \xi \in \mathcal{X}\}$ *(for simplicity, we drop the subscript* 1 *in the presentation when* $n = 1$*). We consider two cases depending on whether the probability of event* S*, denoted by* $P(S)$*, is zero or not. In the first case,* $P(S) > 0$*. Randomly pick* $\hat{\xi} \in S$ *according to the probability distribution of* Ξ *conditional on* S *and define* $\hat{u} = v^*(\hat{\xi})$*. It suffices to show that* $\hat{u}$ *is optimal for the optimization problem* $\min_{u \in \mathcal{F}} f(u)$ *with probability* 1*. Suppose this is not true and* $P(S') > 0$*, where* S' *is the event such that* $\hat{\xi} \in S$ *and* $v^*(\hat{\xi})$ *is not optimal for* $\min_{u \in \mathcal{F}} f(u)$*. We define a new feasible solution of problem (2.3):*

$$\hat{v}(\xi) = \begin{cases} v^*(\xi), & \text{if} \quad \xi \notin S', \\ u^0 \wedge \xi, & \text{if} \quad \xi \in S', \end{cases}$$

where u^0 *is an optimal solution of* $\min_{u \in \mathcal{F}} f(u)$*. If* $\xi \notin S'$*, then* $\hat{v}(\xi) = v^*(\xi)$ *and* $f(\hat{v}(\xi)) = f(v^*(\xi))$*. If* $\xi \in S'$ *and* $\xi > u^0$*,* $f(\hat{v}(\xi)) = f(u^0) < f(v^*(\xi))$*. If* $\xi \in S'$ *and* $\xi \leq u^0$*,* $v^*(\xi) < \xi \leq u^0$ *and the convexity of* f *implies that* $f(\hat{v}(\xi)) = f(\xi) < f(v^*(\xi))$*. Since* $P(S') > 0$*, we have* $E[f(\hat{v}(\xi))] < E[f(v^*(\xi))]$*, which is a contradiction. Therefore, with probability* 1*,* $\hat{u}$ *is optimal for the optimization problem* $\min_{u \in \mathcal{F}} f(u)$*. In the second case,* $P(S) = 0$*. Note that* f *must be decreasing over* $\mathcal{X}$*, otherwise we can easily construct a feasible solution of problem (2.3) with a lower cost. Hence, assumption (a) implies that* $\bar{\xi} < \infty$*, and* $\hat{u} = \bar{\xi}$ *is a minimizer of the function* f *on* $\mathcal{F}$*. For*

$n > 1$, define, for $i = 1, \ldots, n$, event $S_i = \{\xi_i | v_i^*(\xi_i) < \xi_i\}$. If the probability of S_i is positive, randomly pick $\hat{\xi}_i \in S_i$ according to the probability distribution of Ξ conditional on S_i and define $\hat{u}_i = v_i^*(\hat{\xi}_i)$; otherwise, define $\hat{u}_i = \bar{\xi}_i$ (again $\bar{\xi}_i < \infty$). Since the components of the random vector Ξ are independent, we can extend the above analysis to show that, with probability 1, $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_n)$ is an optimal solution of problem (2.1).

We can explicitly incorporate constraints on u in Theorem 2.1 and consider a more general objective function to allow both componentwise minimum and maximum operations. To simplify notations, we define an operator $\Diamond_k$ as $u\Diamond_k\xi \triangleq (u_1 \wedge \xi_1, ..., u_k \wedge \xi_k, u_{k+1} \vee \xi_{k+1}, ..., u_n \vee \xi_n)$, so that the first k terms we have the componentwise minimum operator, while the $n-k$ terms left we have the componentwise maximum operation. The problem of interest is

$$\inf_{u \in \mathcal{U}} E[f(u\Diamond_k\Xi)], \tag{2.4}$$

where $f : \mathcal{F}^n \to \bar{\Re}$ and $\mathcal{U} \subseteq \mathcal{F}^n$. Define a set

$$\mathcal{V} = \{u\Diamond_k\xi | u \in \mathcal{U}, \xi \in \mathcal{X}\}. \tag{2.5}$$

We impose the following assumption:

Assumption 2.1.

(a) For any $u \in \mathcal{F}^n$ such that $u\Diamond_k\xi \in \mathcal{V}$, $\forall \xi \in \mathcal{X}$, there exists $u' \in \mathcal{U}$ such that $u'\Diamond_k\xi = u\Diamond_k\xi$, $\forall \xi \in \mathcal{X}$.

(b) The indicator function of the set $\mathcal{V}$ is componentwise convex (componentwise discrete convex if $\mathcal{F} = \mathcal{Z}$).

Notice that Part (a) of the above assumption implies that if $u\Diamond_k\xi \in \mathcal{V}$, $\forall \xi \in \mathcal{X}$, we do not necessarily need $u \in \mathcal{U}$. Instead, we only require that there exists $u' \in \mathcal{U}$ such that $u'\Diamond_k\xi = u\Diamond_k\xi$, $\forall \xi \in \mathcal{X}$. As can be seen from the proof of Theorem 2.2 below, Assumption 2.1 allows us to convert the constrained optimization problem (2.4) to an equivalent unconstrained optimization problem so that Theorem 2.1 can be applied.

Theorem 2.2. *Consider the optimization problem (2.4), where $f : \mathcal{F}^n \to \bar{\Re}$ and the random vector Ξ in $\mathcal{F}^n$ satisfy the assumptions in Theorem 2.1.*

Suppose that Assumption 2.1 is satisfied. Problem (2.4) and the following optimization problem have the same optimal objective value.

$$\begin{array}{ll} \inf & E[f(v_1(\Xi_1), \ldots, v_n(\Xi_n))] \\ s.t. & v_j(\xi_j) \leq \xi_j \quad \forall\ \xi_j \in \mathcal{X}_j, j = 1, \ldots, k \\ & v_j(\xi_j) \geq \xi_j \quad \forall\ \xi_j \in \mathcal{X}_j, j = k+1, \ldots, n \\ & (v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathcal{V}\ \forall \xi \in \mathcal{X}. \end{array} \tag{2.6}$$

Since Assumption 2.1 may not be easy to check, in the following we provide a nontrivial example with linear constraints under which Assumption 2.1 holds in the following lemma.

Lemma 2.2. *Assume that* $\mathcal{U} = \{u \in \mathcal{F}^n | Au \leq b, u_1 \geq \underline{u}_1, ..., u_k \geq \underline{u}_k, u_{k+1} \leq \bar{u}_{k+1}, ..., u_n \leq \bar{u}_n\}$, *where* $b, \underline{u}_1, ..., \underline{u}_k, \bar{u}_{k+1}, ..., \bar{u}_n$ *are given constants,* $A = (a_{ij})$ *with entries* $a_{ij} \geq 0$ *for any* i *and* $j = 1, ..., k$, *and* $a_{ij} \leq 0$ *for any* i *and* $j = k+1, ..., n$. *In addition* $\xi_j \geq \underline{u}_j\ \forall \xi_j \in \mathcal{X}_j, j = 1, ..., k$, *and* $\xi_j \leq \overline{u}_j\ \forall \xi_j \in \mathcal{X}_j, j = k+1, ..., n$. *Then Assumption 2.1 is satisfied.*

Remind that in Part (a) of Assumption 2.1, we do not need $u \in \mathcal{U}$. We will illustrate this by the following example which satisfies the conditions in Lemma 2.2

Example 2.1. *Suppose* $k = n = 1$. *Let* $\mathcal{U} = \{u | 0 \leq u \leq 1\}$ *and* ξ *is uniformly distributed between* 0 *and* 1. *Then the set* $\mathcal{V} = \{u \wedge \xi | 0 \leq u \leq 1, 0 \leq \xi \leq 1\} = \{v | 0 \leq v \leq 1\}$. *Take* $u = 2$, *notice that* $u \wedge \xi \in \mathcal{V}$ *but* u *does not belong to the set* $\mathcal{U}$. *However, there exists* $u' = 1$ *which belongs to the set* $\mathcal{U}$ *such that* $u \wedge \xi = u' \wedge \xi\ \forall \xi \in \mathcal{X}$.

In Lemma 2.2, an important condition is that $\xi_j \geq \underline{u}_j\ \forall \xi_j \in \mathcal{X}_j, j = 1, ..., k$, and $\xi_j \leq \overline{u}_j\ \forall \xi_j \in \mathcal{X}_j, j = k+1, ..., n$.. To illustrate why such a condition is needed, we provide an example as follows.

Example 2.2. *Suppose that* $k = n = 1$, $f(u) = u^2$, $\mathcal{U} = [1, 2]$, $\xi = 0$ *or* 2 *with equal probability. Here we have* $\xi < \underline{u}$ *when* $\xi = 0$. *For the original problem* (2.4), *the optimal solution is* $u^* = 1$, *the optimal objective value is* 0.5. *For the transformed problem* (2.6), *we have* $\mathcal{V} = \{0\} \cup [1, 2]$, *which implies that the optimal solution is* $v(\xi = 0) = 0$, $v(\xi = 2) = 0$, *and the optimal objective value is* 0.

2.3 Preservation of Structural Properties

One advantage of our transformation technique is that it can be used to establish the preservation of not only convexity and submodularity but also $L^\natural$-convexity under optimization operations, which plays a critical role in characterizing the structure of the optimal policies for many dynamic decision making problems and facilitates their efficient computations. To see this, we first provide a brief review of the concept of $L^\natural$-convexity and some structural properties. $L^\natural$-convexity was defined by Murota (1998) as a fundamental concept to extend convex analysis from real space to spaces with integers (see Murota 2009 for a survey of the recent developments in discrete convex analysis). It was first introduced into the inventory management literature by Lu & Song (2005) and used by Zipkin (2008) to characterize the optimal structural policy of lost-sales inventory models with positive leadtimes. Since then, $L^\natural$-convexity was found to be powerful enough to establish the structures of optimal policies in various other inventory models: serial inventory systems (Huh & Janakiraman 2010); inventory-pricing models with positive leadtimes (Pang et al. 2012); and perishable inventory models (Chen et al. 2014); etc.

In the transformed problem the decisions are $v = (v(\xi)|\xi \in \mathcal{X}) \in (\mathcal{F}^n)^{\mathcal{X}}$. Note that the direct product of lattices is still a lattice under the componentwise partial order (see Example 2.2.3 (d) of Topkis 1998). Therefore, if X_α is a lattice for each $\alpha \in \mathbb{A}$, where $\mathbb{A}$ is an index set, then the direct product of sets X_α, is also a lattice. In the following we present the definition of $L^\natural$-convexity with domain $\mathcal{Y} \triangleq (\mathcal{F}^n)^{\mathcal{X}}$, where $\mathbb{A}$ is any index set.

Definition 2.1. *A function $f : \mathcal{Y} \to \bar{\Re}$ is $L^\natural$-convex if for any $x, x' \in \mathcal{Y}, \lambda \in \mathcal{F}_+$,*

$$f(x) + f(x') \geq f((x + \lambda e) \wedge x') + f(x \vee (x' - \lambda e)),$$

where e is the all-ones vector in $\mathcal{Y}$. A set $\mathcal{V} \subseteq \mathcal{Y}$ is said to be $L^\natural$-convex if its indicator function $\delta_{\mathcal{V}}$ is $L^\natural$-convex.

For an $L^\natural$-convex function f, its effective domain $\mathrm{dom}(f) = \{x \in \mathcal{Y} | f(x) < +\infty\}$ is an $L^\natural$-convex set. We sometimes say a function f is $L^\natural$-convex on a set $\mathcal{V}$ with the understanding that $\mathcal{V}$ is an $L^\natural$-convex set and the extension of f to the whole space by defining $f(\boldsymbol{v}) = +\infty$ for $v \notin \mathcal{V}$ is $L^\natural$-convex. One can also show that an $L^\natural$-convex function restricted to an $L^\natural$-convex set is also

$L^\natural$-convex. Following a similar proof in Simchi-levi et al. (2014), we can show that an equivalent definition of $L^\natural$-convexity is given as follows: A function $f : \mathcal{Y} \to \bar{\Re}$ is $L^\natural$-convex if and only if $g(x, \xi) \triangleq f(x - \xi e)$ is submodular in $(x, \xi) \in \mathcal{Y} \times \mathcal{S}$, where $\mathcal{S}$ is the intersection of $\mathcal{F}$ and any unbounded interval in $\Re$, and e is the all-ones vector in $\mathcal{Y}$.

We now list some of the commonly used properties of $L^\natural$-convexity. To describe the monotonicity of optimal solution sets, we use the *induced set ordering* $\sqsubseteq$ which defines $X' \sqsubseteq X''$ for two nonempty sets X' and X'' if $x' \in X'$ and $x'' \in X''$ imply that $x' \wedge x'' \in X'$ and $x' \vee x'' \in X''$ (see Topkis 1998, p32). For a nonempty set X_t that depends on the parameter t in a partial order set T, we say that X_t is increasing in t on T if $\{X_t, t \in T\}$ has the induced set ordering $\sqsubseteq$.

Proposition 2.1. *(a) Any nonnegative linear combination of $L^\natural$-convex functions is $L^\natural$-convex. That is, if $f_i : \mathcal{Y} \to \bar{\Re}(i = 1, 2, ..., n)$ are $L^\natural$-convex, then for any scalar $\alpha_i \geq 0$, $\sum_{i=1}^m \alpha_i f_i$ is also $L^\natural$-convex.*

(b) If f_k is $L^\natural$-convex for $k = 1, 2, \ldots$ and $\lim_{k\to\infty} f_k(x) = f(x)$ for any $x \in \mathcal{Y}$, then f is $L^\natural$-convex.

(c) Assume that a function $f(\cdot, \cdot)$ is defined on the product space $\mathcal{Y} \times \mathcal{F}^m$. If $f(\cdot, y)$ is $L^\natural$-convex for any given $y \in \mathcal{F}^m$, then for a random vector ζ defined on $\mathcal{F}^m$, $E_\zeta[f(x, \zeta)]$ is $L^\natural$-convex, provided it is well defined.

(d) If $f : \mathcal{Y} \to \bar{\Re}$ is an $L^\natural$-convex function, then $g : \mathcal{Y} \times \mathcal{F} \to \bar{\Re}$ defined by $g(x, \lambda) = f(x - \lambda e)$ is also $L^\natural$-convex.

(e) Assume that $\mathcal{A}$ is an $L^\natural$-convex set of $\mathcal{F}^n \times \mathcal{Y}$ and $f(\cdot, \cdot) : \mathcal{F}^n \times \mathcal{Y} \to \bar{\Re}$ is an $L^\natural$-convex function. Then the function

$$g(x) = \inf_{y:(x,y)\in\mathcal{A}} f(x, y)$$

is $L^\natural$-convex over $\mathcal{F}^n$ if $g(x) \neq -\infty$ for any $x \in \mathcal{F}^n$.

(f) Let e and $\tilde{e}$ be the all-ones vectors corresponding to the state s-pace of x and the decision space of y respectively in (10). Then $\arg\min_{y:(x,y)\in\mathcal{A}} f(x, y)$ is increasing in x and

$$\arg\min_{y:(x+\omega e,y)\in\mathcal{A}} f(x + \omega e, y) \sqsubseteq \omega\tilde{e} + \arg\min_{y:(x,y)\in\mathcal{A}} f(x, y).$$

(g) *Denote x_i a component of $x \in \mathcal{Y}$. A set with a representation $\{x \in \mathcal{Y} : l \le x \le u, x_i - x_j \le v_{ij}, \forall i \neq j\}$, is $L^\natural$-convex in the space $\mathcal{Y}$, where $l, u \in \mathcal{Y}$ and $v_{ij} \in \mathcal{F}$.*

(h) *A smooth function $f : \Re^n \to \Re$ is $L^\natural$-convex if and only if its Hessian is a diagonally dominated M-matrix, where a matrix A with its ij-th component being a_{ij} is called a diagonally dominated M-matrix, if*

$$a_{ij} \le 0, \quad \forall i \neq j, \quad a_{ii} \ge 0, \quad \text{and} \quad \sum_{j=1}^{n} a_{ij} \ge 0, \quad \forall i.$$

We now show how our transformation technique can be used to establish preservation properties of convexity, submodularity, and $L^\natural$-convexity under optimization operations.

Consider the following optimization problem

$$g(x, z) = \inf_{u:(x,z,u)\in\mathcal{A}} E[f(x, u\diamondsuit_k(z + \Xi))], \tag{2.7}$$

where $f(\cdot,\cdot) : \mathcal{F}^m \times \mathcal{F}^n \to \bar{\Re}$, $x \in \mathcal{F}^m, z \in \mathcal{F}^n$ and set $\mathcal{A} \subseteq \mathcal{F}^m \times \mathcal{F}^n \times \mathcal{F}^n$ is non-empty.

Define a set

$$\mathcal{A}^{\Xi} = \{(x, z, w) | w = u\diamondsuit_k(z + \xi), (x, z, u) \in \mathcal{A}, \xi \in \mathcal{X}\}.$$

Similar to Assumption 2.1, we specify the following condition:

Assumption 2.2.

(a) *For any (x, z, u) such that $(x, z, u\diamondsuit_k(z + \xi)) \in \mathcal{A}^{\Xi}$ $\forall \xi \in \mathcal{X}$, there exists $(x, z, u') \in \mathcal{A}$ such that $u'\diamondsuit_k(z + \xi) = u\diamondsuit_k(z + \xi)$ $\forall \xi \in \mathcal{X}$.*

(b) *The indicator function of the set $\mathcal{A}^{\Xi}$ is componentwise convex in w (componentwise discrete convex if $\mathcal{F} = \mathcal{Z}$).*

Similar to Lemma 2.2, we provide an example with linear constraints which satisfies Assumption 2.2. The proof is similar and thus omitted for brevity.

Lemma 2.3. *Assume that $\mathcal{A} = \{(x, z, u) | Au \le b, u_1 \ge \underline{u}_1, ..., u_k \ge \underline{u}_k, u_{k+1} \le \bar{u}_{k+1}, ..., u_n \le \bar{u}_n\}$, where $b, \underline{u}_1, ..., \underline{u}_k, \bar{u}_{k+1}, ..., \bar{u}_n$ are parameters that may depend on x and z, $A = (a_{ij})$ with entries $a_{ij} \ge 0$ for any i and*

$j = 1, ..., k$, and $a_{ij} \le 0$ for any i and $j = k+1, ..., n$. In addition $\mathcal{X}_j$ is contained in $[\underline{u}_j - z_j, +\infty)$ for $j = 1, ..., k$, and $\mathcal{X}_j$ is contained in $(-\infty, \overline{u}_j - z_j]$ for $j = k+1, ..., n$. Then Assumption 2.2 is satisfied.

Now we are ready to present our results on the preservation of structural properties.

Theorem 2.3 (PRESERVATION)**.** *Consider the optimization problem* (2.7)*, where f and Ξ satisfy the assumptions in Theorem 2.1 given any* (x, z)*. If Assumption 2.2 is satisfied, then we have the following results:*

(a) If f and $\mathcal{A}^{\Xi}$ are convex, then g is also convex.

(b) If f is submodular and $\mathcal{A}^{\Xi}$ is a lattice, then g is also submodular.

(c) If f and $\mathcal{A}^{\Xi}$ are $L^{\natural}$-convex, then g is also $L^{\natural}$-convex.

The following theorem characterizes the monotonicity properties of the solution set to the optimization problem (2.7).

Theorem 2.4. *Consider the optimization problem* (2.7)*, where f and Ξ satisfy the assumptions in Theorem 2.1 given any* (x, z)*. Let $\mathcal{U}^*(x, z)$ denote the the optimal solution set of* (2.7)*. If Assumption 2.2 is satisfied, $\mathcal{A}$, $\mathcal{A}^{\Xi}$ are closed, and in addition $u_j \le z_j + \bar{\xi}_j, j = 1, ..., k, u_j \ge z_j + \underline{\xi}_j, j = k+1, ..., n$, then we have the following results:*

(a) If f is a submodular function, and $\mathcal{A}, \mathcal{A}^{\Xi}$ are lattices, then $\mathcal{U}^*(x, z)$ is increasing in (x, z)*. There exist a greatest element and a least element in $\mathcal{U}^*(x, z)$, which are increasing in* (x, z)*.*

(b) If f is an $L^{\natural}$-convex function, and $\mathcal{A}, \mathcal{A}^{\Xi}$ are $L^{\natural}$-convex sets, then $\mathcal{U}^*(x, z)$ is increasing in (x, z) and $\mathcal{U}^*((x, z) + \omega e) \sqsubseteq \mathcal{U}^*(x, z) + \omega e$ for any $\omega > 0$. Within $\mathcal{U}^*(x, z)$, there exist a greatest element and a least element, which have the above monotonicity properties with limited sensitivity.

In the following we provide an example to show that the assumption $u_j \le z_j + \bar{\xi}_j, j = 1, ..., k, u_j \ge z_j + \underline{\xi}_j, j = k+1, ..., n$ is needed.

Example 2.3. *Suppose that $f(u) = u^2$ and the support of Ξ is $[-3, -1]$. Let $\mathcal{U}^*(z) = \arg\min_{u \in \mathcal{U}} E[f(u \wedge (z + \Xi))]$, and $z = 0, \omega = 2$. When $\mathcal{U} = \Re$, we* *have* $\mathcal{U}^*(z) = [-1, \infty)$*,* $\mathcal{U}^*(z + \omega) = \{0\}$*. Notice that* $\mathcal{U}^*(z) \sqsubseteq \mathcal{U}^*(z + \omega)$ *does not hold. However, when* $\mathcal{U}(z) = \{u \in \Re : u \leq z + \bar{\xi}\}$*, we have* $\mathcal{U}(z) = (-\infty, -1], \mathcal{U}(z+\omega) = (-\infty, 1]$*. Then* $\mathcal{U}^*(z) = \{-1\}, \mathcal{U}^*(z+\omega) = \{0\}$*. Clearly* $\mathcal{U}^*(z) \sqsubseteq \mathcal{U}^*(z + \omega)$*.*

Notice that if the conditions in Lemma 2.3 are satisfied, then the assumptions $u_j \leq z_j + \bar{\xi}_j, j = 1, ..., k, u_j \geq z_j + \underline{\xi}_j, j = k + 1, ..., n$ in Theorem 2.4 are without loss of generality. To see this, given any (x, z, u) which is feasible for problem (2.7), $(x, z, u_1 \wedge (z_1 + \bar{\xi}_1), ..., u_k \wedge (z_k + \bar{\xi}_k), u_{k+1} \vee (z_{k+1} + \underline{\xi}_{k+1}), ..., u_n \vee (z_n + \underline{\xi}_n))$ is also feasible and yields the same objective value. In all the applications we are going to present later, the constraint set satisfies the conditions in Lemma 2.3.

Chapter 3

Applications

3.1 Introduction

The transformation technique we developed in Chapter 2 has a wide range of applications. In this chapter, we discuss several applications, which can be divided into two categories: (1) inventory management with supply capacity uncertainties, and (2) capacity allocation in network revenue management.

Supply uncertainty of inventory/production systems can be driven by a variety of factors. Most studies in this literature focus on random yield problems where the supply is a random proportion of the order quantity; see Henig & Gerchak (1990), Federgruen & Yang (2008, 2011), Chen, Feng & Seshadri (2013) and the references therein. Such an issue usually arises from the quantity uncertainty of items produced in a batch. Another important supply uncertainty is the supply capacity uncertainty due to the unreliability of the supply processes (e.g., partial delivery or cancellation of an order by the supplier). In such an environment, the firm has to place orders before knowing the actual supply capacity. There are relatively few papers addressing the random capacity problems.

Ciarallo et al. (1994) consider an inventory control problem, assuming that the replenishment decisions are made before the capacity uncertainty is realized and the replenishment leadtime is zero. They show that the presence of capacity uncertainty does not affect the optimality of a base-stock policy. Wang & Gerchak (1996) extend the analysis to systems with both random supply capacity and random yield. Feng (2010) addresses a joint pricing and inventory control problem with supply capacity uncertainty and zero leadtime and shows that the optimal policy is characterized by two critical values: a reorder point and a target safety stock. The common technical challenge of these models is that with random supply capacity, the corresponding dynamic programming recursions are not convex minimization (concave maximization) problems anymore, and delicate analyses are needed to characterize the structures of optimal policies.

Our transformation technique can be readily applied to the aforementioned models to simplify the structural analysis. More importantly, such an approach allows us to address more general inventory models under supply capacity uncertainty with multi-dimensional state spaces using the concept of $L^\natural$-convexity. In this chapter we will demonstrate three applications in the area of inventory management with supply uncertainty.

Our first application is to provide a new approach to the analysis of optimal joint inventory and transshipment control under uncertain capacity. Specifically, we consider the model studied in Hu et al. (2008). In this model, a firm operates two facilities in separate markets, where the firm produces the same product and sells at constant prices. Both facilities face uncertain demands and uncertain production capacities. The firm needs to determine the production quantities at the beginning of each period. The demand and production uncertainties are then revealed and the firm further decides how much inventory to be transshipped from one facility to another. Demands are satisfied after the transshipment and unfilled demands are lost.

Our second application is an inventory system with two capacitated suppliers, a regular one with a longer leadtime and an expedited one with a shorter leadtime. The two suppliers have independent supply capacity uncertainties. The objective of the firm is to find a dual-sourcing strategy to minimize the total expected cost. There is an extensive literature on the dual sourcing problem. It was first studied by Barankin (1961) in a one-period setting and then extended by Daniel (1963), Fukuda (1964) and Whittmore & Saunders (1977) to various settings with multi-period horizons. Feng & Shi (2012) consider a joint inventory control and pricing problem with multiple suppliers whose replenishment lead times are zero and supply capacities are uncertain. They show that with deterministic capacities a multi-level base-stock list-price policy plus a cost-based supplier selection (i.e., ordering from a cheaper source first) is optimal. However, with general random supply capacities, such a policy is no longer optimal. They show that the optimal policy can be characterized by a *near* reorder point such that a positive order is placed (almost everywhere) if and only if the inventory level is below this point. They also identify a condition under which a strict reorder-point policy and a cost-based supplier-selection criterion become optimal. More recently, Zhou & Chao (2014) address the dual-sourcing problem with price sensitive demand, a regular supplier with one-period leadtime and an expedited supplier with zero leadtime, and characterize the structure of the optimal policy. Gong et al. (2014) further generalize the structural analysis to a dual-sourcing problem with price sensitive demand and Markovian supply interruptions. In both models, there are no capacity limits on the supplies. To the best of our knowledge, our paper is the first addressing the dual-sourcing system with arbitrary deterministic leadtime discrepancies and supply capacity uncertainties.

Our third application is an assemble-to-order inventory system with multiple components and products. The order quantity of each component cannot exceed a random capacity. The firm decides the ordering quantities of all components and then the number of products assembled to minimize the expected cost. The assemble-to-order system is one of the most important production/inventory systems; see Song & Zipkin (2003) for a review of the research literature and applications of assemble-to-order systems up to the early 2000s. Lu & Song (2005) study a continuous-review assemble-to-order system with random demands and lead-times with an order-based approach. Nadar et al. (2014) develop the optimal structural results for a continuous-review assemble-to-order generalized M-system with lost sales. Bollapragada et al. (2004) study multi-echelon assembly systems under installation base-stock policies where the component suppliers have various leadtimes and random supply capacities. They propose a decomposition approach and their numerical study shows that their heuristic performs well in comparison with the optimal base-stock policy. In this paper, we show that our approach applies to the assemble-to-order system with random component capacity. Moreover, for the generalized M-system, we show that the cost-to-go functions are

Revenue Management (RM), also known as Yield Management, has been widely adopted in various industries such as airlines, hotels, car rentals and cruise lines. Driven by its prevalence in service industry, the research interest in RM has been growing rapidly over the last two decades; see Talluri and van Ryzin (2005) for a comprehensive introduction to the practice and theoretical developments of RM.

The forth application is the capacity allocation in network revenue management where fixed capacities of resources are allocated dynamically to different products with random demands. In the airline industry, this corresponds to setting booking limits for each itinerary-fare class combination. The booking limits are truncated by the random demand. The firm aims to maximize the expected total revenue. The network revenue management problem, which involves managing multiple resources (such as airline seats in different leg-cabin combinations), is notoriously challenging. Indeed, as mentioned by Talluri and van Ryzin (2005), "in the network case, exact optimization is for all practical purposes impossible", and thus the literature focuses predominantly on various approximations. One approximation is to formulate a stochastic programming problem (see Cooper & Homem-de Mello 2007, Möller et al. 2008, Chen & Homem-de Mello 2010 and the references therein). For example, one can formulate a two-stage stochastic linear programming problem (SLP) by aggregating the demand over the planning horizon and determining the booking limits at the beginning (see section 3.3.1 of Talluri and van Ryzin 2005). To improve upon the SLP, one can consider a multi-stage stochastic programming (MSSP), in which the policy of booking limits is revised from time to time in order to take into account the information about demand learned so far. The MSSP is challenging, evidenced by Chen & Homem-de Mello (2010): "even the continuous relaxation of that problem does not have a concave expected recourse function", as its objective function and constraints involve booking limits truncated by realized demands. As a compromise, they propose an approximation based on re-solving a sequence of two-stage stochastic programs.

We consider the MSSP with continuous relaxation. In each time period, the firm decides the booking limits allocated to each demand class before the demand is realized. Interestingly, our transformation technique preserves concavity in the dynamic programming recursions, and hence overcomes the difficulty stated by Chen & Homem-de Mello (2010). Under certain network structure, we further show that $L^\natural$-concavity can be preserved and use it to derive some monotonicity properties of the optimal booking limits. Our approach opens the door to the development of effective algorithms to solve MSSP directly.

In all the above applications, we employ the transformation technique to prove that the apparently non-convex minimization problems (or non-concave maximization problems) can be converted to equivalent convex minimization problems (or concave maximization problems), and under some conditions, the optimal decisions are monotone in terms of the state variables with limited sensitivities. Without the transformation technique, the structural analyses would have been much more complicated, if not impossible, to carry out.

3.2 Inventory Transshipment with Random Capacities

The purpose of this section is to apply the transformation technique we developed in Chapter 2 to provide a new approach to the analysis of optimal joint inventory and transshipment control under uncertain capacity.

In this section, we consider the model studied in Hu et al. (2008). They provide a characterization of the structure of the optimal transshipment and production policy. For this purpose, they identify several important properties of the profit-to-go functions, which play a pivotal role in the derivation of the structure of the optimal policy. They spent several pages through a very detailed and complicated analysis of derivatives to prove these properties. By employing the concept of $L^\natural$-convexity, or equivalently $L^\natural$-concavity, in this section, we present a simple yet non-trivial proof of those properties. In particular, we realize that these properties of the profit-to-go functions are nothing but natural consequences of $L^\natural$-concave functions after a proper transformation of the original variables. However, to prove that the profit-to-go functions are $L^\natural$-concave (after variable transformation) is not straightforward. In fact, there are two bottlenecks in showing the $L^\natural$-concavity. First, in the transshipment stage, the equality constraint that guarantees the sum of inventory positions at two facilities must remain unchanged prohibits the feasible set to be sublattice. To tackle this difficulty, we apply a recent result by Chen, Hu & He (2013) that deals with parametric optimizations with non-lattice structures. Second, in the production stage, the realized production quantity is the minimum of the production quantity decision and the realized production capacity. As a result, the objective function is not concave in the decision variables. Interestingly, the transformation technique we developed provides a tool to resolve this issue.

In the following, we present the model formulation in detail. Consider a firm operating two manufacturing facilities in separate markets through multiple time periods. Each facility faces uncertain capacities that are independent in time and of each other. Facilities also face uncertain demands which are independent in time but can be correlated across the two facilities. In each period, the firm's decisions can be divided into two stages. The first stage is the production stage where the firm decides how much it will produce in each of the facilities. After the production stage, the capacities and demands are realized. The firm's actual production quantity, which is the minimum of the planned production quantity and the realized capacity, incurs a unit production cost. The firm then enters the transshipment stage where it decides how much inventory to be transshipped from one facility to another. Finally, the demands are met and unsatisfied demands are lost. The firm receives linear revenue on satisfied demands and pays linear holding and transshipment costs. The problem is then to find the optimal production and transshipment quantities in each period so that the firm maximizes the total discounted profit over the planning horizon.

We now introduce the dynamic programming formulation of the optimization problem in Hu et al. (2008) as follows. Let $G_*^k(x_1^k, x_2^k)$ be the profit-to-go function when the current inventory levels at the two facilities are x_1^k and x_2^k respectively and there are k periods left in the planning horizon.

Production Stage:

$$\begin{aligned} G_*^k(x_1^k, x_2^k) = \max_{y_1^k \geq x_1^k, y_2^k \geq x_2^k} & E_{T_1^k, T_2^k, D_1^k, D_2^k}\{-c_1(y_1^k \wedge (x_1^k + T_1^k) - x_1^k) \\ & - c_2(y_2^k \wedge (x_2^k + T_2^k) - x_2^k) + r_1 D_1^k + r_2 D_2^k \\ & + J_*^k(y_1^k \wedge (x_1^k + T_1^k) - D_1^k, y_2^k \wedge (x_2^k + T_2^k) - D_2^k)\} \end{aligned} \tag{3.1}$$

Transshipment Stage:

$$J_*^k(z_1^k, z_2^k) = \max_{\hat{z}_1^k + \hat{z}_2^k = z_1^k + z_2^k} J^k(z_1^k, z_2^k, \hat{z}_1^k, \hat{z}_2^k), \tag{3.2}$$

where

$$\begin{aligned} J^k(z_1^k, z_2^k, \hat{z}_1^k, \hat{z}_2^k) = & -r_1(\hat{z}_1^k)^- - r_2(\hat{z}_2^k)^- - h_1(\hat{z}_1^k)^+ - h_2(\hat{z}_2^k)^+ \\ & -s_1(z_1^k - \hat{z}_1^k)^+ - s_2(z_2^k - \hat{z}_2^k)^+ + \alpha G_*^{k-1}((\hat{z}_1^k)^+, (\hat{z}_2^k)^+), \end{aligned} \tag{3.3}$$

and $G_*^0(x_1^0, x_2^0) \equiv 0$.

In the production stage, in period k, the target inventory levels at the two facilities y_1^k and y_2^k are decided. They are constrained to be no smaller than the current inventory levels at the two facilities x_1^k and x_2^k. The first two terms on the right hand side of (3.1) are the production costs with c_1, c_2 and T_1^k, T_2^k representing the marginal production costs and random capacities at the two facilities respectively. The next two terms are the full revenue collected over the realized demands, where r_1, r_2 and D_1^k, D_2^k are marginal revenue and random demand respectively. The revenue for the lost sales is deducted in the transshipment stage.

In the transshipment stage, in period k, the transshipment quantities or equivalently, the inventory levels after transshipment $\hat{z}_1^k$ and $\hat{z}_2^k$ are decided, whose sum is constrained to be equal to the inventory levels before transshipment (but after demands realization) z_1^k and z_2^k. The first two terms on the right hand side of (3.3) are the deducted revenue for the lost sales. The next two terms are the holding costs, where h_1 and h_2 are unit holding costs at the two facilities respectively. The two terms following are transshipment costs with s_1 (s_2) being the unit transshipment cost from facility 1 (2) to 2 (1). Finally, α in (3.3) is the discount factor.

Hu et al. (2008), under the assumption of continuous demands and capacities, prove the following properties on the profit-to-go function $G_*^k(x_1, x_2)$, which are essential for their derivation of the optimal transshipment and production policies.

$\mathbb{A}_1$: $G_*^{k-1}(x_1, x_2)$ is jointly concave in x_1 and x_2, and

$$\begin{aligned}\frac{\partial^2}{\partial x_1^2} G_*^{k-1}(x_1, x_2) &\leq \frac{\partial^2}{\partial x_1 \partial x_2} G_*^{k-1}(x_1, x_2),\\ \frac{\partial^2}{\partial x_2^2} G_*^{k-1}(x_1, x_2) &\leq \frac{\partial^2}{\partial x_2 \partial x_1} G_*^{k-1}(x_1, x_2);\end{aligned}$$

$\mathbb{A}_2$: $G_*^{k-1}(x_1, x_2)$ is submodular and

$$\frac{\partial^2}{\partial x_1 \partial x_2} G_*^{k-1}(x_1, x_2) = \frac{\partial^2}{\partial x_2 \partial x_1} G_*^{k-1}(x_1, x_2).$$

Through an inductive argument, their proof relies on a full characterization of the optimal transshipment policy and a rather involved analysis of the derivatives which spans several pages. In the following, we present our new approach by using what we have introduced earlier. Interestingly, our approach does not rely on the characterization of the optimal policy and it applies to discrete demands as well as capacities without any further efforts.

Denote d_i^k as the realization of demand for facility i in period k, and define $q_i^k = z_i^k + d_i^k$, $w_i^k = \hat{z}_i^k + d_i^k$. Furthermore, we change variables by letting $\tilde{y}_2^k = -y_2^k$, $\tilde{x}_2^k = -x_2^k$, $\tilde{T}_2^k = -T_2^k$, $\tilde{q}_2^k = -q_2^k$, $\tilde{w}_2^k = -w_2^k$. Then the original problem can be equivalently reformulated as

$$
\begin{aligned}
&\tilde{G}_*^k(x_1^k, \tilde{x}_2^k) = \max_{y_1^k \geq x_1^k, \tilde{y}_2^k \leq \tilde{x}_2^k} E_{T_1^k, \tilde{T}_2^k, D_1^k, D_2^k}\{-c_1(y_1^k \wedge (x_1^k + T_1^k) - x_1^k) \\
&\quad + c_2(\tilde{y}_2^k \vee (\tilde{x}_2^k + \tilde{T}_2^k) - \tilde{x}_2^k) + \tilde{J}_*^k(y_1^k \wedge (x_1^k + T_1^k), \tilde{y}_2^k \vee (\tilde{x}_2^k + \tilde{T}_2^k))\},
\end{aligned} \tag{3.4}
$$

where $\tilde{G}_*^k(x_1^k, \tilde{x}_2^k) = G_*^k(x_1^k, -\tilde{x}_2^k)$ and by introducing a new variable v

$$
\begin{aligned}
&\tilde{J}_*^k(q_1^k, \tilde{q}_2^k) = \max_{w_1^k, \tilde{w}_2^k, v} \tilde{J}(w_1^k, \tilde{w}_2^k, v) \\
&s.t. \quad w_1^k + v = q_1^k \\
&\qquad\quad \tilde{w}_2^k + v = \tilde{q}_2^k,
\end{aligned} \tag{3.5}
$$

where

$$
\begin{aligned}
&\tilde{J}(w_1^k, \tilde{w}_2^k, v) = r_1(w_1^k \wedge d_1^k) + r_2((-\tilde{w}_2^k) \wedge d_2^k) \\
&- h_1(w_1^k - d_1^k)^+ - h_2(-\tilde{w}_2^k - d_2^k)^+ - s_1 v^+ - s_2(-v)^+ \\
&+ \alpha \tilde{G}_*^{k-1}((w_1^k - d_1^k)^+, -(-\tilde{w}_2^k - d_2^k)^+).
\end{aligned} \tag{3.6}
$$

In the following, we introduce a result developed in Chen, Hu & He (2013). It establishes a preservation property of $L^\natural$-concavity under optimization operations when the constraint set may not be a sublattice.

Proposition 3.1. *Consider the following optimization problem parameterized by a two-dimensional vector x:*

$$
f(x) = \max_{y_1, \ldots, y_N} \{\sum_{n=1}^N f_n(y_n) : \sum_{n=1}^N y_n = x, y_n \in S_n, \forall n\},
$$

where S_n are subsets of $\Re^2$, and f is defined on $S = \{\sum_{n=1}^N y_n : y_n \in S_n, \forall n\}$.

If S_n is of the following form:

$$\{(x_1, x_2) \in \Re^2 : l_1 \le x_1 \le u_1, l_2 \le x_2 \le u_2, l_0 \le x_1 - x_2 \le u_0\},$$

and all f_n are $L^\natural$-concave on S_n, then f is $L^\natural$-concave on S.

Now we are ready to state and prove our main result, which offers a new approach that proves the key properties $\mathbb{A}_1$ and $\mathbb{A}_2$ when demands and capacities are continuous .

Theorem 3.1. *Suppose that $\tilde{G}_*^{k-1}(\cdot,\cdot)$ is $L^\natural$-concave, then $\tilde{G}_*^{k}(\cdot,\cdot)$ is also $L^\natural$-concave.*

Using Proposition 2.1 part (h), it is straightforward to check that Theorem 3.1 then implies the properties $\mathbb{A}_1$ and $\mathbb{A}_2$ of $G_*^k(\cdot,\cdot)$ when demands and capacities are continuous. We also point out that the structure of the optimal policies can be derived from Theorem 3.1 for both continuous and discrete demands as well as capacities with some minor modifications of the analysis in Hu et al. (2008).

Our transformation technique provides a novel way to analyze a joint inventory and transshipment control problem with uncertain capacities. Our approach not only significantly simplifies the structural analysis but also can be easily applied to some extensions as we discuss below, which otherwise may require considerable amount of additional effort.

(1) *Backorder case.* In the case of backorder instead of lost sales, we replace the term $\alpha\tilde{G}_*^{k-1}((w_1^k - d_1^k)^+, -(-\tilde{w}_2^k - d_2^k)^+)$ in (3.6) by $\alpha\tilde{G}_*^{k-1}((w_1^k - d_1^k), -(-\tilde{w}_2^k - d_2^k))$, which is still $L^\natural$-concave by induction hypothesis. Similarly, it is easy to show that adding shortage cost does not change the $L^\natural$-concavity in (3.6). Thus, Theorem 3.1 holds in this case.

(2) *Capacities on the transshipment quantities.* In many practical scenarios, a firm may not have the luxury to transship any arbitrary large amount of quantities from one facility to another because for instance, it has only a few fleet vehicles. In some settings, the transshipment can be restricted to a single direction, i.e., one of the transshipment capacity is zero. Let S_1 (S_2) be the capacity on the transshipment quantities from facility 1 (2) to facility

2 (1). Then problem (3.5) is now reformulated as

$$\begin{aligned}
&\tilde{J}^k_*(q_1, \tilde{q}_2) = \max_{w_1, \tilde{w}_2, v} \tilde{J}(w_1, \tilde{w}_2, v) \\
&s.t. \quad w_1 + v = q_1, \\
&\qquad \tilde{w}_2 + v = \tilde{q}_2, \\
&\quad -S_2 \le v \le S_1.
\end{aligned}$$

It is straightforward to check that Proposition 3.1 still applies and consequently our conclusion still holds.

3.3 Dual Sourcing with Supply Capacity Uncertainty

Consider a firm managing a T-period periodic-review inventory system in the presence of two capacitated suppliers (or delivery modes): a regular supplier with a longer replenishment leadtime of l_R periods and a unit ordering cost c_R, and an expedited (emergency) supplier with a shorter replenishment leadtime of l_E periods and a unit ordering cost c_E, where l_R and l_E are nonnegative integers and $l_R > l_E$. There are no fixed ordering costs. Both suppliers offer limited and uncertain capacities, denoted by $K_{R,t}$ and $K_{E,t}$, $t \in \{1, ..., T\}$, for regular and expedited suppliers, respectively. The processes $\{K_{R,t}\}_{t=1}^T$ and $\{K_{E,t}\}_{t=1}^T$ are both independent over time and independent of each other. Note that the independence assumption on the supply capacity distributions can be justified by the dual sourcing practice with two geographically distant locations, such as China and Mexico in the case study of Van Mieghem (2008), where the production processes are typically independent of each other. Demands of successive periods, denoted by D_t for period t, are stochastic, independent over time, and independent of the supply capacities. For convenience, let $D_{[t,t+l]}$ be the total demand from period t to period $t+l$, i.e., $D_{[t,t+l]} = D_t + ... + D_{t+l}$. We use d_t and $d_{[t,t+l]}$ to denote the realization of D_t and $D_{[t,t+l]}$.

It is notable that a typical assumption in the dual-sourcing literature without capacity limits is that the expedited ordering cost c_E is greater than the regular ordering cost c_R, because otherwise it is trivial for the firm to procure exclusively from the expedited supplier (see, e.g., Veeraraghavan and Scheller-Wolf 2006, Sheopuri et al. 2010). We do not make this assumption here. In fact, if the expedited capacity is limited, even when the regular ordering cost is higher, it may still be beneficial to order from the regular supplier.

The sequence of events is as follows. At the beginning of period t, orders from the regular supplier l_R periods ago and the expedited supplier l_E periods ago (if $l_E \geq 1$) are received. (Note that if $l_E = 0$, we assume that an order from the expedited supplier is received right away.) The firm then reviews the inventory level and the orders outstanding, and determines how much to order from the two suppliers before observing the suppliers' capacities $K_{R,t}$ and $K_{E,t}$. Let q_R and q_E be the (target) order quantities from the regular and expedited channels, respectively. After the orders are placed, the suppliers' capacities $K_{R,t}$ and $K_{E,t}$ are realized. We use $k_{R,t}$ and $k_{E,t}$ to denote realizations of $K_{R,t}$ and $K_{E,t}$ respectively. Then the amounts of inventories shipped from the regular and expedited suppliers are $q_R \wedge k_{R,t}$ and $q_E \wedge k_{E,t}$, respectively. Note that here we assume that the supply capacity uncertainties are resolved in the same period when the orders are placed (see Federgruen and Yang 2011 for a similar treatment for the random yield problem). This is reasonable when the capacity uncertainties are mainly driven by the unreliability of the production process and the production time is no more than the period length while the shipping time is long. The ordering costs are given by $c_R(q_R \wedge k_{R,t})$ and $c_E(q_E \wedge k_{E,t})$. Here we assume that the ordering cost is proportional to the quantity actually delivered, which is a common assumption in the literature of inventory control with random capacities (see Ciarallo et al. 1994, Wang and Gerchak 1996, Feng 2010, and so on). This assumption is appropriate when the payment is made upon the receipt of the shipments and the firms only pay the actual delivered amount. At the end of this period, the demand is realized and met with on-hand inventory (if any). Unmet demand is fully backlogged with a unit shortage cost h^-. Excess inventory is carried over to the next period with a unit holding cost h^+.

The objective of the firm is to find a dual-sourcing strategy so as to minimize the total expected discounted cost, including ordering cost, holding cost and backorder cost, over the planning horizon. To present the dynamic programming model for deriving the optimal strategy, one can naturally describe the system state right before the firm places orders by a vector $s = (s_0, ..., s_{l_R-1})$, where s_i denotes the amount of on-hand net inventory plus outstanding orders that will arrive within i periods, $i = 1, .., l_R - 1$. However, in a backlogging model, since the orders of each period will have an influence only l_E periods later, and the on-hand net inventory level l_E periods later solely depends on s_{l_E}, it suffices to use the now standard accounting technique to discount the future inventory cost to the current period and focus on the pipeline inventory levels $s_{l_E}, ..., s_{l_R-1}$. Specifically, we can reduce the state space to $k = l_R - l_E$ dimensions by defining the system state as $z = (z_1, ..., z_k)$, where $z_i = s_{i+l_E-1}, i = 1, ..., k$. The state space is given by

$$\mathcal{S} = \{(z_1, \ldots, z_k) : z_1 \leq z_2 \leq \ldots \leq z_k\}.$$

Given the system state z, the system state of the next period is given by

$$\tilde{z} = (z_2 + q_E \wedge k_{E,t} - d_t, ..., z_k + q_E \wedge k_{E,t} - d_t, y \wedge (z_k + k_{R,t}) + q_E \wedge k_{E,t} - d_t),$$

where $y = z_k + q_R$ is the (target) order-up-to level from the regular channel. For reasons that will become clear later, we denote $u = -q_E$ and $\tilde{k}_{E,t} = -k_{E,t}$. The dynamics of the system state can be rewritten as

$$\tilde{z} = [(z_2, ..., z_k, y \wedge (z_k + k_{R,t})) - (u \vee \tilde{k}_{E,t} + d_t)e],$$

where e is the k-dimensional all-ones vector.

We are now ready to present the dynamic program to derive the firm's optimal strategy. Let $\alpha \in (0, 1]$ be the discount factor. The optimality equations can be written as follows. For $t = 1, ..., T$,

$$v_t(z) = \min_{y \geq z_k, u \leq 0} \left\{ E[g_t(z, y \wedge (z_k + K_{R,t}), u \vee \tilde{K}_{E,t})] \right\} \ \forall \ z \in \mathcal{S}, \tag{3.7}$$

where

$$g_t(z, y, u) = c_R(y - z_k) - c_E u + B_t(z_1 - u) + \alpha E[v_{t+1}((z_2, ..., z_k, y) - (d_t + u)e)], \tag{3.8}$$

and

$$B_t(x) = \alpha^{l_E} E[h^+(x - d_{[t,t+l_E]})^+ + h^-(d_{[t,t+l_E]} - x)^+].$$

Note that the expectation of the right hand side of equation (3.7) is taken over the random capacities. The function g_t represents the expected total discounted cost after the capacities are realized but before the demand is realized. The first term of the right hand side of equation (3.8) is the ordering cost from the regular supplier, the second term is the ordering cost from the expedited supplier, the third term is the expected discounted holding and shortage cost of period $t + l_E$, and the last term is the expected total discounted future costs. For simplicity, we assume the terminal value function $v_{T+1}(z) = 0$ for any z, which implies that there is no salvage value for leftover inventory and no backlogging cost for unfilled demand after period $T + l_E$. That is, the firm makes decisions in the first T periods but takes into account the inventory cost up to period $T + l_E$. Our structural results and analysis still hold if $v_{T+1}(z)$ is assumed to be $L^\natural$-convex.

Problem (3.7) admits optimal solutions under rather general and standard conditions. Nevertheless, it is a challenging problem. First, the state space is multi-dimensional. A more severe issue is that the objective function of problem (3.7) is not convex. Note that for the last period with $v_{T+1} = 0$, the objective function has a structure similar to that in (2.1), which may not be convex. Thus, it is far from being clear whether the cost-to-go functions v_t are convex, and even if they are, the objective function of problem (3.7) is not. However, with our transformation technique we can convert the non-convex minimization problem (3.7) into an equivalent convex minimization problem and show that v_t is actually $L^\natural$-convex.

In the following analysis, we assume that both c_E and c_R are smaller than $h^-/(1-\alpha)$, which ensures that it is not optimal to never order anything and merely accumulate penalty costs. Let $(y_t(z), u_t(z))$ denote the optimal solution for problem (3.7). When there are multiple optimal solutions, we assume it is the greatest one, which will be shown to be well defined later.

Theorem 3.2. *For all t, $v_t(z)$ is $L^\natural$-convex in $z \in \mathcal{S}$. The optimal solution $(y_t(z), u_t(z))$ is increasing in z with limited sensitivity. (When there are multiple optimal solutions, we assume it is the greatest one.) That is, for any $\omega > 0$,*

$$y_t(z) \le y_t(z+\omega e) \le y_t(z) + \omega, \quad u_t(z) \le u_t(z+\omega e) \le u_t(z) + \omega \tag{3.9}$$

The monotonicity and limited sensitivity of $y_t(z)$ imply that the optimal regular order quantity $q_{R,t}(z)$, which is equal to $y_t(z) - z_k$, increases in $z_1, ..., z_{k-1}$, but decreases in z_k and satisfies $-\omega \le q_{R,t}(z+\omega e) - q_{R,t}(z) \le 0$.

To gain more insights, we can transform the state vector to $x = (x_1, ..., x_k)$

where $x_1 = z_1$ and $x_i = z_i - z_{i-1}$, $i = 2, ..., k$. Note that $x_1 = z_1$ represents the amount of on-hand net inventory plus outstanding orders that will arrive within l_E periods, and x_i represents the size of the outstanding order that will arrive $l_E + i - 1$ periods later. Denote the corresponding optimal order quantities by $\hat{q}_{R,t}(x) = q_{R,t}(z)$ and $\hat{q}_{E,t}(x) = q_{E,t}(z)$. The monotonicity and limited sensitivity of $y_t(z)$ imply the following inequalities.

$$\begin{aligned} & -\omega \leq \hat{q}_{R,t}(x + \omega e_k) - \hat{q}_{R,t}(x) \leq \hat{q}_{R,t}(x + \omega e_{k-1}) - \hat{q}_{R,t}(x) \leq ... \\ & \leq \hat{q}_{R,t}(x + \omega e_1) - \hat{q}_{R,t}(x) \leq 0. \end{aligned} \tag{3.10}$$

The derivation of above equalities are shown in Appendix B. Compare states $z + \omega e_i$ and z. For $i = 1$, the former has ω more units of on-hand inventory or outstanding orders that will arrive within l_E periods. For $i = 2, ..., k$, the former has ω more units of outstanding order that will arrive $l_E + i - 1$ periods later. Thus, inequalities (3.10) imply that the regular order quantity decreases in on-hand inventory level and the sizes of the outstanding orders. The sensitivity decreases in the age of the outstanding order, where the age refers to the number of periods passed since the order was placed. In other words, the regular order quantity is most sensitive to the size of the most recently placed order.

Similarly, for the expedited order quantity $\hat{q}_{E,t}(x) = -u_t(z)$, we have

$$\begin{aligned} & -\omega \leq \hat{q}_{E,t}(x + \omega e_1) - \hat{q}_{E,t}(x) \leq \hat{q}_{E,t}(x + \omega e_2) - \hat{q}_{E,t}(x) \leq ... \\ & \leq \hat{q}_{E,t}(x + \omega e_k) - \hat{q}_{E,t}(x) \leq 0. \end{aligned} \tag{3.11}$$

That is, the expedited order quantity decreases in the sizes of outstanding orders in the pipeline, but the sensitivity increases in the age of the outstanding order. In other words, the expedited order quantity is least sensitive to the most recently placed order, which is opposite to the sensitivity of the regular order quantity. and in the joint inventory-pricing control problem-s with positive leadtime where the replenishment decision has a decreasing sensitivity in the age of the outstanding order whereas the pricing decision has an increasing sensitivity in the age of the outstanding order (see, e.g., Chen et al. 2014). The implication is that the decisions whose immediate impacts are closer to the on-hand stock (e.g., pricing or expedited order) are more sensitive to the on-hand inventory level and older outstanding orders while the decisions whose immediate impacts are further away from the on-hand stock (e.g., regular order) is more sensitive to the younger outstanding orders.

3.4 Assemble-to-Order Systems with Random Capacities

Consider an assemble-to-order (ATO) system over a planning horizon with T periods. The ATO system consists of m components indexed by $i \in \{1, 2, ..., m\}$ and n products indexed by $j \in \{1, 2, ..., n\}$. At the beginning of each period, the firm observes on-hand inventory levels of the m components $x = (x_1, ..., x_m)^T$, and then decides the order-up-to inventory levels of components $y = (y_1, ..., y_m)^T$. The delivered quantity of each component i cannot exceed a random capacity, denoted by $\Xi_{t,i}$, which is realized after the order is placed. The capacities are independent of each other and over time. Inventory replenishment leadtime is assumed to be zero. The demand for product j in period t is $D_{t,j}$ and we assume that they are independent over time and independent of capacities. Let $D_t = (D_{t,1}, ..., D_{t,n})^T$. The bill of materials is specified by an $m \times n$ matrix A, whose component a_{ij} denotes the units of component i required to make one unit of product j. Unmet demands are assumed to be lost. Let c_i and h_i represent the ordering cost and holding cost of component i per unit respectively, and b_j denote the per unit shortage cost of product j. We use c, h, b to denote the vectors $(c_1, ..., c_m)^T$, $(h_1, ..., h_m)^T$, $(b_1, ..., b_n)^T$ respectively. The one-period discount factor is $\alpha \in (0, 1]$. The objective of the firm is to minimize the total expected discounted cost.

Let $f_t(x)$ be the cost-to-go function with initial inventory levels x at the beginning of period t. We omit the subscript t for notational brevity when no ambiguity occurs. The optimality equation is

$$f_t(x) = \min_{y \geq x} \{E[c^T(y \wedge (x + \Xi) - x)] + E[g_t(y \wedge (x + \Xi)|D)]\}, \tag{3.12}$$

where

$$g_t(z|d) = \min_{u:(z,u) \in \mathcal{U}(d)} \{\mathcal{L}(z, u|d) + \alpha f_{t+1}(z - Au)\}, \tag{3.13}$$

and

$$\mathcal{L}(z, u|d) = h^T(z - Au) + b^T(d - u). \tag{3.14}$$

The boundary condition is assumed to be $f_{T+1}(x) = 0$ without loss of generality. The first term in the objective function of (3.12) is the ordering cost. Similar to the dual sourcing model, we assume that the ordering cost is proportional to the quantity actually delivered. The feasible set in (3.13) is given by $\mathcal{U}(d) = \{(z, u)|Au \leq z, 0 \leq u_j \leq d_j, j = 1, 2, ..., n\}$, where z is the on-hand inventory level after the inventory ordered in the current period arrives, and u is the vector of assembled-product quantities. The inventory holding and shortage costs are given in $\mathcal{L}(z, u|d)$.

Due to the complexity of general ATO systems, some important special systems are studied in the literature, one of which is a generalized M-system (see Nadar et al. 2014). A generalized M-system has m components and $m + 1$ products, where each product i requires a single unit of component i for $i \leq m$ and product $m + 1$ consumes one unit of each component. This ATO system reduces to an M-system when $m = 2$. The bill of materials matrix has the following form:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}. \tag{3.15}$$

We summarize the structural results of this section in the following theorem.

Theorem 3.3. *(a) For a general ATO system, the optimal cost function $f_t(x)$ is convex in x for all t.*

(b) For a generalized M-system, the optimal cost function $f_t(x)$ is $L^\natural$-convex in x for all t. The optimal order-up-to level $y_t(x)$ is increasing in x with limited sensitivity. That is, for any $\omega > 0$, $y_t(x) \leq y_t(x + \omega e) \leq y_t(x) + \omega e$. (When there are multiple optimal solutions, we assume it is the greatest one.)

Theorem 3.3 summarizes the sensitivity results for the stage when components are ordered. For a generalized M-system, the order-up-to level y_i of any component i increases in its own inventory level x_i as well as the inventory level of any other component $j \neq i$. The limited sensitivity implies that for any component i the ordering quantity $y_i - x_i$ decreases in the inventory level of any component. One can easily check that the following sensitivity results hold during the stage when products are assembled. For a generalized M-system, the quantity of product $m+1$ increases in the quantity of each component, while the quantity of product $j(\neq m+1)$ increases in the quantity of component j but decreases in the quantity of component $k(\neq j)$.

3.5 Network Revenue Management Using Booking Limits

We consider a network system consisting of m resources (airline seats in different legs), indexed by $i \in \{1, ..., m\}$, with initial capacity levels $C = (C_1, ..., C_m)^T$, and n products (itinerary-class combinations), indexed by $j \in \{1, ..., n\}$. The corresponding prices, denoted by $p = (p_1, ..., p_n)^T$, are exogenously given. Each product needs at most one unit of each resource. Let $A = (a_{ij})$ be the resource coefficient matrix, where $a_{ij} = 1$ if product j uses one unit of resource i and $a_{ij} = 0$ otherwise. Define $D_t = (D_{t,1}, ..., D_{t,n})^T$ where $D_{t,j}$ is demand of product j in period t. Assume that the demands of different products are independent and the demands are independent over time. The objective of the firm is to decide the booking limits for all demand classes dynamically so as to maximize the total expected profit over the planning horizon.

As mentioned earlier, the model we consider here is MSSP in Chen & Homem-de Mello (2010) with continuous relaxations. Chen & Homem-de Mello (2010) point out that the major difficulty of the above model is that it is not a concave maximization problem, since the decisions are truncated by random demands. Therefore, they re-solve a sequence of two-stage stochastic programs for approximation. Interestingly, as we show in this section, our transformation technique can overcome this difficulty and allows us to preserve concavity in the dynamic programming recursions. Under certain network structure, we further demonstrate that $L^\natural$-concavity can be preserved and use it to derive monotone properties of the optimal booking limits. Note that the model considered here is different from the one in section 3.2.1 of Talluri & Van Ryzin (2005). Their model assumes there is at most one demand request in any period. We do not impose this assumption. Since in our model each time period corresponds to the time when the firm needs to revise its capacity allocation policy, it may not be practical to divide the planning horizon so much so that there is at most one demand in any period due to the increased computational complexity.

In the following, we omit the subscript t for notational brevity when no ambiguity occurs. The state variable is denoted by the vector $x = (x_1, ..., x_m)^T$ in which x_i is the capacity level of the resource i in the current period. At the beginning of the planning horizon, we have $x = C$. In each period, the firm observes the current capacity level x and decides the booking limits for different demand classes. The decision variable is denoted by vector $u = (u_1, ..., u_n)^T$ where u_j is the booking limit for class j demand in the current period. The action space can be defined as $\mathcal{A} = \{(x, u)|Au \leq x, u \geq 0\}$. Let $f_t(x)$ be the optimal value. The optimality equations can be expressed as

$$f_t(x) = \max_{u:(x,u)\in\mathcal{A}} E\left[p^T(u \wedge D) + f_{t+1}\left(x - A(u \wedge D)\right)\right], t = 1, ..., T, \quad (3.16)$$

where $f_{T+1}(x) = 0$. For $\xi \in \mathcal{F}_+^m$, define the function $g_t : \mathcal{F}_+^{m+1} \to \Re$ such that

$$g_t(x, \xi) = p^T\xi + f_{t+1}(x - A\xi).$$

Then the optimality equation can be expressed as

$$f_t(x) = \max_{u:(x,u)\in\mathcal{A}} E\left[g_t(x, u \wedge D)\right], t = 1, ..., T. \quad (3.17)$$

We also consider a special case where the resource coefficient matrix has the same format as the bill of materials matrix in the assemble-to-order generalized M-system, i.e., the resource coefficient matrix is given by (3.15). When the number of resources $m = 2$, one can relate this type of resource coefficient matrix to the following setting. There are two legs in the network: A to B and B to C. There are three types of consumers. Type one consumers travel from A to B, type two consumers travel from B to C, and type three consumers travel from A to C with a transition at B.

We summarize the structural results in the following theorem.

Theorem 3.4. *(a) For the network revenue management problem* (3.16)*, the optimal value function* $f_t(x)$ *is concave in* x *for all* t*.*

(b) If, in addition, the resource coefficient matrix is given by (3.15)*, then for all* t*,* $f_t(x)$ *is* $L^\natural$*-concave. The optimal booking limit* $u^*_{m+1}(x)$ *is increasing in* x *with limited sensitivity, i.e., for any* $\omega > 0$*,* $u^*_{m+1}(x) \leq u^*_{m+1}(x + \omega e) \leq u^*_{m+1}(x) + \omega$*. For* $j = 1, ..., m$*,* $u^*_j(x)$ *is increasing in* x_j *and decreasing in* $x_k, k \neq j$*, with limited sensitivity, i.e.,* $u^*_j(x) \leq u^*_j(x + \omega e_j) \leq u^*_j(x) + \omega$ *and* $u^*_j(x) - \omega \leq u^*_j(x + \omega e_k) \leq u^*_j(x)$ *for any* $\omega > 0$*,* $k \neq j$*. (When there are multiple optimal solutions, we choose the one such that* $(-u^*_1(x), ..., -u^*_m(x), u^*_{m+1}(x))$ *is the greatest.)*

The sensitivity result from Theorem 3.4 implies that if the current capacity level of any resource i increases by ω, then the allocated capacity of product i and $m + 1$ should also increase, but the allocated capacity of product $j, j \neq i, j \neq m + 1$ will decrease. All the above changes are bounded by ω because of the limited sensitivity.

Remark 3.1. *Even though the resource coefficient matrix here is the same as the bill of materials matrix in Section 3.4, the analyses of the two models have a significant difference. For the ATO model, the decision variable is truncated by random capacity and the bill of materials matrix does not enter the constraints when we apply the transformation technique. However, for the revenue management model the decision variable is truncated by random demand and the resource coefficient matrix affects the constraints when applying the transformation.*

Chapter 4

Generalizations

4.1 Introduction

In this chapter, we are going to generalize the transformation technique developed earlier in several directions.

First of all, we allow a more general objective function. Take the inventory problem with random capacities as an example, the objective function considered in this chapter allows the ordering cost to depend both on the initial ordering quantity and the quantity actually received. Notice that several papers studying inventory management problems with random yield (for example Henig & Gerchak 1990 and Federgruen & Yang 2011) the firm's ordering cost depends on the quantity actually received, as well as the quantity initially ordered. This general cost structure is necessary if the production is in-house, or the cost consequence of supply uncertainty is shared between the supplier and the firm who places the order. Interestingly, all papers we reviewed in earlier chapters which study inventory models with random capacities assume that the ordering costs only depends on the quantity actually received. To the best of our knowledge, we are the first to allow the ordering costs to depend on both the quantity received and the quantity initially ordered in a random capacity model.

Secondly, we consider the case where the random vector can have dependent components. In particular, we focus on the case where the random variables have "positive dependence". In an inventory system where the firm can order from multiple suppliers or produce in multiple facilities, it is common that the supply capacities are influenced by common factors. For instance, agricultural products in the same region are affected by local weather or natural disasters; The suppliers may share a common second tier supplier or they may import materials from the same country which has export restrictions.

Last but not least, we do not restrict to risk-neutral decision makers and incorporate risk attitude into the model. For studies on inventory management models with risk attitude, see Chen, Sim, Simchi-Levi & Sun (2007) and the references therein. To model the decision maker's risk measure, this chapter considers a commonly used risk measure called Conditional Value-at-risk (CVaR). Studies that address inventory models using CVaR as the risk measure include Ahmed et al. (2007) and Chen et al. (2009), among others. Based on the results under CVaR criterion, we extend results to a more general class of risk measure called distortion risk measure.

4.2 More General Objective Functions, Dependent Random Variables

Consider the following optimization problem

$$\tau^* = \inf_{u \in \mathcal{F}^n} l(u) + E[f(u \wedge \Xi)], \tag{4.1}$$

where $l : \mathcal{F}^n \to \bar{\Re}, f : \mathcal{F}^n \to \bar{\Re}$, and Ξ is a random vector with support $\mathcal{X} \subseteq \mathcal{F}^n$. One may associate problem (4.1) with an inventory management problem with random supply capacities. The firm wants to minimize the cost by choosing the ordering quantities before the random capacities are realized. The effective inventory level after receiving the orders is the minimum of the ordering quantities and the realized supply capacities. For now, we assume that the decision maker is risk-neutral. Therefore, the objective function simply takes expectation of the costs incurred with different realizations of random capacities. Later we will extend to the case where the decision maker is not risk-neutral, and thus minimizes some risk measure.

Different from problem (2.1) introduced in Chapter 2, we include the cost term $l(u)$ in the objective function. This term allows a more general cost structure. For example, the ordering cost may depend on the quantity actually received as well as the quantity initially ordered. Moreover, we do not need to assume that the random vector Ξ has independent components. One technical challenge of problem (4.1) is that even though the function $l(\cdot)$ and $f(\cdot)$ are jointly convex, the objective function may not be convex in u. Therefore, the main purpose of this section is to develop a transformation technique to convert the original problem to an equivalent convex minimization problem.

Our transformation technique requires that the components of random vector Ξ have some specific dependence structure. We begin by introducing some definitions.

Definition 4.1 (Topkis 1998)**.** *Let $\{F_t(w) : t \in T\}$ be a collection of distribution functions on $\Re^n$ that are indexed by a parameter t, with t contained in a subset T of $\Re^m$. If $\int h(w)dF_t(w)$ is increasing in t on T for each increasing real-valued function $h(w)$ on $\Re^n$, then $F_t(w)$ is stochastically increasing in t on T.*

Given some random variables $\{\Xi_1, \Xi_2, ..., \Xi_n\}$ which are dependent of each other, we define that they have "positive dependence" as follows.

Definition 4.2. *Let $\tilde{F}_{\xi_i}(w)$ be the joint distribution of $\Xi_1, ..., \Xi_{i-1}, \Xi_{i+1}, ...\Xi_n$ conditioned on $\Xi_i = \xi_i$. We define that $\{\Xi_1, ..., \Xi_{i-1}, \Xi_{i+1}, ...\Xi_n|\Xi_i\}$ is stochastically increasing if $\{\tilde{F}_{\xi_i}(w) : \xi_i \in \mathcal{X}_i\}$ is a collection of stochastically increasing functions. And $\{\Xi_1, \Xi_2, ..., \Xi_n\}$ have positive dependence if $\{\Xi_1, ..., \Xi_{i-1}, \Xi_{i+1}, ...\Xi_n|\Xi_i\}$ is stochastically increasing for all $i = 1, ..., n$.*

It is well-known that a collection of distribution functions $\{F_t(w) : t \in T\}$ on $\Re^1$ is stochastically increasing in t on a subset T of $\Re^m$ if and only if $1 - F_t(w)$ is increasing in t on T for each w in $\Re^1$. Therefore, if there are only two random variables Ξ_1 and Ξ_2, then $\{\Xi_2|\Xi_1\}$ is stochastically increasing if and only if

$$Pr(\Xi_2 > \xi_2|\Xi_1 = \xi_1) \text{ is increasing in } \xi_1 \ \forall \xi_2.$$

Feng et al. (2015) consider a joint inventory and pricing model with dependent random supply capacities, and show that when the random capacities are positively dependent, the optimal value function in their problem is concave in the state variables. Similar concepts are also used in Li et al. (2013) when studying a supply diversification problem with responsive pricing. We are now ready to present our transformation technique.

Theorem 4.1. *Suppose that (a) the objective function of* (4.1) *is lower semicontinuous and goes to ∞ when $|u| \to \infty$. (b) f is componentwise convex* *(componentwise discrete convex if $\mathcal{F} = \mathcal{Z}$) and supermodular, $l(u)$ is increasing; (c) $\{\Xi_1, \Xi_2, ..., \Xi_n\}$ are positively dependent. Then problem* (4.1) *has the same optimal objective value as the following problem:*

$$\begin{aligned} \inf \quad & l(u) + E[f(v(\Xi))] \\ s.t. \quad & v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{F}^n. \\ & v(\xi) \leq \xi \ \forall \xi \in \mathcal{X}, \\ & v(\xi) \leq u \ \forall \xi \in \mathcal{X}, \\ & v_j(\xi_j) \ \textit{is increasing} \ \forall \xi_j \in \mathcal{X}_j, \ \ j = 1, ...n \end{aligned} \tag{4.2}$$

Given the above theorem, if $l(\cdot)$ and $f(\cdot)$ are jointly convex, then the transformed problem 4.2 is a convex minimization problem. Notice that when Ξ has independent components, condition (c) still holds and thus the transformation applies.

Comparing Theorem 4.1 with Theorem 2.1, there are several differences. First of all, In the transformed problem (4.2), we require that $v(\xi) \leq u \ \forall \xi \in \mathcal{X}$. This is due to the additional term $l(u)$ in the objective function. In addition, $v_j(\xi_j)$ need to be increasing. This is caused by the dependence structure of Ξ.

Next we present an example showing that the monotone constraints are needed when the random variables are dependent.

Example 4.1. *Let $l(u) = 0$. We write f in standard quadratic form*

$$f(x) = \frac{1}{2} x^T H x + c^T x,$$

where $c^T = [-8, -2]$ and

$$H = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

The function f is jointly convex since H is positive definite, and is supermodular since the cross partial is positive. The random variable Ξ_1 can take values $a_1 = 0$ or $b_1 = 2$, Ξ_2 can take values $a_2 = 1$ or $b_2 = 3$. The probability matrix is

$$\begin{bmatrix} P(\Xi_1 = a_1, \Xi_2 = a_2) = 0.4, & P(\Xi_1 = a_1, \Xi_2 = b_2) = 0.2 \\ P(\Xi_1 = b_1, \Xi_2 = a_2) = 0.1, & P(\Xi_1 = b_1, \Xi_2 = b_2) = 0.3 \end{bmatrix}.$$

Note that Ξ_1 and Ξ_2 are positively dependent. In this case, "positively depen-

dent" requires

$$P(\Xi_2 = b_1 | \Xi_1 = a_2) \leq P(\Xi_2 = b_1 | \Xi_2 = b_2),$$

and

$$P(\Xi_2 = b_2 | \Xi_1 = a_1) \leq P(\Xi_2 = b_2 | \Xi_1 = b_1).$$

The optimal solution of the original problem (4.1) *is* $u_1^* = \frac{35}{18}, u_2^* = \frac{1}{9}$ *with optimal objective value* -3.222. *But the optimal solution of the transformed problem* (4.2) *is* $v_1(a_1) = 0, v_1(b_1) = 2, v_2(a_2) = 0.3, v_2(b_2) = -0.1$ *with optimal objective value* -3.3. *Therefore, without the monotone constraints, the transformed problem cannot generate the same optimal value as the original problem.*

However, if we impose the monotonic constraints on $v(\xi)$, *that is,* $v_1(a_1) \leq v_1(b_1), v_2(a_2) \leq v_2(b_2)$, *we have* $u_1^* = \frac{35}{18}, u_2^* = \frac{1}{9}$, *and* $v_1(a_1) = 0, v_1(b_1) = \frac{35}{18}, v_2(a_2) = \frac{1}{9}, v_2(b_2) = \frac{1}{9}$. *One can easily check that the original problem has the same optimal value as the transformed one after adding the monotone constraints.*

We also show an example that if the objective function is not supermodular (submodular in this example), then after the transformation the optimal value may change.

Example 4.2. *Similar to Example 4.1,* $l(u) = 0$,

$$f(x) = \frac{1}{2} x^T H x + c^T x,$$

where $c^T = [-8, -2]$ *and*

$$H = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}.$$

Ξ_1 *can take values* $a_1 = 0$ *or* $b_1 = 2$, Ξ_2 *can take values* $a_2 = 1$ *or* $b_2 = 3$. *The probability matrix is*

$$\begin{bmatrix} P(\Xi_1 = a_1, \Xi_2 = a_2) = 0.4, & P(\Xi_1 = a_1, \Xi_2 = b_2) = 0.2 \\ P(\Xi_1 = b_1, \Xi_2 = a_2) = 0.1, & P(\Xi_1 = b_1, \Xi_2 = b_2) = 0.3 \end{bmatrix}.$$

The optimal solution of the original problem is $u_1^* = 2, u_2^* = 0.9$, *optimal* *objective value of original problem is -4.82. For the transformed problem,* $v_1(a_1) = 0, v_1(b_1) = 2, v_2(a_2) = 0.7, v_2(b_2) = 1.1$*, and the optimal objective value is* -4.9*.*

In the following we show an example that if the random variables do not have positive dependence, then after the transformation the optimal value may change.

Example 4.3. *Similar to Example 4.1, but the probability matrix changes to*

$$\left[\begin{array}{cc} P(\Xi_1 = a_1, \Xi_2 = a_2) = 0, & P(\Xi_1 = a_1, \Xi_2 = b_2) = 0.5 \\ P(\Xi_1 = b_1, \Xi_2 = a_2) = 0.5, & P(\Xi_1 = b_1, \Xi_2 = b_2) = 0 \end{array}\right].$$

Then the optimal solution of the original problem is $u_1^* = 2, u_2^* = 0$*, and the optimal objective value is* -4*. However, for the transformed problem the optimal solution is* $v_1(a_1) = 0, v_1(b_1) = 2, v_2(a_2) = -0.5, v_2(b_2) = 0.5$*, and the optimal objective value is* -4.5*.*

Similar to the constrained case in Chapter 2, we can explicitly incorporate constraints on u in Theorem 4.1 and consider a more general optimization model. The problem of interest is

$$\inf_{u \in \mathcal{U}} l(u) + E[f(u \wedge \Xi)], \tag{4.3}$$

where $f : \mathcal{F}^n \to \bar{\Re}$ and $\mathcal{U} \subseteq \mathcal{F}^n$. Define a set

$$\mathcal{V} = \{u \wedge \xi : u \in \mathcal{U}, \xi \in \mathcal{X}\}. \tag{4.4}$$

The following assumption identifies a condition under which the equivalent transformation in Theorem 4.1 can be generalized to constrained optimization problems.

Assumption 4.1. *for any* $u \in \mathcal{F}^n$ *such that* $u \wedge \xi \in \mathcal{V}\ \forall \xi \in \mathcal{X}$*, there exists* $u' \in \mathcal{U}, u' \leq u$ *such that* $u' \wedge \xi = u \wedge \xi\ \forall \xi \in \mathcal{X}$*. The indicator function of* $\mathcal{V}$ *is componentwise convex and supermodular.*

Notice that we require $u' \leq u$ in the above assumption, which is different from Assumption 2.1. This is due to the additional term $l(u)$ in the objective function. In other words, if $l(u)$ is a constant, then we can remove the requirement $u' \leq u$ in the above assumption. The indicator function of $\mathcal{V}$ is supermodular is equivalent to that $\mathcal{V}$ is a lattice.

The following theorem demonstrates the equivalent transformation for constrained optimization problems.

Theorem 4.2. *Consider the optimization problem (4.3), where $f : \mathcal{F}^n \to \bar{\Re}, l : \mathcal{F}^n \to \bar{\Re}$ and the random vector Ξ in $\mathcal{F}^n$ satisfy the assumptions in Theorem 4.1. Suppose that Assumption 4.1 is satisfied, then problem (4.3) and the following optimization problem have the same optimal objective value.*

$$\begin{array}{ll} \inf & l(u) + E[f(v(\Xi)] \\ s.t. & v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{V} \ \forall \xi \in \mathcal{X}, \\ & v(\xi) \leq \xi \ \forall \xi \in \mathcal{X}, \\ & v(\xi) \leq u \ \forall \xi \in \mathcal{X}, \\ & v_j(\xi_j) \ \textit{is increasing} \ \forall \xi_j \in \mathcal{X}_j, \ j = 1, ...n. \end{array} \tag{4.5}$$

4.3 Incorporating Risk Measure

In problem (4.3) the decision maker's objective is to minimize the expected cost (or equivalently to maximize the expected profit). This is under the assumption that the decision maker is risk-neutral. However, evidently not all decision makers are risk-neutral in real life, and many are willing to sacrifice some expected profits for lower risks. To capture the decision maker's risk attitude, it is suitable to incorporate risk measure into the model. Introduced by Rockafellar & Uryasev (2000), Conditional Value-at-risk (CVaR) is a commonly used risk measure in practice. There are a number of studies that address operations management problems using CVaR (see Chen et al. 2009 and the references therein).

CVaR of a random variable with confidence level α is defined as the mean of the generalized $\alpha-$tail distribution. In the following we present an equivalent definition through a convex optimization problem which is more convenient to work on:

$$CVaR_\alpha(X) = \inf_{\lambda \in \Re} \{\lambda + \frac{1}{1-\alpha} E[(x - \lambda)^+]\},$$

where $\alpha \in [0, 1)$ is the degree of risk aversion. The larger α is, the more risk-averse the decision maker is.

Now we investigate the following optimization problem under the CVaR criterion:

$$\inf_{u \in \mathcal{U}} CVaR_\alpha[l(u) + f(u \wedge \Xi)]. \tag{4.6}$$

Theorem 4.3. *Consider the optimization problem* (4.6)*, where functions* $f : \mathcal{F}^n \to \bar{\Re}, l : \mathcal{F}^n \to \bar{\Re}$ *and the constraint set* $\mathcal{U}$ *satisfy the assumptions in Theorem 4.2, the random vector* Ξ *in* $\mathcal{F}^n$ *has independent components, then problem (4.3) and the following optimization problem have the same optimal objective value.*

$$\begin{aligned}
\inf \quad & l(u) + E[g(v(\Xi), \lambda)] \\
s.t. \quad & v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{V} \ \forall \xi \in \mathcal{X}, \\
& v(\xi) \leq \xi \ \forall \xi \in \mathcal{X}, \\
& v(\xi) \leq u \ \forall \xi \in \mathcal{X}, \\
& \lambda \in \Re,
\end{aligned} \tag{4.7}$$

where $g(u, \lambda) = \lambda + \frac{1}{1-\alpha}(f(u) - \lambda)^+$.

The above results can be extended to a more general class of risk measure, called *distortion risk measure*. A distortion risk measure $\rho(\cdot)$ can be represented as a weighted average of CVaRs with different degrees of risk aversion, i.e.,

$$\rho(X) = \int_0^1 CVaR_\alpha(X) d\mu(\alpha).$$

where $\mu(\cdot)$ is the probability measure function. A decision maker with a distortion risk measure faces the following optimization problem:

$$\inf_{u \in \mathcal{U}} \rho[l(u) + f(u \wedge \Xi)]. \tag{4.8}$$

The next corollary demonstrates our transformation with any distortion risk measure.

Corollary 4.1. *Consider the optimization problem* (4.8)*, where all assumptions in Theorem 4.3 are satisfied, then we have the following equivalent* *formulation*

$$\begin{aligned}
\inf \quad & l(u) + \int_0^1 E[g(v(\Xi), \lambda(\alpha), \alpha)] d\mu(\alpha) \\
s.t. \quad & v(\xi) = (v_1(\xi_1), ..., v_n(\xi_n)) \in \mathcal{V} \ \forall \xi \in \mathcal{X}, \\
& v(\xi) \leq \xi \ \forall \xi \in \mathcal{X}, \\
& v(\xi) \leq u \ \forall \xi \in \mathcal{X}, \\
& \lambda(\alpha) \in \Re \ \forall \alpha \in [0, 1),
\end{aligned} \tag{4.9}$$

where $g(u, \lambda(\alpha), \alpha) = \lambda(\alpha) + \frac{1}{1-\alpha}(f(u) - \lambda(\alpha))^+$.

Chapter 5

Algorithms and Computational Studies

5.1 Inventory Substitution with Random Capacities

The transformation technique enables us to convert the original non-convex optimization problems to convex once, and therefore provide opportunities for developing efficient algorithms. In this chapter we will illustrate the algorithms and computational studies through an inventory substitution model. Based on the transformed optimization problem, we develop a fast and effective heuristic method to solve this problem based on piecewise decision rules. Through detailed computational studies, we show that this heuristic approach can achieve a reasonable accuracy and is much faster than the standard Monte Carlo method.

We briefly review the literature on inventory management with substitutions. Parlar & Goyal (1984) and Pasternack & Drezner (1991) are among the first to consider the two-product substitution problem. Bassok et al. (1999) study a single-period multi-product inventory model with downward substitution and arbitrary starting inventory levels. Hsu & Bassok (1999) and Rao et al. (2004) extend this model by introducing random yield and setup costs respectively. Both of these two papers focus on proposing efficient algorithms. Netessine et al. (2002) consider a capacity investment model with single-level substitution and correlated random demands. Later Shumsky & Zhang (2009) extend the model by Netessine et al. (2002) to multiple periods, and Yu et al. (2015) further extend it to allow general downward substitutions. None of the aforementioned papers consider uncertain capacities. In this chapter, we consider an inventory substitution problem with random capacities and apply our transformation technique to achieve a convex formulation.

The inventory substitution model is as follows. The firm manages N types of products to satisfy customer demands. The products are indexed by $i = 1, ..., n$ with product 1 having the highest quality. Corresponding to each product there is a demand class, indexed by $j = 1, ..., n$. If any demand class j cannot be satisfied, products with higher quality ($i \geq j$) can be used for substitution. (some examples of downward substitution in real life) Let $x = (x_1, ..., x_N)^T$ denote the initial inventory level at the beginning of the period. After observing the initial inventory level, the firm decides the order-up-to inventory levels $y = (y_1, ..., y_N)^T$. The ordering quantity of each product i cannot exceed a random capacity, denoted by K_i. Therefore, the actual amount of product i when the order is received is the minimum of y_i and realized capacity k_i. Let $K = (K_1, ..., K_n)^T$. Demands $D = (D_1, ..., D_N)^T$ are then observed and the firm makes the substitution decision to use existing inventory of different products to satisfy each demand class.

We assume that the demands are independent of capacities, while the random capacities of different products can be dependent of each other. We use $d = (d_1, ..., d_N)^T$ to denote the realized demands. The firm's objective is to minimize the expected total costs. The problem formulation is as follows.

$$\min_{y \geq x} E[c(x, y, K) + L(y \wedge (x + K)|D)], \tag{5.1}$$

where

$$\begin{aligned} L(y|d) = & \min_{w_{ij}, u_i^+, u_j^- \geq 0} \sum_{i=1}^{N} h_i u_i^+ + \sum_{j=1}^{N} p_j u_j^- + \sum_{i=1}^{N} \sum_{j=i}^{N} s_{ij} w_{ij}, \\ s.t. \quad & \sum_{i=1}^{j} w_{ij} + u_j^- = d_j, \ \forall j = 1, ..., N, \\ & \sum_{j=i}^{n} w_{ij} + u_i^+ = y_i, \ \forall i = 1, ..., N. \end{aligned} \tag{5.2}$$

In (5.1), $c(x, y, k)$ is the total ordering cost which depends on the initial inventory x, the target inventory level y, and the realized capacity k. We assume that the ordering cost consists of two parts. The first part is proportional to the quantity received (the effective quantity), while the second part is proportional to the quantity initially ordered. so that $c(x, y, k) = c_e^T(y \wedge (x + k) - x) + c_o^T(y - x)$. The two-part cost structure includes, as special cases, setting where the firm only pay for the effect units or where it pays exclusively for all ordered units. In general, this two-part cost structure allows the cost consequence of capacity uncertainties to be shared between the supplier and the retailer. The second term $L(y|d)$ represents the inventory holding and shortage costs as well as the substitution costs given inventory level y and realized demands d. In (5.2), w_{ij} is the amount of substitution of product i to demand j, u_i^+ is the left-over inventory of product i, u_j^- is the shortage of demand j, s_{ij} is the unit substitution cost to use product i to satisfy demand j, and p_j is the unit shortage cost for demand j. The cost parameter h has two interpretations. If $h_i > 0$, it represents the holding cost of product i; If $h_i < 0$, it represents the salvage value of product i.

As is common in the inventory substitution literature, we make the following assumptions of cost parameters.

Assumption 5.1. *$h_i \leq h_j + s_{ij}$, i.e., holding product i is less costly than converting it to product j and holding product j; $p_j \leq p_i + s_{ij}$, i.e., the shortage cost of product j cannot exceed the cost of using product i to satisfy the demand class j and incurring a shortage of product i.*

Define $g(y) = c_e^T y + E[L(y|D)]$, then (5.1) can be simplified as

$$\min_{y \geq x} E[g(y \wedge (x + K))] + c_o^T y - (c_o^T + c_e^T)x. \tag{5.3}$$

In order to apply our transformation technique, we firstly prove the following lemma.

Lemma 5.1. *The function $L(\cdot|d)$ is convex and supermodular for any d.*

Interestingly, in the proof of the above lemma, we convert the substitution stage problem to an equivalent maximum weight circulation problem, and then applied results from Murota (2005). The theorem below shows the transformed problem for the inventory substitution problem.

Theorem 5.1. *The original problem* (5.3) *is equivalent to the following prob-*

lem

$$\begin{aligned} \min \quad & E[g(v(K))] + c_o^T y - (c_o^T + c_e^T)x \\ s.t. \quad & v(k) = (v_1(k_1), ..., v_n(k_n)) \ \forall k \in \mathcal{K}, \\ & v_i(k_i) \geq x_i, \ \forall k_i \in \mathcal{K}_i, \forall i = 1, ..., n \\ & v_i(k_i) \leq x_i + k_i \ \forall k_i \in \mathcal{K}_i, \forall i = 1, ..., n \\ & v_i(k_i) \leq y_i, \ \forall k_i \in \mathcal{K}_i, \forall i = 1, ..., n \\ & v_i(k_i) \ \textit{is increasing}, \ \forall i = 1, ..., n. \end{aligned} \tag{5.4}$$

5.2 Solution Procedures

The transformation technique we developed can be used to convert the original problem given in (5.1) and (5.2) to an equivalent convex problem (5.4). Notice that in the objective function of (5.4), the last two terms are constants, which can be omitted without changing the optimal solution. Let $\tilde{c}_i = c_{e,i} + h_i^+, \ \forall i = 1, ..., N, \tilde{s}_{ij} = s_{ij} - h_i^+ - p_j, \ \forall i = 1, ..., N, j = i, ..., N$. After incorporating $g(\cdot)$ into (5.4), we have the following reformulation, which is referred to as P1 from now on.

$$\min \quad \{\sum_{i=1}^{n} c_{o,i} y_i + E[\sum_{i=1}^{N} \tilde{c}_i v_i(K_i) + \sum_{i=1}^{N} \sum_{j=i}^{N} \tilde{s}_{ij} w_{ij}(K, D)]\} \tag{5.5}$$

$$\text{s.t.} \quad v(k) = (v_1(k_1), ..., v_n(k_n)) \ \forall k \in \mathcal{K}, \tag{5.6}$$

$$\sum_{j=i}^{N} w_{ij}(k, d) \leq v_i(k), \ \forall i = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \tag{5.7}$$

$$\sum_{i=1}^{j} w_{ij}(k, d) \leq d_j, \ \forall j = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \tag{5.8}$$

$$w_{ij}(k, d) \geq 0, \ \forall i = 1, ..., n, j = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \tag{5.9}$$

$$x_i \leq v_i(k_i) \leq x_i + k_i, \ \forall i = 1, ..., N, \forall k_i \in \mathcal{K}_i, \tag{5.10}$$

$$v_i(k_i) \leq y_i, \ \forall i = 1, ..., N, \ \forall k_i \in \mathcal{K}_i, \tag{5.11}$$

$$v_i(k_i) \text{ is increasing}, \ \forall i = 1, ..., N. \tag{5.12}$$

To solve the above problem, we firstly present a linear programming formulation based on Monte Carlo sampling. Then, we will provide a decision rule heuristic algorithm.

5.2.1 An LP formulation based on Monte Carlo Sampling

We use Monte Carlo methods to generate random samples of capacities $k^m, m = 1, ..., M$, and random samples of demands $d^l, l = 1, ..., L$. Therefore, the total number of scenarios is ML. We use u^{ml+}, u^{ml-}, w^{ml} to represent decisions u^+, u^-, w after observing the realizations of capacities k^m and demands d^l. Given the realization of capacities k^m, we have $v(k^m) = (v_1(k_1^m), ..., v_i(k_i^m), ..., v_n(k_n^m))$.

$$\begin{array}{ll}
\min & \{\sum_{i=1}^N c_{o,i} y_i + \frac{1}{M}\sum_{m=1}^M \sum_{i=1}^N \tilde{c}_i v_i(k_i^m) + \frac{1}{ML}\sum_{m=1}^M \sum_{l=1}^L \sum_{i=1}^N \sum_{j=1}^N \tilde{s}_{ij} w_{ij}^{ml}\} \\
\text{s.t.} & \sum_{j=i}^N w_{ij}^{ml} \le v_i(k_i^m) \ \forall i = 1, ..., N, \forall m, l, \\
& \sum_{i=1}^j w_{ij}^{ml} \le d_j^l \ \forall j = 1, ..., N, \forall m, l, \\
& w_{ij}^{ml} \ge 0 \ \forall i, j, m, l, \\
& x_i \le v_i(k_i^m) \le x_i + k_i^m \ \forall i = 1, ..., N, \forall m, \\
& v_i(k_i^m) \le y_i \ \forall i = 1, ..., N, \forall m, \\
& v_i(k_i^{m'}) \le v_i(k_i^{m''}), \ \forall m', m'' \text{ s.t. } k_i^{m'} \le k_i^{m''}, \forall i = 1, ..., N.
\end{array} \tag{5.13}$$

The last constraint in (5.13) corresponds to the constraints "$v_i(k_i)$ is increasing" in P1. In our computational studies, for each component i we sort the capacity scenarios generated by Monte Carlo methods, then we need at most M inequalities to represent the monotonicity of $v_i(k_i)$. The total number of decision variables and constraints of problem (5.13) scales linearly with the number of generated scenarios ML. Solving problem (5.13) is very time-consuming. To strike a balance between accuracy and computation time, in all our computational experiments we choose $M = 100, L = 100$.

5.2.2 A Heuristic Approach Using Piecewise linear decision rules

This section presents a heuristic approach based on piecewise linear decision rules. Firstly we will briefly review some related papers on stochastic optimization and decision rule approaches. Stochastic optimization problems are notoriously difficult to solve. One suitable method to approximately solve stochastic programs is to impose that the recourse decision is a linear function of the uncertainties. This linear decision rule approximation provides a tractable and scalable methodology and hence attracted considerable interests in recent years. For instance, linear decision rule approaches have been used to solve multistage stochastic linear programs (Ben-Tal et al. 2004, Chen, Sim & Sun 2007). In terms of loss of optimality, for the stochastic optimization with expectation as objective function, Kuhn et al. (2011) propose a primal-dual approach of linear decision rules to find both the upper and lower bounds of the optimal value for linear dynamic systems. To improve the performance of linear decision rules and reduce the approximation error, Chen et al. (2008) introduce deflected and segregated linear decision rules. The theory of piecewise linear decision rules and other nonlinear decision rules has been developed in Goh & Sim (2010) and Georghiou et al. (2015). There are many interesting applications which employ the aforementioned decision rule approaches to conduct computational studies (see Atamtürk & Zhang 2007, Chen & Zhang 2009, and See & Sim 2010 among others). In this paper, we propose a heuristic approach to solve the optimization problem after transformation using piecewise linear decision rules.

We start with a brief introduction of the piecewise linear decision rule approach. Notice that problem (5.5) can be formulated as a multi-stage stochastic linear program with some additional requirements as follows.

$$\begin{array}{ll} \min & E_{\Xi}[\sum_t^T c_t^\intercal u_t(\Xi^t)] \\ \text{s.t.} & \sum_{s=1}^t A_{ts} u_s(\xi^s) \geq b_t(\xi^t) \quad \forall \xi \in \mathcal{X}, t = 1, ..., T, \\ & u_t(\xi^t) \geq 0 \quad \forall \xi \in \mathcal{X}, t = 1, ..., T, \\ & u_t(\xi^t) \in \mathbb{U}_t, \ \forall t = 1, ..., T. \end{array} \tag{5.14}$$

In the above formulation, $\xi^t = (\xi_1, ..., \xi_t)$ is the observation history of random variables up to stage t. There are 3 stages in our problem. In the first stage, before any random variables are realized (for expositional reasons, we assume that there is a dummy random variable realized, which is equal to constant 1), the decision vector is y; In the second stage, after the random capacities are realized, the decision vector is $v(k)$; In the third stage, after both the random capacities and demands are realized, the decision vector is $w(k, d)$. Therefore, we have $\xi_1 = 1, \xi_2 = k, \xi_3 = d$, $u_1 = y, u_2 = v, u_3 = w$. In the objective function, we have $c_1 = c_o$, $c_2 = \tilde{c}, c_3 = \tilde{s}$. Constraints (5.7)-(5.11) in P1 can be easily represented as the linear constraints in (5.14). Constraints (5.6) and (5.12) in P1 are represents as $u_t \in \mathbb{U}_t$ in (5.14). To be more specific, we require $\mathbb{U}_2 = \{(v(k), k \in \mathcal{K}) | v(k) = (v_1(k_1), ..., v_n(k_n)) \ \forall k \in$

$\mathcal{K}, v_i(k_i)$ is increasing, $\forall i.\}$

Let $\xi = (\xi_1, ..., \xi_T)$. Define truncation operator P_t through $P_t\xi = \xi^t$. The support of ξ is denoted by $\mathcal{X}$, which is assumed to be a bounded polyhedron of the form $\mathcal{X} = \{\xi \in \Re^k : W\xi \geq h\}$ for some $W \in \Re^{l\times k}$ and $h \in \Re^l$. We allow $\mathcal{X}$ to contain infinitely many scenarios. Assume that $b_t(\xi^t) = B_t\xi^t$ for some matrix B_t.

Before we introduce the piecewise linear decision rule approach, we firstly demonstrate the linear decision rule formulation. We model the decision rule $x_t(\xi^t)$ as a linear function of ξ^t (an affine function of $(\xi_2, ..., \xi_t)$ since $\xi_1 = 1$), it can thus be expressed as $x_t(\xi^t) = X_t\xi^t$. The linear decision rule problem of (5.14) is equivalent to:

$$\begin{array}{ll} \min & \sum_{t=1}^{T} c_t^\intercal X_t P_t E[\Xi] \\ \text{s.t.} & (\sum_{s=1}^{t} A_{ts} X_s P_s - B_t P_t)\xi \geq 0 \quad \forall \xi \in \mathcal{X}, t = 1, ..., T, \\ & X_t P_t \xi \geq 0 \quad \forall \xi \in \mathcal{X}, t = 1, ..., T, \\ & X_t \in \mathbb{X}_t, \ \forall t = 1, ..., T. \end{array} \tag{5.15}$$

The last constraint $X_t \in \mathbb{X}_t$ corresponds to the last constraint in (5.14), which can be represented as linear constraints. To see this, we need entries $(i, 1), (i, i+1), i = 1, ..., N$ of X_2 are nonnegative, and all other entries are zero. This is represented as the following matrix, where $+$ means that the entry needs to be nonnegative:

$$\begin{pmatrix} + & + & 0 & \cdots & 0 \\ + & 0 & + & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ + & 0 & 0 & \cdots & + \end{pmatrix}.$$

Clearly these requirements can be represented as linear constraints.

Problem (5.15) can have infinitely many constraints parameterised by $\xi \in \mathcal{X}$. Applying Proposition 3.1 in Georghiou et al. (2011), we can reformulate the $\xi-$dependent constraints in (5.15) in terms of a finite number of linear constraints and obtain

$$
\begin{array}{ll}
\min & \sum_{t=1}^{T} c_t^{\intercal} X_t P_t E[\Xi] \\
\text{s.t.} & \sum_{s=1}^{t} A_{ts} X_s P_s - B_t P_t = \Lambda_t W,\ \Lambda_t h \geq 0,\ \Lambda_t \geq 0 \quad \forall t = 1, ..., T, \\
& X_t P_t = \Gamma_t W,\ \Gamma_t h \geq 0,\ \Gamma_t \geq 0, \quad \forall t = 1, ..., T, \\
& X_t \in \mathbb{X}_t,\ \forall t = 1, ..., T.
\end{array}
\tag{5.16}
$$

As mentioned earlier, constraints $X_t \in \mathbb{X}_t$ can be formulated as a finite number of linear constraints. Therefore, problem (5.16) is a finite linear program.

The piecewise linear decision rule approach model the decision rule $x_t(\xi^t)$ as a piecewise linear function of ξ^t. Following Georghiou et al. (2015), we lift the original random vector ξ to a higher space, so that x_t can be expressed as an affine function of lifted random vector ξ'. To define the lifted random vector, select a set of breakpoints for each coordinate axis in $\mathcal{X}$. The breakpoints are denoted by

$$z_1^i < ... < z_{r_i-1}^i \ \forall i = 2, ..., k.$$

The number of breakpoints along the ξ_i axis is $r_i - 1$. The dimension of the lifted space is $k = \sum_{i=1}^{k} r_i$. The lifted random vector can be written as

$$\xi' = (\xi'_{1,1}, \xi'_{2,1}, ..., \xi'_{2,r_2}, ..., \xi'_{k,1}, ..., \xi'_{k,r_k})^{\intercal}.$$

Define the lifting operator $L = (L_{1,1}, ..., L_{i,j}, ..., L_{k,r_k})$ which maps ξ to ξ' as follows

$$
L_{i,j}(\xi) = \begin{cases}
\xi_i, & \text{if} \quad r_i = 1, \\
\min\{\xi_i, z_1^i\}, & \text{if} \quad r_i > 1, j = 1, \\
\max\{\min\{\xi_i, z_j^i\} - z_{j-1}^i, 0\}, & \text{if} \quad r_i > 1, j = 2, ..., r_i - 1, \\
\max\{\xi_i - z_{j-1}^i, 0\}, & \text{if} \quad r_i > 1, j = r_i.
\end{cases}
\tag{5.17}
$$

By construction, the piecewise linear decision rule of the original random vector ξ with breakpoints $\{z_j^i, j = 1, .., r_i - 1, i = 2, ..., k\}$ is equivalent to the linear decision rule of the lifted random vector ξ'. We model the decision rule $x_t(\xi^t)$ as a linear function of ξ'^t, it can thus be expressed as $x_t(\xi^t) = X'_t \xi'^t = X'_t L(\xi^t)$.

The linear retraction operator corresponding to L is denoted by $R = (R_1, ..., R_k)$, where the coordinate mapping R_i corresponds to the ξ_i axis in the original space is defined through

$$R_i(\xi') = \sum_{j=1}^{r_i} \xi'_{i,j}.$$

Define truncation operators P'_t which map ξ' to ξ'^t. The piecewise-linear decision rule problem of (5.14) is equivalent to

$$\begin{array}{ll} \min & \sum_{t=1}^{T} c_t^\intercal X'_t P'_t E[\Xi'] \\ \text{s.t.} & (\sum_{s=1}^{t} A_{ts} X'_s P'_s - B_t P_t R)\xi' \geq 0 \quad \forall \xi' \in \mathcal{X}', t = 1, ..., T, \\ & X'_t P'_t \xi' \geq 0 \quad \forall \xi' \in \mathcal{X}', t = 1, ..., T, \\ & X_t \in \mathbb{X}'_t, \ \forall t = 1, ..., T. \end{array} \tag{5.18}$$

Similar to the linear decision rule approach, the last constraint $X_t \in \mathbb{X}'_t, \ \forall t = 1, ..., T$ can easily formulated using a finite number of linear constraints. According to Georghiou et al. (2011), if the support of random vector ξ is a hyperrectangle of the type $\mathcal{X} = \{\xi \in \Re^k : l \leq \xi \leq u\}$, then the convex hull of the support of lifted random variables has the representation $\{\xi' : W'\xi' \geq h'\}$, and the $\xi'-$dependent constraints in (5.18) can be represented using a finite number of linear constraints as well, which leads to the following reformulation.

$$\begin{array}{ll} \min & \sum_{t=1}^{T} c_t^\intercal X'_t P'_t E[\Xi'] \\ \text{s.t.} & \sum_{s=1}^{t} A_{ts} X'_s P'_s - B_t P_t R = \Lambda_t W', \ \Lambda_t h' \geq 0, \ \Lambda_t \geq 0, \quad \forall t = 1, ..., T, \\ & X'_t P'_t = \Gamma_t W', \ \Gamma_t h' \geq 0, \ \Gamma_t \geq 0 \quad \forall t = 1, ..., T, \\ & X_t \in \mathbb{X}'_t, \ \forall t = 1, ..., T. \end{array} \tag{5.19}$$

5.3 Computational Studies

Our numerical experiments implement the large scale LP formulation using the Monte Carlo methods and the piece-wise linear decision rule approximation. Our objective is to compare the performance of these two methods in terms of running time and accuracy. Our numerical studies are conducted with the number of products $N = 3, 5, 10$ and 20. For each value of N, we vary the model parameters as follows.

Experiment Setup. For each product i, the unit ordering cost $c_i =$

$c_{o,i}+c_{e,i}$, where $c_{o,i}$ is the unit ordering cost for each product initially ordered, and $c_{e,i}$ is the unit ordering cost for each product actually received after the realization of the random capacity. We let $c_{o,i} = c_{e,i} = 0.5c_i$ and $c_i = 1 + \eta(N - i)$ for $\eta = 0.1, 0.2, 0.3, 0.4, 0.5$. Here η measures the variation of products' ordering costs. The cost of using product i to satisfy demand j is chosen as $s_{ij} = 0.5(c_i - c_j)$. We set the holding cost/savage value for each product i as $h_i = -0.5c_i, 0$, or $0.5c_i$. Notice that $h_i = 0.5c_i$ means that product i incurs a holding cost which is equal to 50% of the ordering cost, while $h_i = -0.5c_i$ means that it has a savage value which is 50% of the ordering cost. The penalty costs are chosen as $p_i = 2c_i, 5c_i, 10c_i$.

The demand of each product follows a truncated Gaussian distribution with mean 100, and coefficient of variation, denoted by CV_d of $0.1, 0.2, 0.3$. In order to model the positive dependence of random capacities, we consider the case where products share a common "market risk". In this case, the capacity of each player i is given by $K_i = Y_i + Z$, where $Y_i, i = 1, ..., N$ and Z are mutually independent Gaussian variables. The "market risk", denoted by Z, is embedded in each product's capacity. We let Z and Y_i follow a truncated Gaussian distribution with mean 50 and coefficient of variation $CV_k = 0.1, 0.2, 0.3$. The support of demands is chosen to be a polyhedral $\{d \in \Re^N : 50 \leq d_j \leq 150, j = 1, ..., N\}$. The support of Z is $\{z \in \Re : 25 \leq z \leq 75\}$ and the support of Y_i is $\{y_i \in \Re : 25 \leq y_i \leq 75\}$. For each dimension we divided into $r = 20$ pieces for PLDR.

Performance Metrics. We run all experiments using MATLAB 2014a which calls Gurobi to solve linear programs. The running time of the two methods is measured in CPU seconds of a server with Intel Xeon CPUs E5-4657L v2 2.4GHz and 256GB memory. We use the optimal costs obtained from solving the LP formulation based on Monte Carlo sampling (thereafter as MCLP) as the bench mark, and define the performance error of the piecewise-linear decision rule approach (thereafter as PLDR) as follows:

$$\%error = (\frac{\mathcal{C}(PLDR)}{\mathcal{C}(MCLP)} - 1) \times 100\% \tag{5.20}$$

To calculate PLDR's optimal cost $\mathcal{C}(PLDR)$, we firstly obtain the optimal ordering quantities by solving the PLDR problem, then employ a greedy algorithm to obtain the substitution decisions (See Bassok et al. (1999) or Rao et al. (2004) for details of the greedy algorithm). Then we conduct a simulation with one million cases using the optimal ordering and substitution quantities to obtain the PLDR's optimal cost.

Numerical Results. Table 5.1 demonstrates the average running time for both the PLDR and MCLP methods. The PLDR is much faster than the MCLP method. Notice that when there are $N = 10$ products, it takes about three hours to run each instance using the MCLP method, therefore we only run 15 instances. When $N = 20$, the server is out of memory for the MCLP method. On the other hand, it takes less than 4 minutes to solve the instance with $N = 20$ products using the PLDR method. While the PLDR runs much faster than the MCLP method, it can also achieve a reasonable performance accuracy. Table 5.2 shows the performance error as defined in (5.20). We can see that for both $N = 3$ and $N = 5$ the PLDR have an average error of less than 2.5%.

Table 5.1: Average running time, in CPU seconds, NA = not applicable. For PLDR, 405 instances for each N. For MCLP, 405 instances for $N = 3, 5$, and 15 instances for $N = 10$

	N=3	N=5	N=10	N=20
PLDR	3.34	7.42	20.63	218.05
MCLP	848.42	2107.36	11855.48	NA

Table 5.2: Performance error of PLDR

	Average (%)	Std. dev. (%)	Maximum (%)
% Error for $N = 3$ products	2.19	2.24	10.71
% Error for $N = 5$ products	2.48	2.29	11.83

Chapter 6

Conclusions and Future Research

This thesis studies stochastic optimization problems with decisions truncated by random variables and its applications in operations management. We develop a transformation technique to convert the original non-convex optimization problems to convex ones while preservation some desired structural properties, which are useful for characterizing optimal decision policies and conducting comparative statics. Our transformation technique provides a unified approach to analyze a broad class of models in inventory control and revenue management. In additional, we develop efficient algorithms to solve the transformed stochastic optimization problem.

There are several interesting directions which call for future research. Although in Chapter 4 we provide generalizations of the transformation technique, they are by no means comprehensive. For example, Theorem 4.1 assumes that the random variables are positively dependent, it would be interesting to consider some other dependence structures. In Chapter 4.3 we incorporate distortion risk measures into the model. A natural question is can we manage to incorporate more general risk measures, say coherent risk measures.

Chapter 5 proposes a heuristic approach based on piecewise linear decision rules. Whether there are methods to further improve the performance is an interesting question. Moreover, considering a robust optimization formulation and then applying the piecewise linear decision rule is also worth trying. Apart from the inventory substitution problem, the piecewise linear decision rule approach can also be applied to airline revenue management problems. Further numerical studies need to be done to see whether it performs well in revenue management settings.

References

Ahmed, S., Çakmak, U. & Shapiro, A. (2007), 'Coherent risk measures in inventory problems', *European Journal of Operational Research* **182**(1), 226–238.

Atamtürk, A. & Zhang, M. (2007), 'Two-stage robust network flow and design under demand uncertainty', *Operations Research* **55**(4), 662–673.

Barankin, E. W. (1961), 'A delivery-lag inventory model with an emergency provision (the one-period case)', *Naval Res. Logist. Quart* **8**(3), 258–311.

Bassok, Y., Anupindi, R. & Akella, R. (1999), 'Single-period multiproduct inventory models with substitution', *Operations Research* **47**(4), 632–642.

Ben-Tal, A., Goryashko, A., Guslitzer, E. & Nemirovski, A. (2004), 'Adjustable robust solutions of uncertain linear programs', *Mathematical Programming* **99**(2), 351–376.

Bollapragada, R., Rao, U. S. & Zhang, J. (2004), 'Managing inventory and supply performance in assembly systems with random capacity and demand', *Management Sci* **50**(12), 1729–1743.

Brumelle, S. L. & Mcgill, J. I. (1993), 'Airline seat allocation with multiple nested fare classes', *Oper. Res* **41**(1), 127–137.

Chen, L. & Homem-de Mello, T. (2010), 'Re-solving stochastic programming models for airline revenue managment', *Ann Oper. Res* **177**, 91–114.

Chen, W., Feng, Q. & Seshadri, S. (2013), 'Sourcing from suppliers with random yield for price-dependent demand', *Annals of Operations Research* **208**(1), 557–579.

Chen, X., Hu, P. & He, S. (2013), 'Technical note-preservation of supermodularity in parametric optimization problems with nonlattice structures', *Oper. Res* **61**(5), 1166–1173.

Chen, X., Pang, Z. & Pan, L. (2014), 'Coordinating inventory control and pricing strategies for perishable products', *Oper. Res* **62**(2), 284–300.

Chen, X., Sim, M., Simchi-Levi, D. & Sun, P. (2007), ‘Risk aversion in inventory management’, *Operations Research* **55**(5), 828–842.

Chen, X., Sim, M. & Sun, P. (2007), ‘A robust optimization perspective on 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, X. & Zhang, Y. (2009), ‘Uncertain linear programs: Extended affinely adjustable robust counterparts’, *Operations Research* **57**(6), 1469–1482.

Chen, Y., Xu, M. & Zhang, Z. G. (2009), ‘Technical notea risk-averse newsvendor model under the cvar criterion’, *Operations research* **57**(4), 1040–1044.

Ciarallo, F. W., Akella, R. & Morton, T. E. (1994), ‘A periodic review, production planning model with uncertain capacity and uncertain demandoptimality of extended myopic policies’, *Management Science* **40**(3), 320–332.

Cooper, W. & Homem-de Mello, T. (2007), ‘Some decomposition methods for revenue management’, *Transportation Science* **41**(3), 332–353.

Daniel, K. (1963), A delivery-lag inventory model with emergency, *in* H. Scarf, D. Gilford & M. Shelly, eds, ‘Multistage Inventory Models and Techniques’, Stanford University Press, Stanford, CA, pp. 32–46.

Federgruen, A. & Yang, N. (2008), ‘Selecting a portfolio of suppliers under demand and supply risks’, *Oper. Res* **56**(4), 916–936.

Federgruen, A. & Yang, N. (2011), ‘Technical noteprocurement strategies with unreliable suppliers’, *Operations research* **59**(4), 1033–1039.

Feng, Q. (2010), ‘Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty’, *Management Science* **56**(12), 2154–2172.

Feng, Q., Jia, Z. & George, S. (2015), Managing dependent random supply capacities in dynamic inventory-pricing problems. Presentation at 2015 Mostly OM Conference.

Feng, Q. & Shi, R. (2012), ‘Sourcing from multiple suppliers for price-dependent demands’, *Production and Operations Management* **21**(3), 547–563.

Fukuda, Y. (1964), ‘Optimal policies for the inventory problem with negotiable leadtime’, *Manage. Sci* **10**(4), 690–708.

Georghiou, A., Wiesemann, W. & Kuhn, D. (2011), 'The decision rule approach to optimisation under uncertainty: Methodology and applications in operations management'.

Georghiou, A., Wiesemann, W. & Kuhn, D. (2015), 'Generalized decision rule approximations for stochastic programming via liftings', *Mathematical Programming* **152**(1-2), 301–338.

Goh, J. & Sim, M. (2010), 'Distributionally robust optimization and its tractable approximations', *Operations research* **58**(4-part-1), 902–917.

Gong, X., Chao, X. & Zheng, S. (2014), 'Dynamic pricing and inventory management with dual suppliers of different leadtimes and disruption risks', *Production and Operations Management* **23**(12), 2058–2074.

Henig, M. & Gerchak, Y. (1990), 'The structure of periodic review policies in the presence of random yield', *Operations Research* **38**(4), 634–643.

Hsu, A. & Bassok, Y. (1999), 'Random yield and random demand in a production system with downward substitution', *Operations Research* **47**(2), 277–290.

Hu, X., Duenyas, I. & Kapuscinski, R. (2008), 'Optimal joint inventory and transshipment control under uncertain capacity', *Operations Research* **56**(4), 881–897.

Huh, W. & Janakiraman, G. (2010), 'On the optimal policy structure in serial inventory systems with lost sales', *Oper. Res* **58**(2), 481–491.

Kuhn, D., Wiesemann, W. & Georghiou, A. (2011), 'Primal and dual linear decision rules in stochastic and robust optimization', *Mathematical Programming* **130**(1), 177–209.

Li, T., Sethi, S. P. & Zhang, J. (2013), 'Supply diversification with responsive pricing', *Production and Operations Management* **22**(2), 447–458.

Lu, Y. & Song, J. S. (2005), 'Order-based cost optimization in assemble-to-order systems', *Oper. Res* **53**(1), 151–169.

Möller, A., Römisch, W. & Weber, K. (2008), 'Airline network revenue management by multistage stochastic programming', *Computational Management Science* **5**(4), 355–377.

Murota, K. (1998), 'Discrete convex analysis', *Mathematical Programming* **83**, 313–371.

Murota, K. (2003), *Discrete convex analysis*, SIAM.

Murota, K. (2005), 'Note on multimodularity and l-convexity', *Math. Oper. Res* **30**(3), 658–661.

Murota, K. (2009), Recent developments in discrete convex analysis. in: W, *in* L. L. Cook & J. Vygen, eds, 'Research Trends in Combinatorial Optimization, Bonn 2008', Springer-Verlag, Berlin, Chapter 11, pp. 219–260.

Nadar, E., Akan, M. & Sheller-Wolf, A. (2014), 'Optimal structural results for assemble-to-order generalized m-systems', *Oper. Res* **62**(3), 571–579.

Netessine, S., Dobson, G. & Shumsky, R. (2002), 'Flexible service capacity: Optimal investment and the impact of demand correlation', *Operations Research* **50**(2), 376–388.

Pang, Z., Chen, F. & Feng, Y. (2012), 'A note on the structure of joint inventory-pricing control with leadtimes', *Oper. Res* **60**(3), 581–587.

Parlar, M. & Goyal, S. (1984), 'Optimal ordering decisions for two substitutable products with stochastic demands', *Opsearch* **21**(1), 1–15.

Pasternack, B. A. & Drezner, Z. (1991), 'Optimal inventory policies for substitutable commodities with stochastic demand', *Naval Research Logistics* **38**(2), 221–240.

Rao, U., Swaminathan, J. & Jun, Z. (2004), 'Muldi-product inventory planning with downward substitution, sochastic demand and setup costs', *IIE Transactions* **36**, 59–71.

Robinson, L. (1995), 'Optimal and approximate control policies for airline booking with sequantial nonmonotomic fare classes', *Oper. Res* **43**(2), 252–263.

Rockafellar, R. T. & Uryasev, S. (2000), 'Optimization of conditional value-at-risk', *Journal of risk* **2**, 21–42.

See, C.-T. & Sim, M. (2010), 'Robust approximation to multiperiod inventory management', *Operations research* **58**(3), 583–594.

Shaked, M. & Shanthikumar, G. (2006), *Stochastic Orders*, Springer-Verlag, New York.

Shumsky, R. A. & Zhang, F. (2009), 'Dynamic capacity management with substitution', *Operations research* **57**(3), 671–684.

Simchi-levi, D., Chen, X. & Bramel, J. (2014), *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management*, 3rd edn, Springer-Verlag, New York.

Song, J. & Zipkin, P. (2003), Supply chain operations: Assemble-to-order systems, *in* T. de Kok & S. G. and, eds, 'Supply Chain Management', Handbooks in Operations Research and Management Science, Vol. 30. North-Holland, Amsterdam, pp. 561–596.

Talluri, K. & Van Ryzin, G. (2005), *The Theory and Practice of Revenue Management*, Kluwer Academic Publishers, Boston.

Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton University Press, Princeton, NJ.

Van Mieghem, J. A. (2008), 'Operations strategy: Principles and practice', *Dynamic Ideas, Belmont, MA, pp* pp. 230–232.

Wang, Y. & Gerchak, Y. (1996), 'Periodic review production models with variable capacity, random yield, and uncertain demand', *Management science* **42**(1), 130–137.

Whittmore, A. S. & Saunders, S. C. (1977), 'Optimal inventory under stochastic demand with two supply options', *SIAM J. Appl. Math* **32**(2), 293–305.

Yu, Y., Chen, X. & Zhang, F. (2015), 'Dynamic capacity management with general upgrading', *Operations Research* **63**(6), 1372–1389.

Zhou, X. & Chao, X. (2014), 'Dynamic pricing and inventory management with regular and expedited supplies', *Production and Operations Management* **23**(1), 65–80.

Zipkin, P. (2008), 'On the structure of lost-sales inventory models', *Oper. Res* **56**(4), 937–944.

Zălinescu, C. (2002), *Convex Analysis in General Vector Spaces*, World Scientisfic, Singapore.

Appendix A

A.1 Proof of Lemma 2.1

The quasi-convexity of f implies that $f(x)$ decreases in x as $x \leq x^*$ and increases in x as $x \geq x^*$. If $a \geq x^*$, we have $b \geq a \geq x^*$, which implies that $a \geq x^* \wedge b = x^*$. If $a \leq x^*$, since $a \leq b$, we have $a \leq x^* \wedge b \leq x^*$. In either case, $f(x^* \wedge b) \leq f(a)$. $\square$

A.2 Proof of Theorem 2.1

Let π^* be the optimal objective value of problem (2.3). Since for any $u \in \mathcal{F}^n$, $v(\xi) = u \wedge \xi$ is feasible for problem (2.3), $\pi^* \leq \tau^*$.

It remains to show that $\tau^* \leq \pi^*$. Clearly, it holds when $\pi^* = \infty$. Thus, in the following, we assume that $\pi^* < \infty$, which together with assumption (a) implies that all optimization problems involved below, as well as problems (2.1) and (2.3), admit finite optimal solutions. Given any optimal solution of (2.3) denoted by $v^* = (v^*(\xi)|\xi \in \mathcal{X})$, we will show that we can find a solution $\hat{u} \in \mathcal{F}^n$ such that $E[f(\hat{u} \wedge \Xi)] = E[f(v^*(\Xi))]$.

We first show that it is true for $n = 1$. Let $\hat{u} = \arg\min_{u \in \mathcal{F}} f(u)$ (when there are multiple optimal solutions, we choose the smallest one). Consider any feasible solution $v = (v(\xi)|\xi \in \mathcal{X})$ of problem (2.3). We have $f(\hat{u} \wedge \xi) \leq f(v(\xi))$ for any $\xi \in \mathcal{X}$ according to Lemma 2.1. Hence, $E[f(\hat{u} \wedge \Xi)] \leq \pi^*$. Note that $\hat{u}$ is a feasible solution for problem (2.1), which implies that $\tau^* = E[f(\hat{u} \wedge \Xi)] \leq \pi^*$. Combined with the fact that $\pi^* \leq \tau^*$, we have $\tau^* = \pi^*$.

We now consider the general case with $n \geq 1$. Use v_i^* to represent the ith component of v^* for $i = 1, ..., n$. Starting from the first component, define

$$\pi_1(u_1) = E[f(u_1, v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))].$$

The component-wise convexity of f implies that $\pi_1(u_1)$ is convex in u_1. Since the components of the vector Ξ are independently distributed, $E_{\Xi_1}[\pi_1(v_1(\Xi_1))] = E_{\Xi}[f(v_1(\Xi_1), v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))]$ for any measurable function $v_1(\cdot)$, and the preceding analysis for $n = 1$ implies that there exists a $\hat{u}_1$ such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_1(v_1(\Xi_1))] | v_1(\xi_1) \le \xi_1, v_1(\xi_1) \in \mathcal{F},\ \forall \xi_1 \in \mathcal{X}_1\}\\ &= \min_{u_1 \in \mathcal{F}} E[\pi_1(u_1 \wedge \Xi_1)]\\ &= E[\pi_1(\hat{u}_1 \wedge \Xi_1)].\end{aligned}$$

Next define $\pi_2(u_2) = E[f(\hat{u}_1 \wedge \Xi_1, u_2, v_3^*(\Xi_3), \ldots, v_n^*(\Xi_n))]$. Clearly, π_2 is convex. Following the preceding analysis, there exists a $\hat{u}_2$ such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_2(v_2(\Xi_2))] | v_2(\xi_2) \le \xi_2, v_2(\xi_2) \in \mathcal{F},\ \forall \xi_2 \in \mathcal{X}_2\}\\ &= \min_{u_2 \in \mathcal{F}} E[\pi_2(u_2 \wedge \Xi_2)]\\ &= E[\pi_2(\hat{u}_2 \wedge \Xi_2)].\end{aligned}$$

Continue this process and define $\pi_i(u_i) = E[f(\hat{u}_1 \wedge \Xi_1, ..., \hat{u}_{i-1} \wedge \Xi_{i-1}, u_i, v_{i+1}^*(\Xi_{i+1}), \ldots, v_n^*(\Xi_n))]$. Applying the same approach, we can find $\hat{u}_i, i = 3, ..., n$, such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_i(v_i(\Xi_i))] | v_i(\xi_i) \le \xi_i, v_i(\xi_i) \in \mathcal{F},\ \forall \xi_i \in \mathcal{X}_i\}\\ &= \min_{u_i \in \mathcal{F}} E[\pi_i(u_i \wedge \Xi_i)]\\ &= E[\pi_i(\hat{u}_i \wedge \Xi_i)].\end{aligned}$$

Therefore, we have

$$\begin{aligned}\pi^* &= E[\pi_n(\hat{u}_n \wedge \Xi_n)]\\ &= E[f(\hat{u}_1 \wedge \Xi_1, ..., \hat{u}_n \wedge \Xi_n)].\end{aligned}$$

Since $\hat{u}$ is a feasible solution to (2.1), we have $\tau^* \le E[f(\hat{u} \wedge \Xi)] = \pi^*$. Combined with the fact that $\pi^* \le \tau^*$, we have $\pi^* = \tau^*$. $\square$

A.3 Proof of Theorem 2.2

Problem (2.4) is equivalent to the following unconstrained optimization problem.

$$\inf_{u \in \mathcal{F}^n} \{E[f(u \Diamond_k \Xi)] + \delta_{\mathcal{U}}(u)\}. \tag{A.1}$$

Define for any $v \in \mathcal{F}^n$,

$$\hat{f}(v) = f(v) + \delta_{\mathcal{V}}(v),$$

where $\mathcal{V}$ is defined in (2.5). Then by Assumption 2.1 the optimal objective value of problem (A.1) is equivalent to that of the following problem

$$\inf_{u \in \mathcal{F}^n} E[\hat{f}(u \Diamond_k \Xi)]. \tag{A.2}$$

To see this, note that for any $u \in \mathcal{U}$, we have $u \Diamond_k \xi \in \mathcal{V}$ $\forall \xi \in \mathcal{X}$, and hence $\inf_{u \in \mathcal{F}^n} E[\hat{f}(u \Diamond_k \Xi)] \le \inf_{u \in \mathcal{U}} E[f(u \Diamond_k \Xi)]$. On the other hand, due to Assumption 2.1, we have $\inf_{u \in \mathcal{F}^n} E[\hat{f}(u \Diamond_k \Xi)] \ge \inf_{u \in \mathcal{U}} E[f(u \Diamond_k \Xi)]$.

Define a new random vector $\tilde{\Xi}$ with $(\tilde{\Xi}_1, \ldots, \tilde{\Xi}_k, \tilde{\Xi}_{k+1}, \ldots, \tilde{\Xi}_n) = (\Xi_1, \ldots, \Xi_k, -\Xi_{k+1}, \ldots, -\Xi_n)$ and a new function $\tilde{f} : \mathcal{F}^n \to \bar{\Re}$ by

$$\tilde{f}(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n) = \hat{f}(u_1, \ldots, u_k, -u_{k+1}, \ldots, -u_n).$$

Then problem (A.2) is equivalent to the problem

$$\inf_{\tilde{u} \in \mathcal{F}^n} E[\tilde{f}(\tilde{u} \wedge \tilde{\Xi})].$$

By Theorem 2.1, it has the same optimal objective value with the following problem.

$$\begin{aligned}
\inf \quad & E[\tilde{f}(\tilde{v}(\tilde{\xi}))] \\
\text{s.t.} \quad & \tilde{v}(\tilde{\xi}) \le \tilde{\xi} \quad \forall \tilde{\xi} \in Supp(\tilde{\Xi}), \\
& \tilde{v}(\tilde{\xi}) = (\tilde{v}_1(\tilde{\xi}_1), \ldots, \tilde{v}_n(\tilde{\xi}_n)) \in \mathcal{F}^n,
\end{aligned}$$

which is clearly equivalent to problem (4.5) from the definition of $\tilde{f}$. Notice that the indicator function of the set $\mathcal{V}$ needs to be componentwise convex to ensure that $\tilde{f}$ is componentwise convex. $\square$

A.4 Proof of Lemma 2.2

For notational convenience, we only prove the case where $k = n$, i.e., there is only the $\wedge$ operation. This is because we can apply the same technique used in the proof of Theorem 2.2 to convert a problem with the $\vee$ operation to a new one only with the $\wedge$ operation. In this case the set $\mathcal{U} = \{u | Au \leq b, u_j \geq \underline{u}_j, j = 1, ..., n\}$, where $a_{ij} \geq 0$ for all $i = 1, ..., m$, $j = 1, ..., n$, and $\xi_j \geq \underline{u}_j \; \forall \xi_j \in \mathcal{X}_j$ for $j = 1, ..., n$.

Recall that we define $\bar{\xi}_j = \operatorname{ess\,sup}\{\xi_j | \xi \in \mathcal{X}\}$. We first consider the case where $\bar{\xi}_j < \infty$ for all j. Note that $\mathcal{V} = \{u \wedge \xi | u \in \mathcal{U}, \xi \in \mathcal{X}\}$ is equivalent to the following set $\{w | Aw \leq b, \underline{u}_j \leq w_j \leq \bar{\xi}_j, j = 1, ..., n\}$, denoted by $\mathcal{V}_w$. For any $w = u \wedge \xi \in \mathcal{V}$, we have $Aw = A(u \wedge \xi) \leq b$ since $a_{ij} \geq 0$ for all i, j and $Au \leq b$; $\underline{u}_j \leq u_j \wedge \xi_j = w_j \leq \bar{\xi}_j$ since $\underline{u}_j \leq \xi_j \; \forall \xi_j \in \mathcal{X}_j$. For any $w \in \mathcal{V}_w$, let $u = w$, $\xi_j = \bar{\xi}_j$ for all j. Then $w = u \wedge \xi$ since $w_j \leq \bar{\xi}_j$ for all j, and $u \in \mathcal{U}$. Hence, $\mathcal{V} = \mathcal{V}_w$. Clearly $\mathcal{V}$ is a convex set. Given any u satisfying $u \wedge \xi \in \mathcal{V} \; \forall \xi \in \mathcal{X}$, we define u' such that for $j = 1, ..., n$,

$$u'_j = \begin{cases} u_j, & \text{if} \quad u_j \leq \bar{\xi}_j, \\ \bar{\xi}_j, & \text{if} \quad u_j > \bar{\xi}_j. \end{cases}$$

One can easily check that $u' \wedge \xi = u \wedge \xi \; \forall \xi \in \mathcal{X}$. We only need to show $u' \in \mathcal{U}$. Since $\bar{\xi}_j \geq \underline{u}_j$ and $u_j \geq \underline{u}_j$, we have $u'_j \geq \underline{u}_j$ for $j = 1, ..., n$. Because $A(u \wedge \xi) \leq b \; \forall \xi \in \mathcal{X}$ and Ξ has independent components, we obtain $A(u \wedge \bar{\xi}) \leq b$, which is the same as $Au' \leq b$.

If $\bar{\xi}_j = \infty$ for any j, then $u'_j = u_j$ and following similar arguments we can obtain the desired results. $\square$

Notice that in our proof the assumption Ξ has independent components are needed. To see this, suppose that Ξ can be either $\xi_1 = 4, \xi_2 = 2$ or $\xi_1 = 2, \xi_2 = 4$ with equal probability. Let $A = [1, 1]$, $b = 6$ and $u = (3, 5)$. According to the construction in our proof we have $u' = (3, 4)$. Then $A(u \wedge \xi) \leq b \; \forall \xi \in \mathcal{X}$ but $Au' > b$.

A.5 Proof of Proposition 2.1

Parts (a)-(c) are from Murota (2003). The proofs follow directly from the definition.

(d) We need to show that $g[(x,\lambda)-\xi(e,1)]$ is submodular. Notice that $g[(x,\lambda)-\xi(e,1)] = g(v-\xi e, \lambda-\xi) = f((v-\xi e)-(\lambda-\xi)e) = f(v-\lambda e)$, which is submodular.

(e) We assume without loss of generality that $\mathcal{A} = \mathcal{F}^n \times \mathcal{Y}$; otherwise we can focus on the restriction of f on $\mathcal{A}$ and let f be infinity outside of $\mathcal{A}$. We know that $r(x,y,\xi) = f[(x,y)-\xi e]$ is submodular, and we want to show that $g(x-\xi e)$ is submodular in (x,ξ). We have

$$\begin{aligned} g(x-\xi e) &= \inf_{y\in\mathcal{Y}} f(x-\xi e, y) \\ &= \inf_{y\in\mathcal{Y}} f[(x, y+\xi e)-\xi e] \\ &= \inf_{y\in\mathcal{Y}} r(x, y+\xi e, \xi) \\ &= \inf_{z-\xi e\in\mathcal{Y}} r(x,z,\xi). \end{aligned}$$

Notice that $\{(z,\xi) : z-\xi e \in \mathcal{Y}\}$ is a lattice and $r(x,z,\xi)$ is submodular, it follows from Theorem 2.7.6 of Topkis (1998) that $g(x-\xi e)$ is submodular in (x,ξ).

(f) It follows from Theorem 2.8.2 of Topkis (1998) that $\arg\min_{y:(x,y)\in\mathcal{A}} f(x,y)$ is increasing in x. For any $\omega > 0$ and any $x \in dom(g)$, define $\omega\tilde{e} + \arg\min_{y:(x,y)\in\mathcal{A}} f(x,y)$ as the set $\{u+\omega\tilde{e} : u \in \arg\min_{y:(x,y)\in\mathcal{A}} f(x,y)\}$. Pick y' in $\arg\min_{y:(x,y)\in\mathcal{A}} f(x,y)$ and y'' in $\arg\min_{y:(x+\omega e,y)\in\mathcal{A}} f(x+\omega e,y)$. Then for any $\omega > 0$ such that $(x+\omega e, y'+\omega\tilde{e}) \in \mathcal{A}$ and $(x, y''-\omega\tilde{e}) \in \mathcal{A}$ we have $(x+\omega e, y'' \wedge (y'+\omega\tilde{e})) = (x+\omega e, y'') \wedge (x+\omega e, y'+\omega\tilde{e}) \in \mathcal{A}$, $(x, (y''-\omega\tilde{e}) \vee y') = (x, y''-\omega\tilde{e}) \vee (x,y') \in \mathcal{A}$ and

$$\begin{aligned} 0 &\geq f(x+\omega e, y'') - f(x+\omega e, y'' \wedge (y'+\omega\tilde{e})) \\ &= f[(x, y''-\omega\tilde{e})+\omega(e,\tilde{e})] - f[(x,(y''-\omega\tilde{e})\wedge y')+\omega(e,\tilde{e})] \\ &\geq f(x, y''-\omega\tilde{e}) - f(x,(y''-\omega\tilde{e})\wedge y') \\ &\geq f(x,(y''-\omega\tilde{e})\vee y') - f(x,y') \\ &\geq 0, \end{aligned}$$

where the first and the last inequalities are due to the optimality of y'' and y' for $x+\omega e$ and x respectively, the second inequality is due to the $L^\natural$-convexity of f which implies that $f(x-\omega e, y-\omega\tilde{e})$ is submodular in (x,y,ω), and the third inequality is due to the submodularity of $f(x, y)$ in y. The first and the last inequalities then imply that equality holds throughout the above inequalities and so $y'' \wedge (y' + \omega\tilde{e}) \in \arg\min_{y:(x+\omega e,y)\in\mathcal{A}} f(x+\omega e, y)$, and $(y'' - \omega\tilde{e}) \vee y' \in \arg\min_{y:(x,y)\in\mathcal{A}} f(x, y)$ which then implies that $y'' \vee (y' + \omega\tilde{e}) \in \omega\tilde{e} + \arg\min_{y:(x,y)\in\mathcal{A}} f(x, y)$. Therefore,

$$\arg\min_{y:(x+\omega e,y)\in\mathcal{A}} f(x + \omega e, y) \sqsubseteq \omega\tilde{e} + \arg\min_{y:(x,y)\in\mathcal{A}} f(x, y).$$

(g) Let $\mathcal{A} = \{x \in \mathcal{Y} : l \leq x \leq u, x_i - x_j \leq v_{ij}, \forall i \neq j\}$. For any $x, x' \in \mathcal{A}, \lambda \in \mathcal{F}_+$, we only need to show that $(x + \lambda e) \wedge x', x \vee (x' - \lambda e) \in \mathcal{A}$. Firstly we have $(x+\lambda e) \wedge x' \leq x' \leq u$, and $l \leq x \wedge x' \leq (x+\lambda e) \wedge x'$. For any $i \neq j$, if $x'_i \leq x_i+\lambda, x'_j \leq x_j+\lambda$, then $(x_i+\lambda)\wedge x'_i - (x_j+\lambda)\wedge x'_j = x'_i - x'_j \leq v_{ij}$. If $x'_i \geq x_i + \lambda, x'_j \geq x_j + \lambda$, then $(x_i + \lambda) \wedge x'_i - (x_j + \lambda) \wedge x'_j = x_i - x_j \leq v_{ij}$. If $x'_i \leq x_i + \lambda, x_j + \lambda \leq x'_j$, then $(x_i + \lambda) \wedge x'_i - (x_j + \lambda) \wedge x'_j = x'_i - (x_j + \lambda) \leq x_i + \lambda - (x_j + \lambda) \leq v_{ij}$. If $x'_i \geq x_i + \lambda, x_j + \lambda \geq x'_j$, then $(x_i + \lambda) \wedge x'_i - (x_j + \lambda) \wedge x'_j = (x_i + \lambda) - x'_j \leq x'_i - x'_j \leq v_{ij}$. Thus we have $(x + \lambda e) \wedge x' \in \mathcal{A}$. Similarly we can show that $x \vee (x' - \lambda e) \in \mathcal{A}$.

(h) Please see Proposition 2.3.3 (b) of Simchi-levi et al. (2014).

A.6 Proof of Theorem 2.3

Theorem 2.2 implies that problem (2.7) can be equivalently converted to the following one:

$$\begin{array}{ll} \inf & E[f(x, v_1(\Xi_1), \ldots, v_n(\Xi_n))] \\ \text{s.t.} & v_j(\xi_j) \leq z_j + \xi_j \quad \forall \xi_j \in \mathcal{X}_j, \ \forall\ j = 1, \ldots, k \\ & v_j(\xi_j) \geq z_j + \xi_j \quad \forall \xi_j \in \mathcal{X}_j, \ \forall\ j = k+1, \ldots, n \\ & (x, z, v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathcal{A}^{\Xi} \quad \forall \xi \in \mathcal{X}. \end{array} \tag{A.3}$$

To see this, given fixed (x, z), let $\mathcal{U}(x, z)$ denote the constraint set $\{u : (x, z, u) \in \mathcal{A}\}$, $f_x(u) = f(x, u)$, $\tilde{\Xi}_z = z + \Xi$, and $\tilde{\mathcal{X}}_z = Supp(\tilde{\Xi}_z)$. Then (2.7) is equivalent to

$$\inf_{u\in\mathcal{U}(x,z)} E[f_x(u\Diamond_k\tilde{\Xi}_z)]. \tag{A.4}$$

Let $\mathcal{V}(x, z) = \{u\Diamond_k\tilde{\xi} : u \in \mathcal{U}(x, z), \tilde{\xi} \in \tilde{\mathcal{X}}_z\}$. Given any $u\Diamond_k\tilde{\xi} \in \mathcal{V}(x, z)\ \forall\tilde{\xi} \in \tilde{\mathcal{X}}_z$, we have (x, z, u) satisfying $(x, z, u\Diamond_k(z + \xi)) \in \mathcal{A}^{\Xi}\ \forall\xi \in \mathcal{X}$. According to Assumption 2, there exists $(x, z, u') \in \mathcal{A}$ such that $u' \Diamond_k (z+\xi) = u \Diamond_k (z+\xi)$ $\forall \xi \in \mathcal{X}$. Thus we have $u' \in \mathcal{U}(x, z)$ and $u' \Diamond_k \tilde{\xi} = u \Diamond_k \tilde{\xi}$ $\forall \tilde{\xi} \in \tilde{\mathcal{X}}_z$. If the indicator function of $\mathcal{A}^{\Xi}$ is componentwise convex in w, it is clear that the indicator function of $\mathcal{V}(x, z)$ is also componentwise convex. Therefore, if Assumption 2 is satisfied, then Assumption 1 is also satisfied. According to Theorem 2, we can transform (A.4) into:

$$
\begin{aligned}
\inf \quad & E[f_x(v_1(\tilde{\Xi}_{z1}), \ldots, v_n(\tilde{\Xi}_{zn}))] \\
\text{s.t.} \quad & v_j(\tilde{\xi}_j) \le \tilde{\xi}_j \quad \forall \tilde{\xi}_j \in \tilde{\mathcal{X}}_{zj}, \ \forall\ j = 1, \ldots, k \\
& v_j(\tilde{\xi}_j) \ge \tilde{\xi}_j \quad \forall \tilde{\xi}_j \in \tilde{\mathcal{X}}_{zj}, \quad \forall\ j = k+1, \ldots, n \\
& (v_1(\tilde{\xi}_1), \ldots, v_n(\tilde{\xi}_n)) \in \mathcal{V}(x, z) \quad \forall \tilde{\xi} \in \tilde{\mathcal{X}}_z,
\end{aligned}
$$

which is equivalent to (A.3).

It is straightforward to check that the constraint set involving $(x, z, (v_1(\xi_1), \ldots, v_n(\xi_n))_{\xi \in \mathcal{X}}$ is a convex set, a lattice, and an $L^{\natural}$-convex set (Proposition 2.1 part (g)) on the product set $\mathcal{F}^m \times \mathcal{F}^n \times (\mathcal{F}^n)^{\mathcal{X}}$ for cases (a), (b) and (c) respectively.

In the following we show that the objective function $E[f(x, v_1(\Xi_1), ..., v_n(\Xi_n))]$ is convex, submodular, $L^{\natural}-$convex in $(x, (v_1(\xi_1), ..., v_n(\xi_n))_{\xi \in \mathcal{X}} \in \mathcal{F}^m \times (\mathcal{F}^n)^{\mathcal{X}}$ for cases (a), (b) and (c) respectively. Define $\tilde{f} : \mathcal{F}^m \times (\mathcal{F}^n)^{\mathcal{X}} \times \mathcal{X} \to \bar{\Re}$ such that $\tilde{f}(x, v, \xi) \triangleq f(x, v(\xi))$ $\forall \xi \in \mathcal{X}$. Clearly, $E[\tilde{f}(x, v, \Xi)] = E[f(x, v(\Xi))]$. Given any realization ξ, if $f(\cdot, \cdot)$ is convex/ submodular/ $L^{\natural}-$convex, then one can easily prove by definition that $\tilde{f}(\cdot, \cdot, \xi)$ is also convex/ submodular/ $L^{\natural}-$convex. We show the proof for convexity; the proofs for submodularity and $L^{\natural}-$convex are similar and simply follow their definitions respectively. Given ξ, for any (x, v), (x', v') and $\lambda \in [0, 1]$, we have

$$
\begin{aligned}
& \tilde{f}(\lambda x + (1-\lambda)x', \lambda v + (1-\lambda)v', \xi) \\
& = f(\lambda x + (1-\lambda)x', \lambda v(\xi) + (1-\lambda)v'(\xi)) \\
& \le \lambda f(x, v(\xi)) + (1-\lambda) f(x', v'(\xi)) \\
& = \lambda \tilde{f}(x, v, \xi) + (1-\lambda)\tilde{f}(x', v', \xi).
\end{aligned}
$$

Since $\tilde{f}(\cdot, \cdot, \xi)$ is convex/ submodular/ $L^{\natural}-$convex for any given ξ, we have that the objective function $E[f(x, v(\Xi))] = E[\tilde{f}(x, v, \Xi)]$ is also convex/ submodular/ $L^{\natural}-$convex due to Proposition 2.1 part (c).

Part (a) follows immediately from the theorem of convexity preservation under minimization (see Simchi-levi et al. 2014, Proposition 2.1.15, for the case with finite-dimensional spaces, and Zălinescu 2002, Theorem 2.1.3(v), for the case with general vector spaces). Part (b) follows from Theorem 2.7.6 of Topkis (1998).

Part (c) follows from Proposition 2.1 part (e). □

A.7 Proof of Theorem 2.4

In the following, we provide the proof for part (b). Since part (a) can be proved using almost the same arguments (as $L^\natural$-convexity includes submodularity), its proof is omitted for brevity.

Let $\tilde{\mathcal{V}}(x,z)$ denote the constraint set of the transformed problem (A.3). Define the projection of the solution set of the transformed problem, $\mathcal{S}^*(x,z) = \arg\min_{(v(\xi),\xi\in\mathcal{X})\in\tilde{\mathcal{V}}(x,z)} E[f(x,v(\Xi))]$, on the constraint set $\mathcal{U}(x,z)$ as

$$\Pi_{\mathcal{U}}\mathcal{S}^*(x,z) = \{u \in \mathcal{U}(x,z) | (u\Diamond_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)\}.$$

By Proposition 2.1 we know that $\mathcal{S}^*(x,z)$ is increasing in (x,z) and satisfies the monotone sensitivity property with respect to (x,z) as follows:

$$\mathcal{S}^*((x,z)+\omega e) \sqsubseteq \omega e + \mathcal{S}^*(x,z).$$

We argue that $\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ is the solution set of the original problem for any given (x,z), i.e., $\Pi_{\mathcal{U}}\mathcal{S}^*(x,z) = \mathcal{U}^*(x,z)$. In fact, if u^* is an optimal solution to the original problem, $(u^*\Diamond_k(z+\xi), \xi \in \mathcal{X})$ is a minimizer of the transformed problem, i.e., $(u^*\Diamond_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$. On the other hand, if $u \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$, then $(v(\xi)|v(\xi) = u\Diamond_k(z+\xi), \xi \in \mathcal{X})$ is an optimal solution of the transformed problem due to the definition of $\Pi_{\mathcal{U}}\mathcal{S}^*$. Since $E[f(x,v(\Xi))] = E[f(x,u\Diamond_k(z+\Xi))] = \tau^*$, u is optimal for the original problem. Therefore, our argument is true, which implies that we only need to show that $\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ is increasing in (x,z) and $\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e) \sqsubseteq \omega e + \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$.

We firstly show that

$$\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e) \sqsubseteq \omega e + \Pi_{\mathcal{U}}\mathcal{S}^*(x,z). \tag{A.5}$$

Pick any u' in $\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ and any $u^{''}$ in $\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e)$ respectively. We have

$$(u'\Diamond_k(z+\xi),\xi\in\mathcal{X})\in\mathcal{S}^*(x,z),(u''\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\in\mathcal{S}^*((x,z)+\omega e).$$

It suffices to show that $u^{''}\wedge(u^{'}+\omega e)\in\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e)$ and $(u^{''}-\omega e)\vee u'\in\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$. Since $\mathcal{S}^*((x,z)+\omega e)\sqsubseteq\omega e+\mathcal{S}^*(x,z)$, we have

$$(u^{''}\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\wedge(u'\Diamond_k(z+\xi)+\omega e,\xi\in\mathcal{X})\in\mathcal{S}^*((x,z)+\omega e).$$

Hence,

$$\begin{aligned}&(u^{''}\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\wedge(u'\Diamond_k(z+\xi)+\omega e,\xi\in\mathcal{X})\\&=((u^{''}\Diamond_k(z+\omega e+\xi)\wedge((u'+\omega e)\Diamond_k(z+\omega e+\xi)),\xi\in\mathcal{X})\\&=((u''\wedge(u'+\omega e))\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\\&\in\mathcal{S}^*((x,z)+\omega e).\end{aligned}$$

Since $\mathcal{A}$ is an $L^\natural-$convex set, $(u',x,z)\in\mathcal{A}$, and $(u'',(x,z)+\omega e)\in\mathcal{A}$, we have $(u'',(x,z)+\omega e)\wedge((u',x,z)+\omega e)=(u''\wedge(u'+\omega e),(x,z)+\omega e)\in\mathcal{A}$. Here we use the following property of $L^\natural-$convex set (page 128 of Murota 2003): if $\mathcal{A}$ is an $L^\natural-$convex set, then for any $p,q\in\mathcal{A}$, we have $(p-\omega e)\vee q, p\wedge(q+\omega e)\in\mathcal{A}\ \forall\omega\geq 0$. Hence, $u''\wedge(u'+\omega)\in\mathcal{U}((x,z)+\omega e)$. Together with $((u''\wedge(u'+\omega e))\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\in\mathcal{S}^*((x,z)+\omega e)$ we obtain $u''\wedge(u'+\omega)\in\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e)$.

Similarly, since $\mathcal{S}^*((x,z)+\omega e)\sqsubseteq\omega e+\mathcal{S}^*(x,z)$, we have $((u''\vee(u'+\omega e))\Diamond_k(z+\omega e+\xi),\xi\in\mathcal{X})\in\mathcal{S}^*(x,z)+\omega e$. Hence, $((u''-\omega e)\vee u')\Diamond_k(z+\xi),\xi\in\mathcal{X})\in\mathcal{S}^*(x,z)$. Since $\mathcal{A}$ is an $L^\natural-$convex set, $(u',x,z)\in\mathcal{A}$, and $(u'',(x,z)+\omega e)\in\mathcal{A}$, we have $((u'',(x,z)+\omega e)-\omega e)\vee(u',x,z)=((u''-\omega e)\vee u',x,z)\in\mathcal{A}$. Hence, $(u''-\omega e)\vee u'\in\mathcal{U}(x,z)$. Therefore, $(u''-\omega e)\vee u'\in\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$. This completes the proof of the inequality (A.5).

In the following we show that for any i and $\omega>0$, we have

$$\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)\sqsubseteq\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e_i).$$

It suffices to show that $u'\wedge u''\in\Pi_{\mathcal{U}}\mathcal{S}^*(x,z), u'\vee u''\in\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e_i)$ for any $u'\in\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ and $u^{''}\in\Pi_{\mathcal{U}}\mathcal{S}^*((x,z)+\omega e_i)$.

If the increment ω is associated with a component of x, then we have $\mathcal{S}^*(x,z) \sqsubseteq \mathcal{S}^*(x+\omega e_i, z)$. Hence, $(u'\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \wedge (u''\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) = ((u' \wedge u'')\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$, and $(u'\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \vee (u''\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) = ((u' \vee u'')\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x+\omega e_i, z)$. Since $\mathcal{A}$ is a lattice, $(u', x, z) \in \mathcal{A}$, and $(u'', x+\omega e_i, z) \in \mathcal{A}$, we have $(u' \wedge u'', x, z) \in \mathcal{A}, (u' \vee u'', x+\omega e_i, z) \in \mathcal{A}$. Hence, $u' \wedge u'' \in \mathcal{U}(x,z), u' \vee u'' \in \mathcal{U}(x+\omega e_i, z)$. Therefore, $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z), u' \vee u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x+\omega e_i, z)$.

If the increment ω is associated with a component of z, we firstly show $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$. Applying the previous arguments, we have $u' \wedge u'' \in \mathcal{U}(x,z)$. We only need to show that $((u' \wedge u'')\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$.

Since $\mathcal{S}^*(x,z) \sqsubseteq \mathcal{S}^*(x, z+\omega e_i)$, we have

$$(u'\diamondsuit_k(z+\xi), \xi \in \mathcal{X}) \wedge (u''\diamondsuit_k(z+\omega e_i+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z).$$

Here we have the following observation.

Lemma A.1. *If $u_1 \le u_2, z_1 \le z_2$, then*

$$(u_1 \wedge z_1) \vee (u_2 \wedge z_2) = (u_1 \vee u_2) \wedge (z_1 \vee z_2).$$

Notice that if $u_i' \le u_i''$ or $i \le k$ (corresponding to the $\wedge$ operation), we have

$$(u_i'\diamond(z_i+\xi_i), \xi \in \mathcal{X}) \wedge (u_i''\diamond(z_i+\omega+\xi_i), \xi \in \mathcal{X}) = ((u_i' \wedge u_i'')\diamond(z_i+\xi_i), \xi \in \mathcal{X}),$$

where $\diamond$ denote a $\wedge$ or $\vee$ operation, which is the corresponding operation in $\diamondsuit_k$ for component i.

Therefore, it remains to consider the case where $u_i' > u_i''$ and $i > k$ (corresponding to the $\vee$ operation). Define $\tilde{v}(\xi) \triangleq (u'\diamondsuit_k(z+\xi)) \wedge (u''\diamondsuit_k(z+\omega e_i+\xi))$, we have

$$\begin{aligned} \tilde{v}_i(\xi_i) &= (u_i' \vee (z_i+\xi_i)) \wedge (u_i'' \vee (z_i+\omega+\xi_i)) \\ &= \begin{cases} z_i+\xi_i, & \text{if } \xi_i \ge u_i' - z_i, \\ u_i', & \text{if } u_i' - z_i - \omega \le \xi_i < u_i' - z_i, \\ z_i+\omega+\xi_i, & \text{if } u_i'' - z_i - \omega \le \xi_i < u_i' - z_i - \omega, \\ u_i'', & \text{if } \xi_i < u_i'' - z_i - \omega. \end{cases} \end{aligned}$$

We use $\tilde{v}_{-i}(\xi_{-i})$ to denote $(\tilde{v}_1(\xi_1), ..., \tilde{v}_{i-1}(\xi_{i-1}), \tilde{v}_{i+1}(\xi_{i+1}), ..., \tilde{v}_n(\xi_n))$.

Let π^* denote the optimal objective value of the transformed problem with parameters (x, z). Similar to the arguments in Theorem 2.2, let $\hat{f}(x, z, v) = f(x, v) + \delta_{\mathcal{V}(x,z)}(v)$, where $\mathcal{V}(x, z)$ denotes the constraint set $\{v(\xi)|(x, z, v(\xi)) \in \mathcal{A}^{\Xi}, \xi \in \mathcal{X}\}$. Following the proof in Theorem 2.1, define $\hat{g}(\cdot) = E[\hat{f}(x, z, \tilde{v}_{-i}(\Xi_{-i}), \cdot)]$. Since the function f is componentwise convex, lower semi-continuous with $f(u) \to +\infty$ for $|u| \to \infty$ and the constraint set is componentwise convex and closed, we have that $\hat{g}(\cdot)$ is convex, lower semi-continuous and $\hat{g}(u) \to +\infty$ for $|u| \to \infty$. Therefore, $\min_{u_i \in \mathcal{F}} \hat{g}(u_i)$ has a greatest and a least minimizer, and given any $\tilde{u}_i \in \arg\min_{u_i \in \mathcal{F}} \hat{g}(\cdot)$, we have

$$
\begin{aligned}
\pi^* &= \min\{E[\hat{g}(v_i(\Xi_i))]|v_i(\xi_i) \geq z_i + \xi_i, \forall \xi_i \in \mathcal{X}_i\} \\
&= \min_{u_i \in \mathcal{F}} E[\hat{g}(u_i \vee (z_i + \Xi_i))] \\
&= E[\hat{g}(\tilde{u}_i \vee (z_i + \Xi_i))].
\end{aligned}
$$

We argue that $u_i^{''} \in \arg\min_{u_i \in \mathcal{F}} \hat{g}(u_i)$. Let $\bar{u}_i$ and $\underline{u}_i$ denote the greatest and least minimizer respectively. We will show $\underline{u}_i \leq u_i'' \leq \bar{u}_i$. If, otherwise, $\bar{u}_i < u_i''$, then when $\xi_i < u_i' - z_i$ we have $\bar{u}_i \leq \bar{u}_i \vee (z_i + \xi_i) < \tilde{v}_i(\xi_i)$ and thus $\hat{g}(\bar{u}_i \vee (z_i + \xi_i)) < \hat{g}(\tilde{v}_i(\xi_i))$. By the assumption we have $\underline{\xi}_i + z_i + \omega \leq u_i'' < u_i'$, hence we know $Pr(\xi_i < u_i' - z_i) > 0$. When $\xi_i \geq u_i' - z_i$, we have $\hat{g}(\bar{u}_i \vee (z_i + \xi_i)) = \hat{g}(z_i + \xi_i) = \hat{g}(\tilde{v}_i(\xi_i))$. If $\underline{u}_i > u_i''$, then when $\xi_i < \underline{u}_i - z_i - \omega$, we have $\underline{u}_i \vee (z_i + \xi_i) > \tilde{v}_i(\xi_i)$ and thus $\hat{g}(\underline{u}_i \vee (z_i + \xi_i)) < \hat{g}(\tilde{v}_i(\xi_i))$. Since $\underline{\xi}_i + \omega + z_i \leq u_i'' < \underline{u}_i$, we have $Pr(\xi_i < \underline{u}_i - z_i - \omega) > 0$. In the following we show that when $\xi_i \geq \underline{u}_i - z - \omega$, we have $\hat{g}(\underline{u}_i \vee (z_i + \xi_i)) \leq \hat{g}(\tilde{v}_i(\xi_i))$. When $\underline{u}_i - z_i - \omega \leq \xi_i < \underline{u}_i - z_i$, we have $\hat{g}(\underline{u}_i \vee (z_i + \xi_i)) = \hat{g}(\underline{u}_i) \leq \hat{g}(\tilde{v}_i(\xi_i))$; when $\xi_i \geq \underline{u}_i - z_i$, we have $\hat{g}(\underline{u}_i \vee (z_i + \xi_i)) = \hat{g}(z_i + \xi_i) \leq \hat{g}((u_i' \vee (z_i + \xi_i)) \wedge (z_i + \omega + \xi_i)) = \hat{g}(\tilde{v}_i(\xi_i))$. Therefore, for $\bar{u}_i < u_i''$ or $\underline{u}_i > u_i''$, we have $\pi^* < E[\hat{g}(\tilde{v}_i(\Xi_i))]$, which contradicts the hypothesis that $(\tilde{v}(\xi), \xi \in \mathcal{X})$ is an optimal solution to the transformed problem. Hence, $\pi^* = \min_{u_i \in \mathcal{F}} E[\hat{g}(u_i'' \vee (z_i + \Xi_i))]$ and $((u_i' \wedge u_i'') \vee (z_i + \xi_i), \tilde{v}_{-i}(\xi_{-i}), \xi \in \mathcal{X}) \in \mathcal{S}^*(x, z)$. Since for any $j \neq i$, $\tilde{v}_j(\xi_i) = (u_j' \wedge u_j'') \diamond (z_j + \xi_j), \forall \xi \in \mathcal{X}$, we have $((u' \wedge u'') \Diamond_k (z + \xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x, z)$.

It follows a similar logic to show that $u' \vee u'' \in \Pi_{\mathcal{U}} \mathcal{S}^*(x, z + \omega e_i)$, i.e., $((u' \vee u^{''}) \Diamond_k (z + \omega e_i + \xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x, z + \omega e_i)$. We omit the details for brevity.

In the following we show that the optimal solution set $\mathcal{U}^*(x, z)$ is a lattice, and it has a greatest element and a least element. Given fixed (x, z), let $h(u) = E[f(x, u\diamondsuit_k(z + \Xi))]$. For any realization of Ξ, denoted by ξ, given any u' and u'', we have

$$
\begin{aligned}
& f(x, u'\diamondsuit_k(z+\xi)) + f(x, u''\diamondsuit_k(z+\xi)) \\
= \; & f(x, u'_1 \wedge (z_1+\xi_1), ..., u'_n \vee (z_n+\xi_n)) \\
& + f(x, u''_1 \wedge (z_1+\xi_1), ..., u''_n \vee (z_n+\xi_n)) \\
\geq \; & f(x, (u'_1 \wedge (z_1+\xi_1)) \wedge (u''_1 \wedge (z_1+\xi_1)), ..., \\
& (u'_n \vee (z_n+\xi_n)) \wedge (u''_n \vee (z_n+\xi_n))) \\
& + f(x, (u'_1 \wedge (z_1+\xi_1)) \vee (u''_1 \wedge (z_1+\xi_1)), ..., \\
& (u'_n \vee (z_n+\xi_n)) \vee (u''_n \vee (z_n+\xi_n))) \\
= \; & f(x, (u'_1 \wedge u''_1) \wedge (z_1+\xi_1), ..., (u'_n \wedge u''_n) \vee (z_n+\xi_n)) \\
& + f((u'_1 \vee u''_1) \wedge (z_1+\xi_1), ..., (u'_n \vee u''_n) \vee (z_n+\xi_n)).
\end{aligned}
$$

The inequality is due to the submodularity of f. Then $h(u') + h(u'') \geq h(u' \wedge u'') + h(u' \vee u'')$. Therefore, the objective function of (2.7) is submodular in u. By our assumptions the function f satisfies $f(x) \to +\infty$ for $|x| \to \infty$, and the constraint set is closed and is a lattice, it is equivalent to restricting our constraint set to a compact sublattice of $\Re^n$. By Corollary 2.7.1 of Topkis (1998) the solution set $\mathcal{U}^*(x, z)$ is a compact sublattice of $\Re^n$, and there exist a greatest element and a least element in the solution set. $\square$

Remark A.1. *Firstly, even though we know that if f is submodular then $E[f(u \wedge \Xi)]$ is submodular (see Proposition A.1), unfortunately $E[f(u \wedge (z + \Xi))]$ may not be submodular in (u, z). To see this, we provide an example here. Let $f(x) = x$, and $\Xi = 0$ with probability 1. Choose $u = 3, u' = 2, z = 1, z' = 4$. We have*

$$
\begin{aligned}
& f(u \wedge z) + f(u' \wedge z') \\
&= f(1) + f(2) \\
&< f(1) + f(3) \\
&= f((u \wedge u') \wedge (z \wedge z')) + f((u \vee u') \wedge (z \vee z')).
\end{aligned}
$$

Therefore, the objective function of problem (11) may not be submodular in (x, z, u).

A.8 A Comparison to the Stochastic Linearity Approach

Feng and Shanthikumar, (hereafter referred to as FS, 2016) use the notion of *stochastic linearity in mid-point* to develop a different technique to show that a class of nonlinear supply and demand functions (in the almost sure sense) are in fact linear in the stochastic sense. Like ours, their approach allows them to convert some non-convex minimization problems, including those in Ciarallo et al. (1994), Wang & Gerchak (1996), Feng (2010) and Feng & Shi (2012), into convex minimization problems. Treating the means of the supply and demand functions as decision variables instead of the original decisions (ordering quantity and price), they show that supply and demand functions are stochastically linear in mid point with respect to their means and the objective functions are concave in the means of supply and demand. Note that they focus on the concavity property but do not touch upon supermodularity or $L^\natural-$convavity. Different from their approach, our approach works on the original decision variables and transforms the original optimization problem into an equivalent constrained optimization problem, which allows us to readily show the preservation of convexity, submodularity and $L^\natural$-convexity.

Here we provide a detailed comparison between our transformation technique and their approach. In particular, we show that although their approach can also preserve convexity and submodularity, it does not preserve $L^\natural-$convexity.

We start by a brief introduction about FS's method. FS also consider optimization problems with objective function $E[f(u \wedge \Xi)]$ and convert them to convex minimization problems using stochastic linearity in mid-point (SL(mp)). Specifically, given a stochastic function $Y(u) \triangleq \psi(u, \Xi)$, let $\mu(u) = E[\psi(u, \Xi)]$, and $u(\mu)$ be the inverse of $\mu(u)$, i.e., $u(\mu) = \inf\{u|E[\psi(u, \Xi)] \geq \mu\}$. Then $g(\mu) \triangleq E[f(Y(u(\mu)))]$ is convex in μ as long as $Y(u(\mu))$ is SL(mp). FS prove that, along with several other supply functions, if $\psi(u, \Xi) = u \wedge \Xi$, then $Y(u(\mu))$ is SL(mp). This allows them to convert non-convex minimization problems to equivalent convex minimization problems by a variable transformation.

Our transformation technique can preserve convexity as well as submodularity and $L^\natural-$convexity. FS do not mention whether their approach can preserve submodularity or $L^\natural-$convexity. It turns out that their approach can preserve submodularity of the objective function, but can *not* preserve $L^\natural-$convexity. These are shown in the following Proposition A.1 and Example A.1 respectively.

Proposition A.1. *Suppose that $f : \Re^n \to \bar{\Re}$ is a submodular function, and Ξ is a random vector with support $\mathcal{X} \in \Re^n$, in which any component Ξ_i is independent of each other. Let $\mu(u) = E[u \wedge \Xi]$ and $u(\mu)$ be the inverse of $\mu(u)$. Then*

(a) $h(u) = E(f(u \wedge \Xi))$ is submodular.

(b) $g(\mu) = E(f(u(\mu) \wedge \Xi))$ is submodular.

Proof. (a) For any realization ξ, given any u and u', we have $f(u\wedge\xi)+f(u'\wedge\xi) \geq f((u\wedge\xi)\wedge(u'\wedge\xi))+f((u\wedge\xi)\vee(u'\wedge\xi)) = f((u\wedge u')\wedge\xi)+f((u\vee u')\wedge\xi)$.

(b) Notice that for any component $i = 1, ..., n$, $u_i(\mu_i)$ is increasing. It follows from section 9.A.4 of Shaked & Shanthikumar (2006) and part (a) that $g(\mu)$ is also submodular. $\square$

Example A.1. *Consider $E[f(u_1 \wedge \Xi_1, u_2 \wedge \Xi_2)]$, where $f(u_1, u_2) = e^{u_1-u_2}$ is an $L^\natural-$convex function. Suppose that both Ξ_1 and Ξ_2 follow exponential distribution with mean 1, and they are independent of each other. Then $\forall i = 1, 2, \mu_i(u_i) = E[u_i \wedge \Xi_i] = \int_0^{u_i} \xi_i e^{-\xi_i} d\xi_i + \int_{u_i}^{\infty} u_i e^{-\xi_i} d\xi_i = 1 - e^{-u_i}$. We have $u_i(\mu_i) = -\ln(1-\mu_i)$ and*

$$
\begin{aligned}
&E[f(u_1 \wedge \Xi_1, u_2 \wedge \Xi_2)] \\
&= \int_0^{u_2}\int_0^{u_1} e^{\xi_1-\xi_2}e^{-\xi_1}e^{-\xi_2}d\xi_1 d\xi_2 + \int_0^{u_2}\int_{u_1}^{\infty} e^{u_1-\xi_2}e^{-\xi_1}e^{-\xi_2}d\xi_1 d\xi_2 \\
&+ \int_{u_2}^{\infty}\int_0^{u_1} e^{\xi_1-u_2}e^{-\xi_1}e^{-\xi_2}d\xi_1 d\xi_2 + \int_{u_2}^{\infty}\int_{u_1}^{\infty} e^{u_1-u_2}e^{-\xi_1}e^{-\xi_2}d\xi_1 d\xi_2 \\
&= \frac{1}{2}(1+u_1)(1+e^{-2u_2}) \\
&= \frac{1}{2}(1-\ln(1-\mu_1))(1+(1-\mu_2)^2)
\end{aligned}
$$

Define $g(\mu_1, \mu_2) = \frac{1}{2}(1-\ln(1-\mu_1))(1+(1-\mu_2)^2)$. Let $\mu = [0.7, 0.2], \mu' = [0.8, 0.4], \alpha = 0.1$. We have $g(\mu)+g(\mu') \approx 3.5817$ while $g((\mu+\alpha e)\wedge\mu')+g(\mu\vee(\mu'-\alpha e)) \approx 3.5860$. Therefore,

$$g(\mu)+g(\mu') < g((\mu+\alpha e)\wedge\mu')+g(\mu\vee(\mu'-\alpha e)),$$

which means that $g(\mu_1, \mu_2)$ is not $L^\natural-$convex.

Notice that the approach from FS requires computing the inverse of $\mu(u)$, which may not have a closed form solution. If we consider a constrained optimization problem, even if all the constraints in the original problem are linear, the approach from FS will very likely add non-linear constraints explicitly. However, our transformation technique only adds linear constraints though potentially infinite number of them. More importantly, under the conditions in Lemma 2.2, the constraint set can also preserve $L^\natural-$convexity with our transformation technique, but this may not hold using the approach in FS. We illustrate this in the following example.

Example A.2. *Consider $\inf_{u\in\mathcal{U}} E[f(u_1 \wedge \Xi_1, u_2 \vee \Xi_2)]$, where $f(\cdot,\cdot)$ is an $L^\natural-$convex function and $\mathcal{U} = \{(u_1, u_2)|u_1 - u_2 \le \frac{1}{2}, 0 \le u_1 \le 1, 0 \le u_2 \le 1\}$. Suppose Ξ_1 and Ξ_2 are both uniformly distributed between 0 and 1, and they are independent of each other. Applying our transformation technique, we have*

$$
\begin{aligned}
\inf \quad & E[f(v_1(\Xi_1), v_2(\Xi_2))] \\
s.t. \quad & v(\xi_1) \le \xi_1 \ \forall \xi_1 \in [0,1], \\
& v_2(\xi_2) \ge \xi_2 \ \forall \xi_2 \in [0,1], \\
& (v_1(\xi_1), v_2(\xi_2)) \in \mathcal{V} \ \forall \xi \in [0,1]\times[0,1],
\end{aligned}
\tag{A.6}
$$

where $\mathcal{V} = \{(v_1, v_2)|v_1 - v_2 \le \frac{1}{2}, 0 \le v_1 \le 1, 0 \le v_2 \le 1\}$. All constraints in the transformed problem are linear, and they form an $L^\natural-$convex set.

Next we apply the transformation of FS. We have

$$\mu_1(u_1) = E[u_1 \wedge \Xi_1] = u_1 - \frac{1}{2}u_1^2,$$

$$\mu_2(u_2) = E[u_2 \vee \Xi_2] = \frac{1}{2}(1 + u_2^2).$$

Then we have

$$u_1(\mu_1) = 1 - \sqrt{1 - 2\mu_1},$$

$$u_2(\mu_2) = \sqrt{2\mu_2 - 1}.$$

Hence, the constraint set after the transformation becomes

$$\tilde{\mathcal{U}} = \{(\mu_1, \mu_2)|\sqrt{1 - 2\mu_1} + \sqrt{2\mu_2 - 1} \ge \frac{1}{2}, 0 \le \mu_1 \le \frac{1}{2}, \frac{1}{2} \le \mu_2 \le 1\},$$

which consists of non-linear constraints. One can also check that $\tilde{\mathcal{U}}$ *is not an* $L^\natural-$*convex set. To see this, notice that* $\mu = [0.3, 0.5] \in \tilde{\mathcal{U}}$ *and* $\mu' = [0.41, 0.51] \in \tilde{\mathcal{U}}$*, but* $\mu \vee (\mu' - \alpha e) \notin \tilde{\mathcal{U}}$ *with* $\alpha = 0.01$.

Appendix B

B.1 Proof of Theorem 3.1

For notational brevity, we omit the superscript k in the proof when there is no ambiguity. Define u_1, u_2 as the inventory level after the sales assuming that *the firm can hold inventory with some demand unsatisfied*. Then the realized sales are given by $w_1 - u_1$ and $w_2 - u_2$ at the two facilities respectively. By letting $\tilde{u}_2 = -u_2$, we claim that $\tilde{J}(w_1, \tilde{w}_2, v)$ equals the optimal objective value of the following problem:

$$\begin{aligned}
&\max_{u_1, \tilde{u}_2} r_1(w_1 - u_1) - r_2(\tilde{w}_2 - \tilde{u}_2) - h_1 u_1 + h_2 \tilde{u}_2 \\
&- s_1 v^+ - s_2(-v)^+ + \alpha \tilde{G}_*^{k-1}(u_1, \tilde{u}_2) \\
&s.t. \quad 0 \leq u_1, u_1 - w_1 \leq 0, \\
&\qquad\ \tilde{w}_2 - \tilde{u}_2 \leq 0, \tilde{u}_2 \leq 0, \\
&\qquad\ w_1 - u_1 \leq d_1, u_2 - \tilde{w}_2 \leq d_2.
\end{aligned} \tag{B.1}$$

Note that facing a stationary system, the firm should never hold inventory and reject demand at the same time since it is always more profitable to satisfy the current demand than holding the inventory to fulfill future demands. Therefore, the optimal solution is $u_1 = (w_1 - d_1)^+$, $u_2 = -(-\tilde{w}_2 - d_2)^+$ and our claim is correct.

We further claim that the objective function of the problem (B.1) is $L^\natural$-concave in $(w_1, \tilde{w}_2, v, u_1, \tilde{u}_2)$. To see this, note that $\tilde{G}_*^{k-1}(u_1, \tilde{u}_2)$ is $L^\natural$-concave by our induction hypothesis. The $L^\natural$-concavity of the rest of terms in the objective function is straightforward to verify. The constraint set is $L^\natural$-convex according to Proposition 2.1 part (g). Then the $L^\natural$-concavity of $\tilde{J}(w_1, \tilde{w}_2, v)$ follows from Proposition 2.1 part (e).

Note that the objective function in (3.5) is separable in variables $(w_1, \tilde{w}_2)$

and (v, v). Thus, the $L^\natural$-concavity of $\tilde{J}^k_*(q_1, \tilde{q}_2)$ follows from Proposition 3.1.

By defining $\tilde{G}(y_1, y_2) = E_{D_1,D_2}\{-c_1y_1 + c_2y_2 + \tilde{J}^k_*(y_1, y_2)\}$, (3.4) can be expressed as

$$\tilde{G}_*(x_1, \tilde{x}_2) = \max_{y_1 \geq x_1, \tilde{y}_2 \leq \tilde{x}_2} E_{T_1,\tilde{T}_2}\{\tilde{G}(y_1 \wedge (x_1 + T_1), \tilde{y}_2 \vee (\tilde{x}_2 + \tilde{T}_2))\} + c_1x_1 - c_2\tilde{x}_2$$

Clearly $\tilde{G}(y_1, y_2)$ is $L^\natural$-concave in (y_1, y_2). Moreover, $y_1 \wedge (x_1 + T_1) = (y_1 - x_1) \wedge T_1 + x_1$ and $\tilde{y}_2 \vee (\tilde{x}_2 + \tilde{T}_2) = (\tilde{y}_2 - \tilde{x}_2) \vee \tilde{T}_2 + \tilde{x}_2$. It is easy to see that by transforming the variables $\hat{y}_1 = y_1 - x_1$ and $\hat{y}_2 = \tilde{y}_2 - \tilde{x}_2$, the above problem can be expressed in the form of (2.7). Then Proposition 2.3 implies that the profit-to-go function $\tilde{G}_*(x_1, \tilde{x}_2)$ is $L^\natural$-concave. $\square$

B.2 Proof of Theorem 3.2

The proof is by induction. Suppose that v_{t+1} is $L^\natural$-convex. By Proposition 2.1 (d), for any d_t, $v_{t+1}[(z_2, ..., z_k, y) - (d_t + u)e]$ is $L^\natural$-convex in (z, y, u) and so is $\alpha v_{t+1}[(z_2, ..., z_k, y) - (d_t + u)e]$. Clearly all the other terms of g_t are $L^\natural$-convex in (z, y, u) (That's why we define $u = -q_E$ and $\tilde{k}_{e,t} = -k_{E,t}$). Thus, g_t is $L^\natural$-convex in (z, y, u). Let $\mathcal{A} = \{(z, y, u)|y \geq z_k, u \leq 0\}$ and

$$\mathcal{A}^\Xi = \{(z, y \wedge (z_k + k_{R,t}), u \vee \tilde{k}_{e,t})|y \geq z_k, u \leq 0, k_{R,t} \in Supp(K_{R,t}), \tilde{k}_{e,t} \in Supp(\tilde{K}_{E,t})\}.$$

Since $K_{R,t} \geq 0$ and $\tilde{K}_{E,t} \leq 0$ almost surely, it is easy to see that

$$\mathcal{A}^\Xi = \{(z, w_1, w_2)|z_k + k^u_{R,t} \geq w_1 \geq z_k, \tilde{k}^l_{E,t} \leq w_2 \leq 0\},$$

where $k^u_{R,t} = \operatorname{ess\,sup} Supp(K_{R,t})$ and $\tilde{k}^l_{E,t} = \operatorname{ess\,inf} Supp(\tilde{K}_{E,t})$.

The constraint set $\mathcal{A}^\Xi$ forms an $L^\natural$-convex set because of Proposition 2.1 (g). It is straightforward to see that the set $\mathcal{A} = \{(z, y, u)|y \geq z_k, u \leq 0\}$ is of the form in Lemma 2.3. Applying Theorem 2.3, we know $v_t(z)$ is $L^\natural$-convex in $z \in \mathcal{S}$. According to Theorem 2.4, the greatest optimal solution $(y_t(z), u_t(z))$ is well defined and has the desired monotonicity property with limited sensitivity. $\square$

B.3 Derivation of Inequalities (3.10)

The state of the system is $(z_1, ..., z_k)$, which is the cumulated inventory level. Decisions are $y = z_k + q_r$ and $u = -q_e$. The $L^\natural-$convexity gives us $y(z) \leq y(z+\omega e) \leq y(z)+\omega$, $u(z) \leq u(z+\omega e) \leq u(z)+\omega$. Let $x_1 = z_1, x_2 = z_2-z_1, ...,$ $\hat{q}_r(x) = q_r(z)$, we only need to show that

$$-\omega \leq \hat{q}_r(x + \omega e_k) - \hat{q}_r(x) \leq ... \leq \hat{q}_r(x + \omega e_1) - \hat{q}_r(x) \leq 0.$$

We know that

(a) if $e_k = 1$, $-\omega \leq q(z + \omega e) - q(z) \leq 0$,

(b) if $e_k = 0$, $0 \leq q(z + \omega e) - q(z) \leq \omega$.

Then (a) implies that

$$-\omega \leq \hat{q}_r(x + \omega e_i) - \hat{q}_r(x) \leq 0, \ \forall i = 1, ..., k.$$

And (b) implies that

$$\hat{q}_r(x + \omega e_i) - \hat{q}_r(x) \leq ... \leq \hat{q}_r(x + \omega e_j) - \hat{q}_r(x), \ \forall 1 \leq j \leq i \leq k.$$

B.4 Proof of Theorem 3.3

(a) We prove by induction. Assume that f_{t+1} is convex. It is easy to see that $g_t(z|d)$ in (3.13) is convex in z for any demand realization d since the objective function is jointly convex in (z, u) and the constraints form a convex set. Define $G_t(z) \triangleq c^T z + E[g_t(z|D)]$, which is convex in z. Then $f_t(x) = \min_{y \geq x} E[G_t(y \wedge (x + \Xi))] - c^T x$. The constraint set is $\mathcal{A} = \{(x, y)|y \geq x\}$. By definition, $\mathcal{A}^\Xi = \{(x, y \wedge (x + \xi))|y \geq x, \xi \in Supp(\Xi)\}$. This is equivalent to the set $\{(x, w)|x_i \leq w_i \leq x_i + \bar{\xi}_i, \forall i = 1, ..., m\}$, which is convex. In addition, Assumption 2.2 is satisfied since $\mathcal{A}$ is of the form given in Lemma 2.3. Therefore, following Theorem 2.3 we know that $f_t(x)$ is convex in x.

(b) For $j = 1, ..., m$, define $\hat{u}_j = z_j - u_j$. Let $\hat{u}_{m+1} = u_{m+1}$ and

$$\hat{A} = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & -1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}. \tag{B.2}$$

Then $g_t(y|d)$ can be written as

$$g_t(z|d) = \min_{\hat{u}:(z,\hat{u})\in\hat{\mathcal{U}}(d)} \{\hat{\mathcal{L}}(z, \hat{u}|d) + \alpha f_{t+1}(\hat{u}_1 - \hat{u}_{m+1}, ..., \hat{u}_m - \hat{u}_{m+1})\}, \tag{B.3}$$

where

$$\hat{\mathcal{L}}(z, \hat{u}|d) = \sum_{i=1}^{m} h_i(\hat{u}_i - \hat{u}_{m+1}) + \sum_{j=1}^{m} b_j(d_j - z_j + \hat{u}_j) + b_{m+1}(d_{m+1} - \hat{u}_{m+1}), \tag{B.4}$$

and

$$\hat{\mathcal{U}}(d) = \{(z, \hat{u})|\hat{A}\hat{u} \leq 0, 0 \leq z_j - \hat{u}_j \leq d_j, j = 1, 2, ..., m, 0 \leq \hat{u}_{m+1} \leq d_{m+1}\}. \tag{B.5}$$

We then prove by induction. Clearly $f_{T+1}(x)$ is $L^\natural$-convex. If f_{t+1} is $L^\natural$-convex then the objective function of (B.3) is also $L^\natural$-convex in $(z, \hat{u})$ due to Proposition 2.1 (a) and (d). The constraint set $\hat{\mathcal{U}}(d)$ forms a $L^\natural$-convex set by Proposition 2.1 (g). Therefore, $g_t(z|d)$ is $L^\natural$-convex in z for any d according to Proposition 2.1 (e). Similar to part (a) we have $\mathcal{A}^\Xi = \{(x, w)|x_i \leq w_i \leq x_i + \bar{\xi}_i, \forall i = 1, ..., m\}$, which is $L^\natural$-convex following Proposition 2.1 (g). One can easily check that $\mathcal{A}$ is of the form given in Lemma 2.3. Therefore, applying Theorem 2.3 and 2.4, we know that $f_t(x)$ is also $L^\natural$-convex and the sensitivity results hold. □

B.5 Proof of Theorem 3.4

(a) We prove by induction. Assume f_{t+1} is concave. In the objective function of (3.17), $g_t(\cdot, \cdot)$ is clearly concave. Since $\mathcal{A} = \{(x, u)|Au \leq x, u \geq 0\}$, $\mathcal{A}^\Xi = \{(x, u \wedge d)|(x, u) \in \mathcal{A}, d \in Supp(D)\}$, which is equivalent to the convex set $\{(x, w) : Aw \leq x, 0 \leq w_j \leq \bar{d}_j, \forall j = 1, ..., n\}$. In addition, Assumption 2.2 is satisfied since $\mathcal{A}$ is of the form given in Lemma 2.3. Then it follows from Theorem 2.3 that $f_t(x)$ is concave.

(b) For $j = 1, ..., m$, define $\hat{u}_j = x_j - u_j$, $\hat{u}_{m+1} = u_{m+1}$ and $\hat{A}$ is given in (B.2). The optimality equations can be rewritten as

$$f_t(x) = \max_{\hat{u}:(x,\hat{u})\in\mathcal{A}} E\left[\pi_{m+1} + \sum_{j=1}^{m} p_j x_j - \sum_{j=1}^{m} p_j(\hat{u}_j \vee (x_j - d_j)) + f_{t+1}(\tilde{x})\right], \tag{B.6}$$

where $\pi_{m+1} = p_{m+1}(\hat{u}_{m+1} \wedge d_{m+1})$, $\hat{\mathcal{A}} = \{(x, \hat{u}) | \hat{u} \geq 0, \hat{A}\hat{u} \leq 0, x_j - \hat{u}_j \geq 0, j = 1, ..., m\}$ and

$$\tilde{x} = [\hat{u}_1 \vee (x_1 - d_1), ..., \hat{u}_m \vee (x_m - d_m)] - (\hat{u}_{m+1} \wedge d_{m+1})e.$$

Define $h_t(x, \xi) = p_{m+1}\xi_{m+1} + \sum_{j=1}^{m} p_j x_j - \sum_{j}^{m} p_j \xi_j + f_{t+1}((\xi_1, ..., \xi_m) - \xi_{m+1}e)$. Then for $t = 1, ..., T$,

$$f_t(x) = \max_{\hat{u}:(x,\hat{u})\in\mathcal{A}} E[h_t(x, \hat{u}_1 \vee (x_1 - d_1), ..., \hat{u}_m \vee (x_m - d_m), \hat{u}_{m+1} \wedge d_{m+1})].$$

Clearly $f_{T+1}(x)$ is $L^\natural$-concave. If $f_{t+1}(x)$ is $L^\natural$-concave, then $h_t(x, \xi)$ is also $L^\natural$-concave by Proposition 2.1 (a) and (d). We have that $\mathcal{A}^\Xi$ can be expressed as

$$\{(x, \hat{u}_1 \vee (x_1 - d_1), ..., \hat{u}_m \vee (x_m - d_m), \hat{u}_{m+1} \wedge d_{m+1}) | (x, \hat{u}) \in \hat{\mathcal{A}}\}.$$

Notice that $\mathcal{A}^\Xi$ is equivalent to the following set

$$\{(x, w) | (x_j - \bar{d}_j) \vee w_{m+1} \leq w_j \leq x_j, w_j \geq 0, j = 1, ..., m, 0 \leq w_{m+1} \leq \bar{d}_{m+1}\},$$

where $\bar{d}_j = \operatorname{ess\,sup}\{d_j | d \in Supp(D)\}$. It follows from Proposition 2.1 (g) that $\mathcal{A}^\Xi$ is $L^\natural$-convex. One can easily check that $\mathcal{A}$ is of the form in Lemma 2.3. Therefore, Theorem 2.3 can be applied to show that the $L^\natural$-concavity of $f_t(x)$ is preserved. It follows from Theorem 2.4 that there exists a greatest solution $\hat{u}^*(x)$ such that $\hat{u}_j^*(x)$ is increasing in x for all j with limited sensitivity, which implies that $u_{m+1}^*(x) = \hat{u}_{m+1}^*(x)$ is increasing in x with limited sensitivity while $u_j^*(x) = x_j - \hat{u}_j^*(x)$ is increasing in x_j, and decreasing in $x_k, k \neq j$ with limited sensitivity. $\square$

Appendix C

C.1 Proof of Theorem 4.1

Let π^* be the optimal objective value of problem (4.2). Since for any $u \in \mathcal{F}^n$, $(u, (v(\xi), \xi \in \mathcal{X})) = (u, (u \wedge \xi, \xi \in \mathcal{X}))$ is feasible for problem (4.2), we have $\pi^* \leq \tau^*$.

It remains to show that $\tau^* \leq \pi^*$. Clearly, it holds when $\pi^* = \infty$. Thus, in the following, we assume that $\pi^* < \infty$, which together with assumption (a) implies that all optimization problems involved below, as well as problems (4.1) and (4.2), admit finite optimal solutions.

We first show that it is true for $n = 1$. Let $\hat{u} \in \arg\min_u g(u)$ and $u^* \in \arg\min_u \{l(u) + E[g(u \wedge \Xi)]\}$. Given any solution of problem (4.2), denoted by $(u', (v'(\xi), \xi \in \mathcal{X}))$, we are going to show that

$$l(u^*) + E[g(u^* \wedge \Xi)] \leq l(u') + E[g(v'(\Xi))].$$

For this purpose, we discuss two cases.

Case 1: if $u' \leq \hat{u}$, then

$$l(u^*) + E[g(u^* \wedge \Xi)] \leq l(u') + E[g(u' \wedge \Xi)] \leq l(u') + E[g(v'(\Xi))].$$

The first inequality is due to the optimality of u^*. In the following we prove the second inequality. If $\xi \leq u'$, then $g(u' \wedge \xi) = g(\xi) \leq g(v'(\xi))$ since $v'(\xi) \leq \xi \leq u' \leq \hat{u}$ and $g(\cdot)$ is convex. If $\xi > u'$, then $g(u' \wedge \xi) = g(u') \leq g(v'(\xi))$ since $v'(\xi) \leq u' \leq \hat{u}$ and $g(\cdot)$ is convex. Hence, $g(u' \wedge \xi) \leq g(v'(\xi))$ for any $\xi \in \mathcal{X}$.

Case 2: if $u' > \hat{u}$, then

$$l(u^*) + E[g(u^* \wedge \Xi)] \leq l(\hat{u}) + E[g(\hat{u} \wedge \Xi)] \leq l(u') + E[g(v'(\Xi))].$$

The first inequality is due to the optimality of u^*. Since $v'(\xi) \le \xi$, it is easy to see that $g(\hat{u} \wedge \xi) \le g(v'(\xi))$ for any $\xi \in \mathcal{X}$, and thus $E[g(\hat{u} \wedge \Xi)] \le E[g(v'(\Xi))]$. Then the second inequality follows from $l(\cdot)$ is increasing.

Combing the above two cases we have $\tau^* \le \pi^*$. This completes the proof when $n = 1$. Notice that when $n = 1$, the last constraint in (4.2) is redundant.

We now consider the general case with $n > 1$. We start from the first component. Let $(u', v') = (u'_1, ..., u'_n, ((v'_1(\xi_1), ..., v'_n(\xi_n)), \xi \in \mathcal{X})$ denote an optimal solution of problem (4.2). Given $(u'_2, ..., u'_n, ((v'_2(\xi_2), ..., v'_n(\xi_n)), \xi \in \mathcal{X}))$, π^* is equal to the optimal objective value of the following problem

$$\begin{array}{ll} \inf & l(u_1, u'_2, ..., u'_n) + E[f(v_1(\Xi_1), v'_2(\Xi_2), \ldots, v'_n(\Xi_n))] \\ \text{s.t.} & v_1(\xi_1) \in \mathcal{F}, \\ & v_1(\xi_1) \le \xi_1 \ \forall \xi_1 \in \mathcal{X}_1, \\ & v_1(\xi_1) \le u_1 \ \forall \xi_1 \in \mathcal{X}_1, \\ & v_1(\xi_1) \text{ is increasing } \forall \xi_1 \in \mathcal{X}_1. \end{array} \tag{C.1}$$

We need to show that there exists a u_1^* such that

$$l(u_1^*, u'_2, ..., u'_n) + E[f(u'_1 \wedge \Xi_1, v'_2(\Xi_2), \ldots, v'_n(\Xi_n))] = \pi^*.$$

For this purpose, we firstly prove the following lemma.

Lemma C.1. *Suppose that the function $g(v, \xi) : \mathcal{F} \times \mathcal{X} \to \bar{\Re}$ is supermodular in (v, ξ), and for any $\xi \in \mathcal{X}$, $g(v, \xi)$ is convex in v. Consider the following problem:*

$$\begin{array}{ll} \inf & l(u) + E[g(v(\Xi), \Xi)] \\ s.t. & v(\xi) \in \mathcal{F}, \\ & v(\xi) \le \xi \ \forall \xi \in \mathcal{X}, \\ & v(\xi) \le u \ \forall \xi \in \mathcal{X}, \\ & v(\xi) \textit{ is increasing } \forall \xi \in \mathcal{X}. \end{array} \tag{C.2}$$

Then there exists a u^* such that the optimal objective value of problem (C.2) *is equal to $l(u^*) + E[g(u^* \wedge \Xi, \Xi)]$.*

Proof of Lemma C.1: We start by consider the following problem:

$$\begin{array}{ll} \inf & E[g(v(\Xi), \Xi)] \\ \text{s.t.} & v(\xi) \in \mathcal{F} \ \forall \xi \in \mathcal{X} \end{array} \tag{C.3}$$

Let $\tilde{\xi} = -\xi$ and $\tilde{g}(v, \tilde{\xi}) = g(v, -\tilde{\xi})$, then $\tilde{g}(v, \tilde{\xi})$ is submodular in $(v, \tilde{\xi})$, and the optimal solution of $\min_v g(v, \xi)$ is equivalent to that of $\min_v \tilde{g}(v, \tilde{\xi})$. It then follows from Theorem 2.2.8 of Simchi-levi et al. (2014) that there exists an optimal solution of $\min_v \tilde{g}(v, \tilde{\xi})$ which is increasing in $\tilde{\xi}$. Equivalently, there exists an optimal solution of $\min_v g(v, \xi)$ which is decreasing in ξ, denoted by $\hat{v}(\xi)$. Notice that $(\hat{v}(\xi)|\xi \in \mathcal{X})$ is an optimal solution of problem (C.3).

Let $(u', (v'(\xi), \xi \in \mathcal{X}))$ be an optimal solution of problem (C.2). We are going to show that there exists a u^* such that $l(u^*) + E[g(u^* \wedge \Xi, \Xi)] \leq l(u') + E[g(v'(\Xi), \Xi)]$. We divide our discussion into the following three cases and pick u^* accordingly. (1) When exists some $\hat{\xi}$ such that $\hat{v}(\xi) \geq v'(\xi)$ for $\xi < \hat{\xi}$ and $\hat{v}(\xi) \leq v'(\xi)$ for $\xi \geq \hat{\xi}$, choose $u^* = \sup_{\xi < \hat{\xi}} v'(\xi)$; (2) When $v'(\xi) \leq \hat{v}(\xi)$ $\forall \xi \in \mathcal{X}$, choose $u^* = \sup v'(\xi)$; (3) When $v'(\xi) \geq \hat{v}(\xi)$ $\forall \xi \in \mathcal{X}$, let $u^* = \inf v'(\xi)$.

Case (1): If $\xi \geq \hat{\xi}$, then $u^* \wedge \xi = u^*$ since $u^* \leq v'(\hat{\xi}) \leq \hat{\xi}$. We also have $\hat{v}(\xi) \leq u^*$ since $\hat{v}(\xi)$ is decreasing, and $v'(\xi) \geq u^*$ since $v'(\xi)$ is increasing. Since the function $g(v, \xi)$ is convex in v for any ξ, and $\hat{v}(\xi) \leq u^* \wedge \xi \leq v'(\xi)$, we have $g(u^* \wedge \xi, \xi) \leq g(v'(\xi), \xi)$. If $\xi \leq \hat{\xi}$, we have $v'(\xi) \leq u^* \wedge \xi$ since $v'(\xi) \leq \xi$, $v'(\hat{\xi}) \leq u^*$, and $v'(\xi)$ is increasing. Because $v'(\xi) \leq u^* \wedge \xi \leq \hat{v}(\xi)$ and $g(v, \xi)$ is convex in v for any $\xi \in \mathcal{X}$, we obtain that $g(u^* \wedge \xi, \xi) \leq g(v'(\xi), \xi)$ $\forall \xi \in \mathcal{X}$. Since $u^* \leq u'$ and $l(\cdot)$ is increasing, we have $l(u^*) + E[g(u^* \wedge \Xi, \Xi)] \leq l(u') + E[g(v'(\Xi), \Xi)]$.

Case (2): When $v'(\xi) \leq \hat{v}(\xi)$ $\forall \xi \in \mathcal{X}$, we have $v'(\xi) \leq u^* \wedge \xi \leq \hat{v}(\xi)$ $\forall \xi \in \mathcal{X}$. Since $g(v, \xi)$ is convex in v for any ξ, we have $g(u^* \wedge \xi, \xi) \leq g(v'(\xi), \xi)$ $\forall \xi \in \mathcal{X}$. Notice that in this case we must have $u' = u^*$, hence $l(u^*) + E[g(u^* \wedge \Xi, \Xi)] \leq l(u') + E[g(v'(\Xi), \Xi)]$.

Case (3): When $v'(\xi) \geq \hat{v}(\xi)$ $\forall \xi \in \mathcal{X}$, we have $u^* \wedge \xi = u^*$. Hence, $\hat{v}(\xi) \leq u^* \wedge \xi \leq v'(\xi)$ $\forall \xi \in \mathcal{X}$ and $g(u^* \wedge \xi, \xi) \leq g(v'(\xi), \xi)$ $\forall \xi \in \mathcal{X}$. Since $u^* \leq u'$, we obtain $l(u^*) + E[g(u^* \wedge \Xi, \Xi)] \leq l(u') + E[g(v'(\Xi), \Xi)]$.

This completes the proof of Lemma C.1.

Now back to problem (C.1). By the law of iterative expectation, we can rewrite the second term in the objective function of problem (C.1) as

$$\begin{aligned}
&E[f(v(\Xi))] \\
&= E_{\Xi_1}[E_{\Xi_2,\ldots,\Xi_n|\Xi_1}[f(v_1(\Xi_1), v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))]] \\
&= \int_{\mathcal{X}_1} g(v_1(\xi_1), \xi_1) dF_{\Xi_1}(\xi_1),
\end{aligned}$$

where

$$g(v_1, \xi_1) = E_{\Xi_2,\ldots,\Xi_n|\Xi_1=\xi_1}[f(v_1, v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))].$$

Since f is componentwise convex, $g(v_1, \xi_1)$ is convex in v_1 for any ξ_1. In the following we show that $g(v_1, \xi_1)$ is supermodular in (v_1, ξ_1). Define $w = (w_2, ...w_n)$ and

$$\tilde{f}(v_1, w) = f(v_1, v_2^*(w_2), ..., v_n^*(w_n)).$$

Then

$$g(v_1, \xi_1) = \int \tilde{f}(v_1, w) d\tilde{F}_{\xi_1}(w),$$

where $\tilde{F}_{\xi_1}(w)$ is the joint distribution of $\Xi_2, ..., \Xi_n$ conditional on $\Xi_1 = \xi_1$. By assumption (c), we have that $\{\tilde{F}_{\xi_1}(w)|\xi_1 \in \mathcal{X}_1\}$ is a stochastically increasing collection of distribution functions on $\Re^{n-1}$. Because $v_j^*(\xi_j)$ are increasing for $j = 2, ..., n$ and f is supermodular, $\tilde{f}$ is supermodular in (v_1, w) (See section 9.A.4 of Shaked & Shanthikumar 2006). Then it follows from Theorem 3.10.1 of Topkis (1998) that $g(v_1, \xi_1)$ is supermodular. Therefore, Lemma C.1 implies that there exists a u_1^* such that

$$l(u_1^*, u_2', ..., u_n') + E[f(u_1' \wedge \Xi_1, v_2'(\Xi_2), \ldots, v_n'(\Xi_n))] = \pi^*.$$

Next move on to the second component. Similar to (C.1), π^* is equal to the optimal objective value of the following problem

$$\begin{array}{ll} \inf & l(u_1^*, u_2, u_3', ..., u_n') + E[f(u_1^* \wedge \Xi_1, v_2(\Xi_2), v_3'(\Xi_3), \ldots, v_n'(\Xi_n))] \\ \text{s.t.} & v_2(\xi_2) \in \mathcal{F}, \\ & v_2(\xi_2) \leq \xi_2 \ \forall \xi_2 \in \mathcal{X}_2, \\ & v_2(\xi_2) \leq u_2 \ \forall \xi_2 \in \mathcal{X}_2, \\ & v_2(\xi_2) \text{ is increasing } \forall \xi_2 \in \mathcal{X}_2. \end{array} \tag{C.4}$$

Following the proceeding analysis, there exists a u_2^* such that

$$\pi^* = l(u_1^*, u_2^*, u_3', ..., u_n') + E[f(u_1^* \wedge \Xi_1, u_2^* \wedge \Xi_2, v_3'(\Xi_3), ..., v_n'(\Xi_n))].$$

Continue this process and applying the same approach we can find $u_j^*, j = 1, ..., n$ such that $\pi^* = l(u_1^*, ..., u_n^*) + E[f(u_1^* \wedge \Xi_1, ..., u_n^* \wedge \Xi_n)]$. Therefore, we have $\pi^* = \tau^*$. $\square$

C.2 Proof of Theorem 4.2

Problem (4.3) is equivalent to the following unconstrained optimization problem.

$$\inf_{u \in \mathcal{F}^n} l(u) + E[f(u \wedge \Xi)] + \delta_{\mathcal{U}}(u). \tag{C.5}$$

For any $v \in \mathcal{F}^n$, let $\hat{f}(v) = f(v) + \delta_{\mathcal{V}}(v)$, where $\mathcal{V}$ is defined in (4.4). Then the optimal objective value of problem (C.5) is equivalent to that of the following problem

$$\inf_{u \in \mathcal{F}^n} l(u) + E[\hat{f}(u \wedge \Xi)]. \tag{C.6}$$

To see this, note that for any $u \in \mathcal{U}$, we have $u \wedge \xi \in \mathcal{V}\ \forall \xi \in \mathcal{X}$. Hence, given any feasible solution u of $\inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$, we can always find a $\hat{u} \in \mathcal{F}^n$ (simply let $\hat{u} = u$) such that $l(\hat{u}) + E[\hat{f}(\hat{u} \wedge \Xi)] = l(u) + E[f(u \wedge \Xi)]$. Therefore, $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\} \leq \inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$. On the other hand, due to Assumption 4.1, given any feasible solution $\hat{u}$ of $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\}$, we can always find a $u \in \mathcal{U}$ such that $l(\hat{u}) + E[\hat{f}(\hat{u} \wedge \xi)] \geq l(u) + E[f(u \wedge \Xi)]$. Therefore, $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\} \geq \inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$. Applying Theorem 4.1 to problem (C.6), we can obtain the transformed problem (4.5). $\square$

C.3 Proof of Theorem 4.3

By the definition of CVaR, problem (4.6) is equivalent to

$$\inf_{u\in\mathcal{U}}\inf_{\hat{\lambda}\in\Re}\{\hat{\lambda}+\frac{1}{1-\alpha}E[(l(u)+f(u\wedge\Xi)-\hat{\lambda})^{+}]\}. \tag{C.7}$$

Define $\lambda=\hat{\lambda}-l(u)$ and $g(u,\lambda)=\lambda+\frac{1}{1-\alpha}(f(u)-\lambda)^{+}$, the above problem can be reformulated as

$$\inf_{\lambda\in\Re,u\in\mathcal{U}} l(u)+E[g(u\wedge\Xi,\lambda)].$$

Notice that if f is componentwise/jointly convex, then $g(u,\lambda)$ is also componentwise/jointly convex. By Theorem 4.2, given any λ the problem $\inf_{u\in\mathcal{U}} l(u)+E[g(u\wedge\Xi,\lambda)]$ can be transformed to

$$\begin{aligned}
\inf \quad & l(u)+E[g(v(\Xi),\lambda)]\\
\text{s.t.} \quad & v(\xi)=(v_1(\xi_1),...,v_n(\xi_n))\in\mathcal{V}\ \forall\xi\in\mathcal{X},\\
& v(\xi)\le\xi\ \forall\xi\in\mathcal{X},\\
& v(\xi)\le u\ \forall\xi\in\mathcal{X}.
\end{aligned}$$

Therefore, problem (4.6) is equivalent to problem (4.7). □

C.4 Proof of Corollary 4.1

Based on the definition of the distortion risk measure, problem (C.7) can be formulated as

$$\inf_{u\in\mathcal{U}}\int_0^1\inf_{\lambda(\alpha)\in\Re,\ \forall\alpha\in[0,1)}[l(u)+E[g(u\wedge\Xi,\lambda(\alpha),\alpha)]]d\mu(\alpha),$$

which is equivalent to

$$\inf_{\lambda(\alpha)\in\Re,\forall\alpha\in[0,1)}\inf_{u\in\mathcal{U}}\left[l(u)+\int_0^1 E[g(u\wedge\Xi,\lambda(\alpha),\alpha)]d\mu(\alpha)\right].$$

We have that $\int_0^1 E[g(u,\lambda(\alpha),\alpha)]d\mu(\alpha)$ is componentwise convex in u. Then the desired results follow. □

Appendix D

D.1 Proof of Lemma 5.1

Consider the following maximum weight circulation problem:

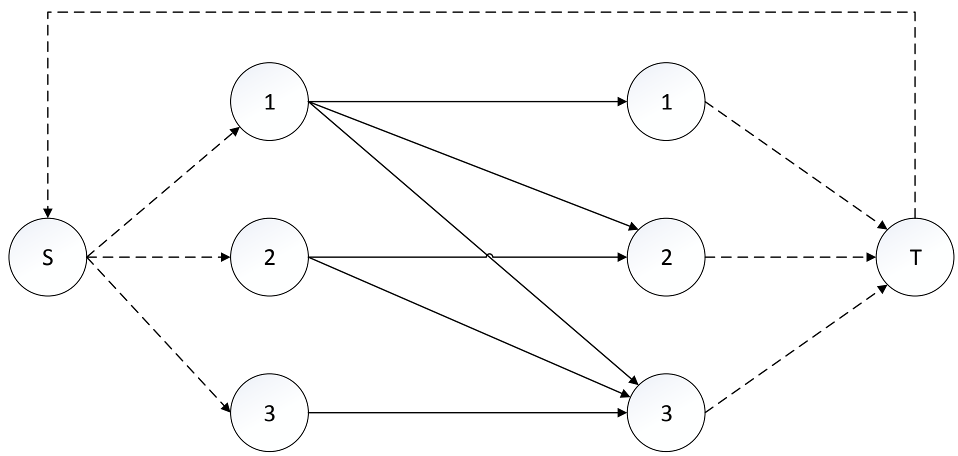
$$\begin{aligned}
F = \max_{w_{ij}, f_{Si}, f_{jT}} & -\sum_{i=1}^{N}\sum_{j=1}^{N} s_{ij}w_{ij} + \sum_{i}^{N} h_i f_{Si} + \sum_{j=1}^{N} p_j f_{jT} \\
s.t. \quad & \sum_{j=i}^{N} w_{ij} = f_{Si},\ \forall i = 1, ..., N \\
& \sum_{i=1}^{j} w_{ij} = f_{jT},\ \forall j = 1, ..., N \\
& 0 \leq f_{Si} \leq y_i, 0 \leq f_{jT} \leq d_j, 0 \leq w_{ij}\ \forall i, j = 1, ..., N.
\end{aligned} \tag{D.1}$$

In this maximum weight circulation problem, we add two more nodes S and T to the original product-demand graph (see Figure D.1 for an example with $n = 3$). Let w_{ij} be the flow from product i to demand j, f_{Si} be the flow from node S to product i, and f_{jT} be the flow from demand j to node T. Flows f_{Si} is bounded above by y_i and f_{jT} is bounded above by d_j. The weights for w_{ij}, f_{Si}, f_{jT} are $-s_{ij}, h_i, p_j$ respectively. since the arcs associated with flow f_{Si} are parallel, then by Theorem 1.3 and 1.4 of Murota (2005), F is concave and submodular. Notice that $L(y|d) = -F + \sum_{i=1}^{N} h_i y_i + \sum_{j=1}^{N} p_j d_j$, thus $L(y|d)$ is convex and supermodular in y. This completes the proof. $\square$

D.2 Proof of Theorem 5.1

Let $\mathcal{K}$ denote the support of the random capacities K, and define $\bar{k}_j = \operatorname{ess\,sup}\{k_j | k \in \mathcal{K}\}$. Note that $\mathcal{V} = \{y \wedge (x + k) | y \geq x, k \in \mathcal{K}\}$ is equivalent to

Figure D.1: The equivalent maximum weight circulation problem



the set $\{w | x_j \leq w_j \leq x_j + \bar{k}_j, j = 1, ..., n\}$, whose indicator function is convex and supermodular. For any $y \in \mathcal{F}^n$ such that $y \wedge (x + k) \in \mathcal{V}$ $\forall k \in \mathcal{K}$, we must have $y \geq x$. To see this, suppose that $y_j < x_j$ for some $j \in \{1, ..., n\}$. Since the random capacities have a positive support, pick any $k \in \mathcal{K}$, we have $y_j \wedge (x_j + k_j) = y_j < x_j$, which is a contradiction. Hence, Assumption 4.1 is satisfied. Therefore, applying Theorem 4.2, the original problem can be transformed to (5.4). $\square$

*