



**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**DYNAMIC REVENUE MANAGEMENT OF AIRLINE
AND PRODUCTION SYSTEMS**

**LIN BING
NANYANG BUSINESS SCHOOL
2008**

Dynamic Revenue Management of Airline and Production Systems

Lin Bing

Nanyang Business School

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2008

Abstract

This dissertation investigates revenue management of airline seat inventory control problems and stock rationing of production/inventory problem. In a booking horizon, in order to maximize the total revenue through selling a finite number of seats of a plane, management must repeatedly decide whether to accept a booking request or reserve the seat for later higher-fare bookings. Correspondingly, for a production/inventory system, in terms of minimizing the penalty cost, it is often important to satisfy a class of demand with higher priority than a lower one. Therefore, inventories could be rationed accordingly. In this dissertation, Markov Decision Processes (MDP) has been used to formulate the problems of flight seat inventory control and rationing of the make-to-stock production/inventory system. Optimal policies for these problems have been completely or partially characterized. For each of the problems, numerical studies verify our analytical results.

Firstly, we study a two-flight problem where there are two parallel flights with same origin and destination in a day (e.g., one departs at 9:00am and another at 11:00am) and booking requests in each fare class arrive randomly. For these two flights, the airline can offer three products for passengers. Two of them are specific products for passengers who are dedicated to one of the two flights only; whereas the third one is a flexible product for those who are willing to take either flight. When a booking request arrives, upon its realization of the product type, the airline has to decide whether accept it, and in case a flexible type is accepted, which flight to accommodate it. Analytical study has characterized the structure of the optimal booking policy as four state-dependent monotone switching curves. We also extend the model to a case that flexible demand segment is purely induced by price discounts and another case of implicit flexible demand segment that is caused by cross-flight buying downs through price incentives. Customer choice has also been introduced into the model in the third extension. For all these extensions, the optimal booking policies can still be characterized as switching curves as well with minor adjustments. Finally, numerical results illustrate the derivation and dynamics of the optimal booking policies.

Secondly, a round-trip/two-leg airline revenue management problem has been analyzed through the discrete-time MDP (Markov Decision Processes). By inductive approach, we completely characterize the structure of the optimal policy as three state-dependent monotone switching curves. Furthermore, multiple-seat booking for round-trip/two-leg problem has been investigated. The structure of the optimal booking policy has also been shown to be the monotone switching curves if booking requests for multiple seats can be partially satisfied. Regarding the problems with three legs or more, under a reasonable assumption, we derive the monotone structural properties of the optimal value function and partially characterize the optimal booking policy as switching surfaces.

Thirdly, we consider rationing of a two-item, two production-facility, and make-to-stock production/inventory system. Demands are classified as three types with each type of m classes. Two types are dedicated to only one of the two items while the third type is the flexible demand which can be satisfied by any of the two items. Upon an arrival of customer demand, management should decide whether satisfy the demand or hold the stock for later higher priority demands. Here, we only consider the no-backlog case in which demand is lost if it is not satisfied. This problem is formulated as a continuous-time Markov Decision Process. Based on some structure properties of the optimal value function, we completely characterize the optimal policy as switching curves for each class of demand and each production process.

Finally, we give concluding remarks on our major results and the future directions of our research.

Acknowledgements

It is a great pleasure to acknowledge the contributions of many people to my research work. First and foremost, I would like to express my sincerest gratitude and greatest appreciation to my supervisor, Professor Chen Shaoxiang, for his continuous direction, guidance, and encouragement, and for his insightful comments and valuable suggestions. I also would like to thank Professor Arun Kumar at our School of Engineering, Professors Li, Zhi-feng, Michael, S. Viswanathan, Wang Qinan at Nanyang Business School, and three anonymous referees for this dissertation. Their comments and suggestions greatly improve this thesis.

I am grateful to the staff in the PhD office of Nanyang Business School for their continuous managerial and financial support.

Special thanks are given to the Division of Information Technology and Operations Management, and Nanyang Business School of Nanyang Technological University, for their continuous support for my research.

I would like to thank the technical staff and colleagues in Nanyang Business School for their kindly assistance. I also wish to thank Feng Yan, Xu Jianjun and Zhuang Weifen, the Ph.D. students in our division, Feng Yi and Zhang Weifeng in the Center for Supply Chain Management of Nanyang Technological University for their valuable suggestions during my research.

I would like to give my special thanks to Professor Yang Xin, my former supervisor in the School of Economics and Management of Tsinghua University. With his encouragement, I am on my way of research.

I am grateful to my parents and many of my former teachers and classmates for their encouragement and love.

Especially, I will thank my wife Wu Wei and my little child Yu-chen for their patience, love and continuous support. I am indebted to them too much.

Finally, this dissertation is dedicated to the memory of my Parents-in-Law.

Table of Content

Abstract	1
Acknowledgements	3
Table of Content	4
List of Figures	6
List of Tables	7
1. Introduction	8
2. Methodology in Revenue Management	13
2.1 Optimization Methods	13
2.2 Discrete-time Markov Decision Process	13
2.3 Semi-Markov Decision Process	15
2.4 Optimal Stopping	15
2.5 Continuous-time Markov Decision Process	15
2.6 Controlled Poisson Process	18
2.7 Fluid Control and Diffusion Control	19
3. Literature Review	26
3.1 Introduction	26
3.2 Optimal Seat Allocation for a Two-Flight Problem	27
3.3 Optimal Seat Allocation for the Round-Trip/Two-Leg Problem	29
3.4 Stock Rationing of a Make-To-Stock Production/Inventory System	30
4. Optimal Seat Allocation for a Two-Flight Problem	32
4.1 Introduction	32
4.2 Model Formulation	35
4.2.1 Dynamic Model Formulation	35
4.2.2 Advantages of the Model with Identified Flexible Demand	37
4.2.3 Structures of the Optimal Booking Policy	40
4.2.4 Properties of the Optimal Value Function and Switching Curves	48
4.3 Flexible Segment induced by Discount	51

4.4	An Implicit Flexible Segment Induced by Cross-Flight Buying Down	53
4.5	Model with Passenger Choice	54
4.6	Numerical Example	55
4.7	Conclusions	61
4.8	Appendix:	63
5.	Optimal Seat Allocation for Round-Trip/Two-Leg Problems.....	71
5.1	Introduction	71
5.2	The Round-Trip Problem	72
5.2.1	Dynamic Model Formulation	72
5.2.2	Structures of the Optimal Booking Policy.....	73
5.2.3	Further Properties of the Optima Value Function	80
5.3	The Two-Leg Problem	84
5.4	Deterministic Linear Programming to Solve the Two-Leg Problem	84
5.5	Extension to the Multiple-Seat Booking Problem.....	85
5.6	Extension to the Multi-Leg Problem	88
5.7	Numerical Example	91
5.8	Conclusions	102
5.9	Appendix:	104
6.	Stock Rationing of a Make-To-Stock Production/Inventory System	115
6.1	Introduction	115
6.2	Model Formulation.....	116
6.2.1	Dynamic Programming Equation	116
6.2.2	Structures of the Optimal Policy	119
6.3	The Case of Average Cost Criterion	125
6.4	Numerical Example	126
6.5	Conclusions	132
6.6	Appendix:	134
7.	Concluding Remarks.....	139
	References.....	142

List of Figures

Figure 4.1: Booking Procedure for Passengers	34
Figure 4.2: Switching Curve for the Booking Limits of Fare Class i on Flight 1	45
Figure 4.3: Switching Curve for the Booking Limits of Fare Class i on Flight 2	46
Figure 4.4: Switching Curve for the Booking Limits of Class i Flexible Passengers	46
Figure 4.5: Properties of the Switching Curve	47
Figure 4.6: Switching Curves for the Price Discount Case	52
Figure 4.7: Switching Curves for the Cross-Flight Buying Down Case	54
Figure 4.8: Optimal Decisions for Flight 1 Booking Requests at $n = 15$	56
Figure 4.9: Optimal Decisions for Flight 2 Booking Requests at $n = 15$	57
Figure 4.10: Optimal Decisions for Flexible Booking Requests at $n = 15$	58
Figure 4.11: Optimal Decisions for Flexible Booking Requests at $n = 10$	58
Figure 4.12: Switching Curves with Discount Fares at $n = 10$	59
Figure 4.13: Marginal Expected Value of $V(x_1, 6) - V(x_1+1, 6)$ at period 15	60
Figure 4.14: Marginal Expected value of $V(3, x_2) - V(3, x_2+1)$ at Period 15	60
Figure 4.15: Marginal Expected Seat Revenue of $V(7, x_2) - V(6, x_2)$	61
Figure 4.16: Marginal Expected Seat Revenue of $V(x_1, 9) - V(x_1, 8)$	61
Figure 5.1: The Scenario of the Round-Trip Problem	72
Figure 5.2: The ‘Slope’ of the Switching Curve $S_{1n}(x_2)$	78
Figure 5.3: The Switching Curves for Three Types of Booking Requests	79
Figure 5.4: The Scenario of the Two-Leg Problem	84
Figure 5.5: The Switching Curves for Type 3 Booking Requests	92
Figure 5.6: The Switching Curve for Type 1 Booking Requests	93
Figure 5.7: The Switching Curve for Type 2 Booking Requests	94
Figure 5.8: The Monotonicity of Switching Curve S_3	94
Figure 5.9: Switching Curve for the Lower Fare Class of Type 1 at the 7 th period	95
Figure 5.10: Switching Curve for the Lower Fare Class of Type 1 at the 8 th period	95
Figure 5.11: Switching Curve for the Lower Fare Class of Type 1 at the 9 th period	95
Figure 5.12: Monotone Properties of $V(x_1, x_2) - V(x_1+1, x_2)$	96
Figure 5.13: Monotone Properties of $V(x_1, x_2) - V(x_1, x_2 + 1)$	96
Figure 5.14: Monotone Properties of $V(x_1, x_2) - V(x_1+1, x_2 + 1)$	97
Figure 5.15: Monotonicity of $V(x_1, x_2) - V(x_1+1, x_2 + 1)$ with respect to n and $(x_1 + x_2)$	97
Figure 5.16: Decisions for Multiple-Seat Booking of Type 1 Booking Request	98
Figure 5.17: Decisions for Multiple-Seat Booking of Type 2 Booking Request	99
Figure 5.18: Decisions for Multiple-Seat Booking of Type 3 Booking Request	99
Figure 5.19: The Projected Switching Curve for Type 1	100
Figure 5.20: The Projected Switching Curve for Type 2	101
Figure 5.21: The Projected Switching Curve for Type 3	101
Figure 5.22: The Projected Switching Curve for Type 4	102
Figure 6.1: Three Types of Demands for Two Items	116
Figure 6.2: Switching curves for admission and routing controls	122
Figure 6.3: The decision curves for production of item 1	123
Figure 6.4: The decision curves for production of item 2	124
Figure 6.5: Switching curves for Lower-Priority Demand on Item 1	127
Figure 6.6: Switching curves for Lower-Priority Demand on Item 2	127
Figure 6.7: Switching curves for flexible Lower-Priority Demand	128
Figure 6.8: Switching Curve for Production of Item 1 with Lower Cost	129

Figure 6.9: Switching Curve for Production of Item 1 with Higher Cost.	129
Figure 6.10: Switching Curve for Production of Item 2 with Lower Cost.	130
Figure 6.11: Switching Curve for Production of Item 2 with Higher Cost.	130
Figure 6.12: Monotone Property of $V(x_1+1, x_2) - V(x_1, x_2)$	131
Figure 6.13: Monotone Property of $V(x_1, x_2+1) - V(x_1, x_2)$	131

List of Tables

Table 4.1: Probabilities for the Booking Requests	56
Table 5.1: Probabilities for the Booking Requests	92
Table 6.1: Arrival Rates for Three Types of Demands	126

1. Introduction

Revenue Management (RM) is an approach to maximizing revenue that is based on managing revenues around capacity and timing for different markets. According to Talluri and van Ryzin (2004), it is concerned with demand-management decisions with the objective of increasing revenues. The decisions in revenue management can be classified into three categories: structural decisions which address the issues of selling format, segmentation or differentiation mechanisms; price decisions which address issues on setting different prices and pricing over time; quantity decisions which are concerned with admission/rejection control, allocation of capacity among different segments, etc.

Revenue management first receives attention in the airline industry. It has a long history of more than forty years. Since the first publication on overbooking control of airline seat capacity, there is a substantial literature on study of revenue management. It is increasingly gaining attention from both academia and industries nowadays. It is a widely accepted method that provides dramatic revenue and profitability in various industries such as airlines, hospitality, car rental, cruise line, railroad, retailing, air cargo, flexible manufacturing, natural gas and electricity transmission and television broadcast.

In practice, firms usually have limited resources such as the number of seats of a flight, the number of rooms in a hotel, the number of machine hours in a factory, and the number of cars to be rented, etc. Management tries to maximize the expected revenue through acceptance/rejection control or pricing strategies for a product or service depending on its remaining capacity and time-to-go period.

In this dissertation, we will discuss two types of airline seat inventory control problems. One is on the seat capacity control of two parallel flights and the other is the round-trip/multi-leg problem. In addition, we are concerned with stock rationing of a production/inventory system which is closely related to revenue management in analysis.

Modern airlines must repeatedly decide whether or not to accept a seat booking request or refuse it and reserve the seat in hope of the later, higher-fare bookings. For the first topic of this

thesis, we study a two-flight model where there are two parallel flights between two cities in a day (e.g., one departs at 9:00am and another at 11:00am) and booking requests arrive randomly. With two flights, the airline can offer three products for passengers. Two are specific products for booking requests who are dedicated to one of the two flights only; whereas the third is a flexible product for those willing to take either flight. Upon a booking request realizing its product type, the airline has to decide whether to accept it, and in case a flexible type is accepted, which flight to accommodate it. We characterize the structure of optimal booking policies as four monotone switching curves. Extensions of the basic model have been made to the case with a flexible segment that is purely induced by price discounts and another case with implicit flexible segment that is caused by cross-flight buying downs through price incentives. In the third extension, customer choice is incorporated into our original model. We demonstrate that the optimal booking policies under these three cases can still be characterized by switching curves as well with minor adjustments. Finally, numerical study illustrates the derivation and the dynamics of the optimal booking policies.

For the second topic, we consider the multi-fare, round-trip/two-leg airline revenue management problem. Single-leg and one-way trip problems are intensively studied. A set of critical numbers called booking limits are shown to characterize the optimal booking policies. Here, we formulate the round-trip/two-leg model as a discrete-time two-dimension Markov decision process. By the inductive approach of dynamic programming, we derive structure properties of the optimal value function and completely characterize the optimal policy as three monotone switching curves. These switching curves are state and time dependent. Moreover, multiple-seat booking problem has been further investigated and the optimal booking policy is also characterized as switching curves if booking requests for multiple seats can be partially satisfied. For three or more legs problem, under a reasonable assumption, we derive the general structural properties for the optimal value function. Further, optimal booking policy has been partially characterized as switching surface.

So far, industries other than airlines have seen a growing interest in the adoption of Revenue Management strategies. More recently, manufacturing firms have started to evaluate such tactical optimization tools. Here, we consider the stock rationing problem of a make-to-stock inventory/production system which is closely related to the practice of Revenue Management through demand segmentation. The difference is that RM considers maximizing the total revenue while rationing is concerned with minimizing total cost. Here, we propose the stock rationing problem for a two-item and two production facility system. In our model formulation, demands are classified as three types with each type of m classes. Two types are inflexible and favor only one of the two items respectively. The third type is the flexible demand which can accept either of the two items. Upon an arrival of the customer demand, management should decide whether to satisfy the demand or hold the stock for later higher priority demands. In our case, we assume that demand is lost if it is not satisfied, i.e. no backlogging. Demands are assumed to arrive in accordance with the Poisson process. Production time is exponentially distributed. We formulate this problem as a continuous-time Markov Decision Process (MDP). Based on the structure properties of the optimal value function, we characterize the optimal policy as switching curves for each demand class and each production process.

This dissertation examines three problems on airline seat inventory control and rationing of the make-to-stock production/inventory system. The contribution is three-fold as follows.

First, previous analytical results on single-flight seat inventory control problem have been extended to the two-flight problem. Optimal booking policies have been completely characterized as switching curves/surfaces. For round-trip/two-leg problem, we completely characterize the structure of the optimal booking policy. For multi-leg (more than two legs) problem, we partially characterize the booking policy under a reasonable assumption. Moreover, as far as we know, we are the first one to completely characterize the structure of the optimal booking policy for the multiple-seat booking for the round-trip/two-leg problem.

For rationing of the make-to-stock production/inventory system, we extend research results from the single-item single-production facility system to the two-item two production

facility system. Optimal policies for both rationing decision and production decision have been completely characterized.

Second, it contributes to Revenue Management with the introduction of a flexible demand segment. The differentiation of customer demand can bring in significant revenue and improve customer service level as well. Such a topic is not fully addressed by previous academic research and industrial applications.

Third, it contributes to mathematical models of the airline seat inventory control systems and rationing of the make-to-stock production/inventory system.

In this dissertation, all our problems are formulated as Markov Decision Process (MDP) which is a very popular in the Operations Research and Management Science Society. MDP is a sequential decision process of a system that evolves over time, with opportunities to influence its future evolution by taking actions at various (discrete or continuous) points in time. The objective of the decision maker may be to maximize the total expected reward or minimize the total expected cost over a certain time horizon. The major result in this dissertation is the proof of the optimality of switching curve/surface policy for all three problems. We adopt the methodology of Monotone Comparative Statics developed by Topkis (1978, 1998). The approach to characterize the switching curve policies in all three problems is similar to that of Hajek (1984). Similar methodology and analysis could also be found in Ha (1997c) and Chen (2004). However, the settings are essentially different between ours and Ha (1997c) and Chen (2004) in that we consider two flights, two legs and two production facilities while they consider a single production facility for two items. Hence, the interpretation to the optimal policy could be significantly different.

The organization of this dissertation is as follows:

In chapter 2, we introduce the mathematical tools widely used in Revenue Management problems.

In chapter 3, we give an overview of the literature relevant to the airline seat inventory control problems and rationing of the production/inventory systems.

In chapter 4, we study the seat inventory control of two-parallel-flight problem. Results on the structure of optimal booking policy are obtained by the inductive approach of dynamic programming.

In chapter 5, we investigate the round-trip/multi-leg problem, and characterize the structures of optimal booking policy as switching curves based on the structure properties of the optimal value function.

In chapter 6, we use continuous-time Markov Decision Process to formulate a make-to-stock production/inventory system with stock rationing. Admission/rejection and production/idle decisions are made in the model. Structures of the optimal policy are shown to be switching curves.

In chapter 7, we summarize our major results and give some comments on future research and applications.

2. Methodology in Revenue Management

In this chapter, we will review mathematical models and methods widely used in Revenue Management. These methods and models are mainly taken from the research papers published on major journals.

2.1 Optimization Methods

Mathematical programming is a commonly used optimization method in the airline revenue management. It is widely used in leg-based control, bid-price control, origin-destination control and airline network control. For the application of this method in the airline seat inventory control problems, readers refer to Wollmer (1986), Curry (1990) and van Ryzin and Liu (2004) for details.

2.2 Discrete-time Markov Decision Process

In the dynamic control problem, discrete-time Markov Decision Processes (MDP) is a popular analytical tool used in airline seat inventory control. In a typical modeling of the airline seat inventory problem, the booking period is divided into such small subintervals. It is assumed that, within each subinterval, two or more passenger arrivals can be ignored. The optimality of booking limit policy can be derived through the concavity of optimal value function which is proved by the inductive approach of dynamic programming. Examples include Lee and Hersh (1993), Lautenbacher and Stidham (1999), Subramanian, Stidham and Lautenbacher (1999), You (1999), Talluri and van Ryzin (2004). For instance, we give explanation of this type of model formulation from Lautenbacher and Stidham (1999).

The booking horizon, $[0, T]$, is divided into N subintervals. These subintervals are numbered in the reverse order, with period N corresponding to the commencement of the booking horizon, and period 0 indicates the departure time of the air plane or the end of the booking horizon. In each subinterval, the probability of two or more requests can be ignored. The reason is as follow. We can make an assumption that the arrival processes of the booking requests are Poisson processes. For a Poisson process, the probability of one event in a small interval and the probability of more than two events are $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$, respectively. If different

fare-classes of passengers come with different arrival rates, we can unify these Poisson processes into a new one. We still have small interval satisfying the condition that the probability for two or more requests can be ignored. For $n = 1, 2, \dots, N$, let

p_{in} = Probability {class i arrival during the period n }, $i = 1, \dots, m$

p_{0n} = Probability {no arrivals during the period n } and

$$\sum_{i=0}^m p_{in} = 1$$

$U_n(x)$ = The maximal undiscounted expected revenue over periods $n, n-1, \dots, 1, 0$, where x denotes the number of seats that have already been booked, $x = 0, 1, \dots, C$. C is the total seat capacity of the flight. Without loss of generality, it is assumed that $r_1 > r_2 > \dots > r_m$, where revenue r_i is the fare for class i .

A discrete-time Markov decision process (MDP) is used to formulate the above problem. The dynamic programming equation is

$$U_n(x) = \sum_{i=1}^m p_{in} \max \{r_i + U_{n-1}(x+1), U_{n-1}(x)\} + p_{0n} U_{n-1}(x)$$

$$U_0(x) = \begin{cases} 0 & x \leq C \\ -R(x-C) & x > C \end{cases} \quad \text{where } R \geq \max_i \{r_i\}$$

The boundary condition $U_0(x)$ means that at the departure time, if overbooking happens, a penalty cost larger than the revenue would be incurred. It is a technical treatment equivalent to that boundary condition is zero.

The proof of the existence of booking limits is through the concavity of optimal value function derived by the inductive approach of dynamic programming. Based on the concavity of $U_0(x)$, inductively, $U_n(x)$ is concave for any period n . Concavity of the optimal value function guarantees the existence of booking limits which is a threshold from acceptance/rejection control. When current state x is smaller than the booking limit, it is optimal to accept a booking request, otherwise reject the booking request.

2.3 Semi-Markov Decision Process

Generally, if we regard the arrival of passengers as a semi-Markov arrival process, then, Semi-Markov decision process (SMDP) is used to study such kind of problem. Model formulation and analytical results can be found in Brummelle and Walczak (2003). Structure properties of the optimal policy remain the same as those of the problems with discrete-time MDP formulation.

2.4 Optimal Stopping

Brummelle and McGill (1993) investigate the case of single-leg, multi-fare seat inventory control. They formulate it as an optimal stopping problem with the assumption that low-fare class of booking requests arrive earlier than the higher-fare class of booking requests. They derive a closed-form solution of the optimal booking policy based on optimal stopping rule.

Feng and Gallego (1995) consider the problem of selling a fixed number of items in a finite period. In their case, it is needed to decide the optimal timing of a single price change from a given initial price to either a lower or higher second price. Based on the continuous-time dynamic programming equation, they show the optimality of the state-dependent time-threshold policy. Moreover, Feng and Xiao (2000) extend similar results to the multiple price case.

Zhao and Zheng (2001) study a two-class dynamic seat allocation model under the assumption that once the low fare class is closed first, it can not be reopened. Thus, it is an optimal stopping problem.

2.5 Continuous-time Markov Decision Process

Markov decision process (MDP) is commonly used in stock rationing of the production/inventory systems. Stock rationing is a kind of strategy equivalent to the demand segmentation of revenue management. In the context of stock rationing, researchers attempt to minimize the production cost, inventory holding cost and penalty cost incurred by rejection of customer demand. Customer demands are often assume to arrive as Poisson processes. The production time follows the exponential distribution. Thereafter, continuous-time Markov Decision Process could be used to formulate the problem. Since the system is memoryless, a Markovian policy will be optimal. Based on the uniformization technique developed by Lippman

(1975), we can derive the dynamic programming optimality equation. Applying the inductive approach, we can prove that the optimal value function has certain properties. This is a convenient technique used in optimal control of queueing systems and production systems. Based on the optimality equation, structures of the optimal policy can be characterized if the function of holding cost per unit time has special structures. It is often assumed to be convex. As a result, we can characterize the optimal policy for this Markovian production/inventory system as threshold control or critical level control, e.g., produce when inventory is below a threshold level or stop otherwise; accept the demand if inventory is above certain critical level, otherwise reject it.

Let us consider the model formulation in Ha (1997a). He considers a single-item multi-class stock rationing problem with lost sales. It is assumed that there are m fare classes. The lost sales costs for the m fare classes are $c_1 > c_2 > \dots > c_m$. $h(x)$ is the inventory holding cost per unit time. The objective is to find the optimal control policy that minimizes the total discounted expected cost over infinite horizon.

$$J_u(x) = E_x^u \left[\int_0^\infty e^{-\alpha t} h(x) dt + \sum_{i=1}^m \int_0^\infty e^{-\alpha t} c_i dM_i^t \right]$$

Where M_i^t indicate, up to time t , the number of rejected i -th class customers.

And value function $V(x)$ is the infimum of $J_u(x)$ over all admissible policy $u \in U$

$$V(x) = \inf_{u \in U} J_u(x)$$

where U is the set of nonanticipating admissible control policies.

Let μ be the production rate. The arrival rates for each class of customer demands are λ_i respectively with $i = 1, \dots, m$. Following the uniformization technique in Lippman (1975), let

$$\Lambda = \mu + \sum_{i=1}^m \lambda_i.$$

Now the transition time τ between two transitions is exponentially distributed with the same rate Λ .

Let t_k be the k -th transition time. The objective function can be rewritten as follows:

$$\begin{aligned}
 J_u(x) &= \sum_{k=0}^{\infty} E \left[\int_{t_k}^{t_{k+1}} e^{-\alpha t} h(x) dt + \sum_{i=1}^m \int_{t_k}^{t_{k+1}} e^{-\alpha t} c_i dM_i^t \right] \\
 &= \sum_{k=0}^{\infty} \left[E \left(\int_{t_k}^{t_{k+1}} e^{-\alpha t} dt \right) E(h(x_k)) + E \left(\int_{t_k}^{t_{k+1}} e^{-\alpha t} dt \right) \sum_{i=1}^m c_i \right] \\
 &= \frac{1}{\alpha + \Lambda} \left\{ \sum_{k=0}^{\infty} \left[\beta^k E(h(x_k)) + \beta^k \sum_{i=1}^m c_i \right] \right\}
 \end{aligned}$$

where $\beta = \frac{\Lambda}{\alpha + \Lambda}$. Thus, a continuous-time Markov Chain problem is equivalent to a discrete-

time Markov Chain problem with discount factor β .

The derivation for the last equality is based on following computation

$$\text{Observing } E(e^{-\alpha\tau}) = \int_0^{\infty} e^{-\alpha\tau} \Lambda e^{-\Lambda\tau} d\tau = \frac{\Lambda}{\alpha + \Lambda}, \text{ then}$$

$$E \left(\int_{t_k}^{t_{k+1}} e^{-\alpha t} dt \right) = \frac{E(e^{-\alpha t_k}) - E(e^{-\alpha(t_k + \tau)})}{\alpha} = \frac{E(e^{-\alpha k\tau})(1 - E(e^{-\alpha\tau}))}{\alpha} = \frac{1}{\alpha} \left(\frac{\Lambda}{\alpha + \Lambda} \right)^k \left(1 - \frac{\Lambda}{\alpha + \Lambda} \right)$$

Regarding the holding cost per stage as $\frac{h(x)}{\alpha + \Lambda}$, the dynamic programming equation takes the form

$$V(x) = \left\{ \frac{h(x)}{\alpha + \Lambda} + \beta \sum_{i=1}^m \frac{\lambda_{1i}}{\Lambda} T_i V(x) + \beta \frac{\mu}{\Lambda} \min \{V(x+1), V(x)\} \right\}$$

Where $\frac{\lambda_{1i}}{\Lambda}$ and $\frac{\mu}{\Lambda}$ are transition probabilities for arrival and production events respectively.

Without loss of generality, by scaling the time one can assume that $\alpha + \Lambda = 1$. Then the dynamic programming equation has been reduced to

$$V(x) = h(x) + \sum_{i=1}^m \lambda_{1i} T_i V(x) + \mu \min \{V(x+1), V(x)\}$$

Where T_i is the operators defined as

$$T_i V(x) = \begin{cases} V(x) & \text{if } x \leq 0 \\ \min \{V(x-1), V(x) + c_i\} & \text{if } x > 0 \end{cases}$$

For this model formulation, it can be proved that there exist critical inventory levels for admission/rejection of various customer demands and decisions of production/idle of the machine. To prove optimality of such control policy, it is critical to show that optimal value function is concave. This can be done by the inductive approach based on the truth of $h(x)$.

2.6 Controlled Poisson Process

Continuous-time Poisson controlled process is a popular analytical tool. Liang (1999), Van Slyke and Yung (2000), Zhao and Zheng (2001) and Feng and Xiao (2001) consider the continuous-time stochastic control model for airline seat inventory control. Dynamic programming equations derived in differential form are the so-called Hamilton-Jacobi-Bellman (HJB) equation. Optimal booking policy is of the threshold type which could be capacity-thresholds or time-thresholds. The capacity-threshold points or time-threshold points are inductively computed in close-form solutions to the differential equations.

Poisson control process is also the commonly used model in dynamic pricing. Gallego and van Ryzin (1994) is a typical one address the issue of dynamic pricing. Based on Bellman's Principle of Optimality, a dynamic programming equation in differential form has been derived. It is often named as Hamilton-Jacobi equation or Hamilton-Jacobi-Bellman (HJB) equation. Under some conditions, close-form solution could be derived. More discussion on the analysis of such stochastic point control problem can be found in Bremaud (1980). Extensions and variations are further analyzed in Gallego and van Ryzin (1997) and Zhao and Zheng (2000), etc. For example, the model formulation in Gallego and van Ryzin (1994) is illustrated as follows:

A firm is selling n items over a finite period horizon. Demand follows Poisson process with rate $\lambda_s = \lambda(p_s)$ where rate $\lambda(p_s)$ at time s is a function of price p_s . The firm uses pricing policy to adjust the demand arrival rate. Both price and arrival rates are belonging to a closed set. The objective is to find a pricing policy u to maximize the expected revenue over a finite horizon.

$$J_u(t, n) = E_u \left[\int_0^t p_s dN_s \right]$$

Where N_s indicate, up to time s , the number of items sold. And $V(x)$ is the supreme of $J_u(x)$ over all admissible policy $u \in U$

$$V(t, n) = \sup_{u \in U} J_u(t, n)$$

where U is the set of nonanticipating admissible pricing policies.

In a small time interval Δt , demand arrives with probability $\lambda\Delta t$ and no demand arrives with probability $1 - \lambda\Delta t$. Applying Bellman's Principle of Optimality,

$$V(t, n) = \sup_{\lambda} [\lambda\Delta t [p(\lambda) + V(t - \Delta t, n - 1)] + (1 - \lambda\Delta t)V(t - \Delta t, n) + o(\Delta t)]$$

Rearranging the terms and letting $\Delta t \rightarrow 0$, we derive

$$\frac{\partial V(t, n)}{\partial t} = \sup_{\lambda} \lambda [p(\lambda) + V(t, n - 1) - V(t, n)]$$

Now the dynamic programming equation is established for this dynamic pricing problem.

2.7 Fluid Control and Diffusion Control

For problems in the production/inventory queueing system, Markov decision process (MDP) is a natural model for detailed description of the system dynamic. Alternatively, Caldentey and Wein (2002, 2005) use diffusion control model to study revenue management of the make-to-stock production systems. Maglaras (2003) applies both fluid control and diffusion control models to revenue management of the single-server make-to-order system. The major difference between fluid control and diffusion control lies in the demand arrival process. In Maglaras (2003), stochastic demand arrival rate is replaced by its deterministic rate while in Caldentey and Wein (2002, 2005) the demand process follows a diffusion process.

Let us compare the difference between fluid and diffusion models with previously studied models with respect to time and state space. For the discrete-time MDP formulation, both time and state space are discrete. For the controlled Poisson process, time is continuous, whereas, state space is discrete. To make problems more analytically tractable, fluid control and diffusion control consider both continuous time and continuous state space. Typically, heavy traffic approximation has been employed to rescale the time and state space. Then, applying the Principle of Optimality,

the dynamic programming equation called Hamilton-Jacobi-Bellman (HJB) equation is derived for researchers to investigate the properties of optimal value function. Sometimes, close-form solutions could be obtained. Unfortunately, we can not derive the close-form solution for HJB equation generally. Alternatively, numerical method like Markov chain approximation is used to approximate the original problem.

Caldentey and Wein (2002, 2005) consider the combined admission/rejection and production/idle control of the single-product, single-server make-to-stock queue. In their model, both demand and price are random. The price is assumed to follow geometric Brownian motion (GBM). Under the heavy traffic condition, the original problem is transformed into a diffusion control problem with two state variables, the inventory level and the price. For the HJB equation, they resort to Markov chain approximation technique to solve it. The solution is characterized by two switching curves used for admission/rejection and production/idle controls.

Here, we consider the model formulation in Caldentey and Wein (2005). Firstly, we need to introduce one-dimension and two-dimension diffusion control models. Consider the following optimization model:

$$V(t, x) = \max_u E \left\{ \int_t^T R(x, u, s) ds + B(T, x(T)) \right\}$$

$$\text{Subject to } dx = \mu(x, u, s) ds + \sigma(x, u, s) dZ$$

In the above equation, $R(x, u, s)$ is the revenue function and u is the decision variable or control variable. State variable x follows a diffusion process and $V(t, x)$ represents the optimal reward obtainable in $[t, T]$ given current state is (t, x) . $B(T, x(T))$ is the terminal reward. Applying Bellman's Principle of Optimality, we can rewrite the objective function as

$$V(t, x) = \max_u E \left\{ \int_t^{t+\Delta t} R(x, u, s) ds + V(t + \Delta t, x(t + \Delta t)) \right\}$$

Observing $x(t + \Delta t) = x + \Delta x$, we apply Taylor series expansion to $V(t + \Delta t, x + \Delta x)$ yielding

$$\begin{aligned} V(t + \Delta t, x + \Delta x) &= V(t, x) + V_t(t, x)\Delta t + V_x(t, x)\Delta x \\ &+ \frac{1}{2}V_{tt}(t, x)\Delta t^2 + \frac{1}{2}V_{xx}(t, x)\Delta x^2 + V_{xt}(t, x)\Delta t\Delta x + o(\Delta t) \end{aligned}$$

Then substitute the above formula and $\int_t^{t+\Delta t} R(x, u, s) ds = R(x, u, s)\Delta t + o(\Delta t)$ back into the objective function, obtaining

$$V(t, x) = \max_u E \left\{ \begin{aligned} &R(x, u, s)\Delta t + V(t, x) + V_t(t, x)\Delta t + V_x(t, x)\Delta x \\ &+ \frac{1}{2}V_{tt}(t, x)\Delta t^2 + \frac{1}{2}V_{xx}(t, x)\Delta x^2 + V_{xt}(t, x)\Delta t\Delta x + o(\Delta t) \end{aligned} \right\}$$

Based on the standard results that

$$E(\Delta x) = \mu(x, u, t)\Delta t$$

$$E(\Delta x^2) = \sigma^2(x, u, t)\Delta t$$

$$E(\Delta x\Delta t) = \mu(x, u, t)\Delta t^2 = o(\Delta t)$$

We derive the Hamilton-Jacobi-Bellman (HJB) equation which is a second-order partial differential equation (PDE).

$$0 = \max_u \left\{ R(x, u, s) + V_t(t, x) + \mu(x, u, t)V_x(t, x) + \frac{1}{2}\sigma^2(x, u, t)V_{xx}(t, x) \right\}$$

When there are two state variables, Hamilton-Jacobi-Bellman equation becomes a two-dimension PDE. It is derived in the same way by applying Bellman's Principle of Optimality. We directly write down the two-dimension equation.

$$0 = \max_u \left\{ \begin{aligned} &R(x, u, s) + V_t(t, x) + \mu_1(x, u, t)V_{x_1}(t, x) + \mu_2(x, u, t)V_{x_2}(t, x) \\ &+ \frac{1}{2}\sigma_1^2(x, u, t)V_{x_1x_1}(t, x) + \frac{1}{2}\sigma_2^2(x, u, t)V_{x_2x_2}(t, x) + \rho\sigma_1(x, u, t)\sigma_2(x, u, t)V_{x_1x_2}(t, x) \end{aligned} \right\}$$

where x is a state vector (x_1, x_2) .

From the above induction for Hamilton-Jacobi-Bellman (HJB) equation, we assume that optimal value function $V(t, x)$ exists as the twice-continuously-differentiable function.

To simplify the derivation of HJB equation, we can define the infinitesimal operators for the diffusion processes. For one-state and two-state variables, define operators, respectively, as

$$\Gamma_1 \equiv \mu \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}$$

$$\Gamma_2 \equiv \mu_1 \frac{\partial}{\partial x_1} + \mu_2 \frac{\partial}{\partial x_2} + \frac{1}{2} \sigma_1^2 \frac{\partial^2}{\partial x_1^2} + \frac{1}{2} \sigma_2^2 \frac{\partial^2}{\partial x_2^2} + \rho \sigma_1 \sigma_2 \frac{\partial^2}{\partial x_1 \partial x_2}$$

With the infinitesimal operator, the HJB equation with one state variable is

$$0 = \max_u \{ R(x, u, s) + V_t(t, x) + \Gamma_1 V(t, x) \}$$

If objective function has a discount factor as

$$J(t, x) = \max_u E \left\{ \int_t^T e^{-\alpha s} R(x, u) ds \right\}$$

We can firstly derive the corresponding HJB equation without discount factor. Thereafter, at time 0, present value of the objective function is given by $V(0, x) = e^{-\alpha t} J(x)$. Observe that $V_t(0, x) = -\alpha e^{-\alpha t} J(x)$, $V_x(0, x) = e^{-\alpha t} J_x(x)$ and $V_{xx}(0, x) = e^{-\alpha t} J_{xx}(x)$, substituting these formulae back into the original HJB equation, we have

$$0 = \max_u \{ R(x, u) - \alpha J(x) + \Gamma_1 J(x) \}$$

This is the present-value version of HJB equation.

In the scenario of Caldentey and Wein (2005), inventory process $X(t)$ is a Brownian motion under the heavy traffic condition and price $R(t)$ follows the geometric Brownian motion (GBM). For the discounted problem,

$$\min_{L, U} E \left\{ \int_0^\infty e^{-\alpha t} R(t) dL(t) + \int_0^\infty e^{-\alpha t} c(X(t)) dt \right\}$$

$$\text{Subject to } X(t) = Z(t) + L(t) - U(t)$$

$$dX(t) = \mu dt + \sigma_1 dW_1$$

$$dY(t) = \sigma_2 dW_2$$

$R(y) = R_1(r) + R_0 e^y$ represents the price in spot-market associated with special value y of $Y(t)$.

$Z(t)$ is the inventory process under which all demand orders are accepted and the machine is always active.

$L(t)$ is the cumulative rejected demand during $[0, t]$.

$U(t)$ is the cumulative loss in production output due to machine idling during $[0, t]$.

The decisions are associated with admission/rejection and production/idle. Consider a small time interval Δt at state (x, y) . There are three options: Increase the idling process $U(t)$, moving inventory instantaneously to $(x - \Delta x, y)$, or increase rejection process $L(t)$, moving inventory to $(x + \Delta x, y)$, or do nothing until the point of time $t + \Delta t$. In these three cases, we have, respectively,

$$V(x - \Delta x, y) = V(x, y) - V_x(x, y)\Delta x + o(\Delta x)$$

$$R(y)\Delta x + V(x + \Delta x, y) = V(x, y) + (R(y) + V_x(x, y))\Delta x + o(\Delta x)$$

$$E \left\{ \int_0^{t+\Delta t} e^{-\alpha s} c(X(s)) ds + e^{-\alpha(t+\Delta t)} V(X(t+\Delta t), Y(t+\Delta t)) \right\}$$

Based on the above three functions, applying infinitesimal operator defined previously, yields the Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \min \{ c(x) - \alpha V(x, y) + \Gamma_2 V(x, y), R(y) + V_x(x, y), -V_x(x, y) \}$$

Finally, numerical solution has been applied to solve the original problem. The structure of the optimal policy is characterized by two switching curves.

Maglaras (2003) considers both single-product and multi-product single-server make-to-order production system. For the single-product system, he formulates it as a fluid control problem by scaling time and state space. The linear fluid approximation yields a close-form solution. For the multi-product case, it is formulated as the diffusion control model with pricing decisions and the sequencing decisions. Before the introduction to fluid control model, we firstly take a look at the finite-horizon continuous-time deterministic control model.

$$\text{Objective function: } V(t, x) = \max_u \left\{ \int_t^T R(x, u, s) ds + B(T, x(T)) \right\}$$

$$\text{Subject to: } \dot{x}(t) = f(x(t), u(t))$$

Applying Bellman's Principle of Optimality to the above objective function,

$$\begin{aligned} V(t, x) &= \max_u \left\{ \int_t^{t+\Delta t} R(x, u, s) ds + V(t + \Delta t, x(t + \Delta t)) \right\} \\ &= \max_u \left\{ R(x, u, s)\Delta t + V(t + \Delta t, x(t + \Delta t)) \right\} \end{aligned}$$

Based on Principle of Smooth Fit, we assume that $V(t, x)$ has the required differentiability properties. Then, expand $V(t + \Delta t, x + \Delta x)$ into a first order Taylor series, obtaining

$$V(t + \Delta t, x(t + \Delta t)) = V(t, x) + V_t(t, x)\Delta t + V_x(t, x)f(x, u)\Delta t + o(\Delta t)$$

Substitute the above formula back into the objective function, yielding

$$V(t, x) = \max_u \left\{ R(x, u, s)\Delta t + V(t, x) + V_t(t, x)\Delta t + V_x(t, x)f(x, u)\Delta t + o(\Delta t) \right\}$$

Let $\Delta t \rightarrow 0$, and we have the Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \max_u \left\{ R(x, u, s) + V_t(t, x) + V_x(t, x)f(x, u) \right\}$$

For the finite-horizon problem, the optimal value function $V(t, x)$ depends on both time t and state variable x . For the infinite-horizon problem, the optimal value function $V(x)$ depends only on state variable x since from any time onward future looks the same. Here, we introduce an infinite-horizon problem in Maglaras (2003).

For simplification, we directly give the objective function and constraints regardless of its implication. The associated deterministic fluid control problem is as follows:

$$\max_{\lambda(t)} \int_0^{\infty} (r(\lambda(t)) - r(\mu) - hx(t)^2) dt$$

Subject to $\dot{x}(s) = \lambda(s) - \mu, \quad x(0) = z$

$\lambda(t)$ and μ are the demand arrival rate and production rate respectively. $r(\lambda(t))$ and $r(\mu)$ are the realized revenue and production cost respectively. $x(t)$ is the queue length. $hx(t)^2$ is the quadratic holding cost. Applying Bellman's Principle of Optimality, rewrite the objective function as:

$$\begin{aligned} V(x(t)) &= \max_{\lambda(s)} \int_t^{t+\Delta t} (r(\lambda(s)) - r(\mu) - hx(s)^2) ds + V(x(t + \Delta t)) \\ &= \max_{\lambda(t)} \left\{ (r(\lambda(t)) - r(\mu) - hx(t)^2)\Delta t + V(x(t)) + V'(x(t))(\lambda - \mu)\Delta t + o(\Delta t) \right\} \end{aligned}$$

Here, Taylor's series expansion is applied to $V(x(t + \Delta t))$ to obtain the above formula.

Rearrange the terms and let $\Delta t \rightarrow 0$, we derive the Hamilton-Jacobi-Bellman (HJB) equation.

$$r(\mu) + hx^2 = \max_{\lambda} \{r(\lambda) + V'(x)(\lambda - \mu)\}$$

Under some conditions, close-form solution can be obtained for this HJB equation.

3. Literature Review

3.1 Introduction

Revenue Management originated in problems on airlines seat bookings. In industry practice, revenue management methods prove to be very effective. Motivated by its success in airline booking system, management tries to extend the methodology of Revenue Management to other industries such as hospitality management, rental business, manufacturing systems, and so on. Currently, it gains increasing attentions from both academia and industries. Substantial literature has been discussing the theory and application of Revenue Management nowadays.

For airline seat booking problem, the objective is to maximize revenue by setting booking limits for the admission control of booking requests. Review will be given in section 3.2 on optimal seat allocation of two parallel flights.

For the network revenue management, the problem could be much more complex due to multiple legs involved in the system. It needs significant computation to find the exact solution. To make it applicable, researchers have proposed many heuristic and approximation algorithms. For such kind of problems and computation algorithms, readers can refer to Gallego and van Ryzin (1997), Bertsimas and Popescu (2003) and van Ryzin and Liu (2004). Multi-leg and round-trip problems could be classified into the problems of network revenue management. Readers can refer to section 3.3 for the review. One of the customer choice models has been developed in Talluri and van Ryzin (2004). It attempts to consider consumer behavior in airline seat booking. It arouses broad interest among researchers. Stochastic knapsack problem comes from the research of telecommunication network. It is related to multiple seats booking which could be a knapsack problem. Readers are referred to Kleywegt and Papastavrou (2001, 1998) and van Slyke and Young (2000).

For a more comprehensive review in airline revenue management, readers can refer to McGill and van Ryzin (1999). It gives a clear classification and discussion on the history, basic terminology and development of various subjects of revenue management.

Parallel to research on airline seat inventory control, another important stream of research in revenue management addresses the issues of dynamic pricing of perishable inventory. Typical examples are Gallego and van Ryzin (1994, 1997), Feng and Gallego (1995), Feng and Xiao (2000), Zhao and Zheng (2000), Lin (2004).

Recently revenue management has been extended into other industries such as rental business, retail, hotel, and restaurant. The sample research in these areas are Bitran and Gilbert (1996), Bertsimas and Shioda (2003), Gallego and Phillips (2004), Gallego et al. (2004) and Savin et al. (2005).

More recently, manufacturing firms have started to evaluate the use of revenue management. We will study a stock rationing problem in production/inventory system which is closely related to Revenue Management. A comprehensive review will be given in section 3.4.

3.2 Optimal Seat Allocation for a Two-Flight Problem

The research on airline revenue management can be traced back to Rothstein (1971) and Littlewood (1972) where Littlewood is concerned with a static two-fare, single-leg problem and Rothstein studies an overbooking problem. Belobaba (1989) extends Littlewood's work to the multi-fare situation. Curry (1990) studies seat allocation for a multi-fare, single-leg flight problem with continuous random demands, while the discrete random demand case is studied by Wollmer (1992). Brumelle and McGill (1993) investigate the single-leg, multi-fare flight. By using the optimal stopping rule, they derive a closed-form solution. Li and Oum (2002) review the works of Curry (1990), Wollmer (1992) and Brumelle and McGill (1993) and prove the equivalence of the optimality conditions for the above three static models.

Different from the static model, in a dynamic model, booking requests for all fare class may arrive at any time. In Lee and Hersh (1993), the booking horizon is divided into discrete time periods such that at most one event can occur in each period and the problem is formulated as discrete-time MDP. In the same way, Subramanian, et. al. (1999) and Lautenbacher and Stidham (1999) formulate the multi-fare single-leg flight problem as discrete-time Markov decision processes as well. Other works on single-leg, single-flight problems, see Kleywegt and

Papastavrou (1998), Chatwin (1999), Liang (1999), Zhao and Zheng (2001), and Brumelle and Walczak (2003).

There is a stream of research that addresses the interaction between consumers and the sellers. Examples are Talluri and van Ryzin (1999, 2004), Bertsimas and Popescu (2003), Gallego and van Ryzin (1994, 1997), Zhao and Zheng (2000), Feng and Xiao (2000), Feng and Gallego (2000), Maglaras (2003), and Zhang and Cooper (2005). For a more comprehensive review on this subject, refer to McGill and van Ryzin (1999).

There are many other papers that deal with the multi-flight problems. Feng and Xiao (2001) consider the case of multiple origins, one hub and one destination with capacity constraints on the origin-hub flights. The problem was formulated with a multi-dimension dynamic programming equation, which can then be decomposed into one-dimension equations due to independent demands and capacities for the origin-hub flights. In our formulation, the two-dimension problem can not be decomposed into one-dimension ones due to demand interactions between the two flights. Zhang and Cooper (2005) concentrate on the customer choice models and computational methods, whereas our paper aims to characterize the structure of the optimal booking policies. Gallego and Phillips (2004) consider a demand postponement problem with a similar scenario as ours. In their model it also considers three products with two specific products and one flexible product. Reservation horizon is divided into two periods only. Customers committing to the flexible product accepted in the first period will be assigned to a flight in the second period or denied with a penalty payment. In our model, the reservation horizon is divided into many periods, and the decision on whether to accept the booking request and which flight to accommodate it if the booking request is accepted has to be made instantaneously upon receiving the booking request. Iyer, Deshpande and Wu (2003) also discuss the demand postponement strategy for capacity planning. It is different from ours strategy which is on demand shifting in airline capacity planning. Talluri (2001) considers the network system with flexible type of demand. He proposes heuristic approximation to solve the problem. Different from the above papers considering revenue management at the operational level, Dana (1999) discusses the demand shifting issue which is

closely related to our topic from the economic perspective and demonstrates the benefit of price dispersion to reduce capacity costs through demand shifting.

3.3 Optimal Seat Allocation for the Round-Trip/Two-Leg Problem

In the 1970s, as a result of the deregulation of airline industry, airlines offered a variety of discounted fares of tickets to attract potential passengers to book in advance. To maximize the total revenue from selling the pool of tickets, it needs optimization techniques to allocate the seats to different fare classes. This comes about the field of airline yield management. The earliest academic research works are Rothstein (1971) and Littlewood (1972), the latter applies the marginal seat revenue (MSR) rule to a two-fare, single-leg problem. Belobaba (1987, 1989) extends Littlewood's work to multi-fare class problems and proposes the expected marginal seat revenue (EMSR) rule.

Generally, previous research on seat inventory control models can be classified as the static type models and dynamic type models. For the static models, booking requests are often assumed to arrive sequentially in order of fare classes. Low-fare classes arrive earlier than high-fare classes. By ignoring evolution details of the random arrival process of booking requests, static models are only concerned with total demand of each fare class in the booking period. Along this line of research, Curry (1990) obtains booking limits used in seat inventory control for single-leg flights. Wollmer (1992) considers seats allocation among multi-fare classes. Brumelle and McGill (1993) further investigate the multi-fare single-leg problems. Li and Oum (2002) prove the equivalence of the optimality conditions for these static models.

Compared with static models, dynamic models remove the assumption that booking requests of various fare classes arrive sequentially. They assume that booking requests in each fare class can arrive throughout the whole booking horizon. Lee and Hersh (1993) characterize the structure of the optimal booking policy as booking limits. Moreover, Subramanian, Stidham and Lautenbacher (1999) consider overbooking and cancellation in their studies. Lautenbacher and Stidham (1999) link both static and dynamic models through the common underlying dynamic program. Liang (1999) shows the optimality of booking limit policy in the continuous-time model.

Seat inventory control in the round-trip and multi-leg problems is one of the major issues in airline industry. Compared with the research on single-leg problems, there are only a few publications on the multi-leg problems. Ladany and Bedi (1977) and Hersh and Ladany (1978) deal with two-leg problems under the assumption that no passengers board on the flight at the intermediate stop. Wong, Koppelman and Daskin (1993) study the single fare class and multi-leg problem. You (1999) considers dynamic pricing model of the multi-fare class, multi-leg problem and derive structure results for the two-leg case. Feng and Lin (2001) investigate the monotone properties of the optimal booking policy for multi-leg flight problem in a continuous-time model.

For the multiple seats booking problems, Lee and Hersh (1993) and Brumelle and Walczak (2003) give an investigation on characterizing the optimal booking policy. It is shown in Lee and Hersh(1993) that booking limit policy is optimal when all booking requests demand identical number of seats and booking control is to accept all or none. If passenger demand can be partially satisfied, Brumelle and Walczak (2003) show that the booking limit policy, e.g. the so-called booking curve in their paper, is still optimal.

3.4 Stock Rationing of a Make-To-Stock Production/Inventory System

Stock rationing problems have been extensively studied in literatures. For the periodic-review models, Veinott (1965), Topkis (1968), Cohen, Kleindorfer and Lee (1988) and Frank, Zhang and Duenyas (2003) study the rationing policy for a dynamic single-product inventory system with several demand classes. Nahmias and Demmy (1981) and Deshpande, Cohen and Donohue (2003) consider the continuous review model. Gerchak, Parlar and Yee (1985) also give insightful results on inventory rationing and production in the discrete-time case. More recently, in the make-to-stock queue setting, Ha (1997a) formulates the single-item multi-class stock rationing problem with lost sales. The structure of the optimal policy is characterized as critical numbers called rationing levels. Ha (1997b) considers stock rationing with backordering under the presence of two priority demand classes. Ha (2000) further extends previous results with exponential processing time to the case of Erlang distributed processing time. de Vericourt, Karaesmen and

Dallery (2002) consider an extension of Ha (1997b) to multiple-demand class. Zhao, Deshpande and Ryan (2005) study stock rationing of the network systems.

Stock rationing is closely related to Revenue Management both in idea and methodology. In airline revenue management, seat inventory is often classified as several classes in terms of fares while in rationing problems, customers are classified according to their priority. Revenue Management tries to maximize total revenue through selling limited items or capacities while rationing tries to minimize total costs through the reservation of limited items and production capacities for high priority customers.

4. Optimal Seat Allocation for a Two-Flight Problem

4.1 Introduction

In airline seat reservation system, the fundamental problem of seat inventory control is to make the decisions whether to accept or deny a booking request for a particular fare at any point of time during the booking period. Airlines commonly offer a variety of fares of tickets to capture demands from different market segments. The objective is to maximize revenue by utilizing controls such as booking limits at various price levels. Most of previous studies have primarily focused on the single-flight, multi-fare and single-leg problems, which, according to Lautenbacher and Stidham (1999), can be classified into either static models where fare classes of booking requests come sequentially or dynamic models where requests for each fare class can arrive throughout the whole reservation period.

In this part, we study a two-flight model where there are two parallel flights between two cities in a day (e.g., one departs at 9:00am and another at 11:00am) and booking requests in each fare class arrive according to a random process. For two flights, the airline can offer three products for passengers to choose. Two are specific products for booking requests who are dedicated to one of the two flights only; whereas the third is a flexible product for those willing to take either flight. Upon a booking request realizing its product type, the airline has to decide whether to accept it, and in case a flexible type is accepted, which flight to accommodate it. This paper uncovers the structure of optimal booking policies through four monotone switching curves. We also present extensions of our basic model to the case of a flexible segment that is purely induced by price discounts and another case of implicit flexible segment that is caused by cross-flight buying downs through price incentives. Moreover, customer choice has been introduced into our original model. We demonstrate that the optimal booking policies under these three cases can still be characterized by switching curves with minor adjustments. Finally, numerical examples are used to illustrate the derivation and dynamics of the optimal booking policies.

Customers are usually willing to be *flexible* and purchase the *flexible product* because they are indifferent between two flights as long as they can book a seat at a fare class they want. This

may arise from many circumstances, even if there may be no price discount since the characteristic of a fare class normally goes beyond the price dimension, for example, using various fences and other conditions. Sometimes, knowing that his/her booking request for a specific flight may be rejected, a customer who is desperate to make the trip would be willing to purchase the flexible product so as to minimize the risk of denial. The incentive from price discount in these cases is relatively insignificant. Usually, the flexible booking request comes in a form of desired time of arrival through a direct call to the carrier, e.g., “I want to be at New York by 5:00pm, please book me a flight.”

We shall demonstrate that our model formulation may be easily modified to cater for the scenario that some customers are flexible purely by price factors, namely, the discount case. Most airlines are fully aware of these flexible segments for most of origin-destination markets. As far as we know, the existing literature on the multi-flight problem has not intensively addressed this phenomenon yet.

There are various ways to differentiate the demand. For example, the online booking form can be designed to capture such information as follows:

Your preferable flight

- Flight 1 Flight 2

Are you willing to take the flight other than your choice if it is necessary?

- Yes No

Figure 4.1 displays the framework of the booking procedure. Note that it is different from a sequential booking procedure where the booking requests are checked by Flight 1 and Flight 2 (or vice versa) sequentially.

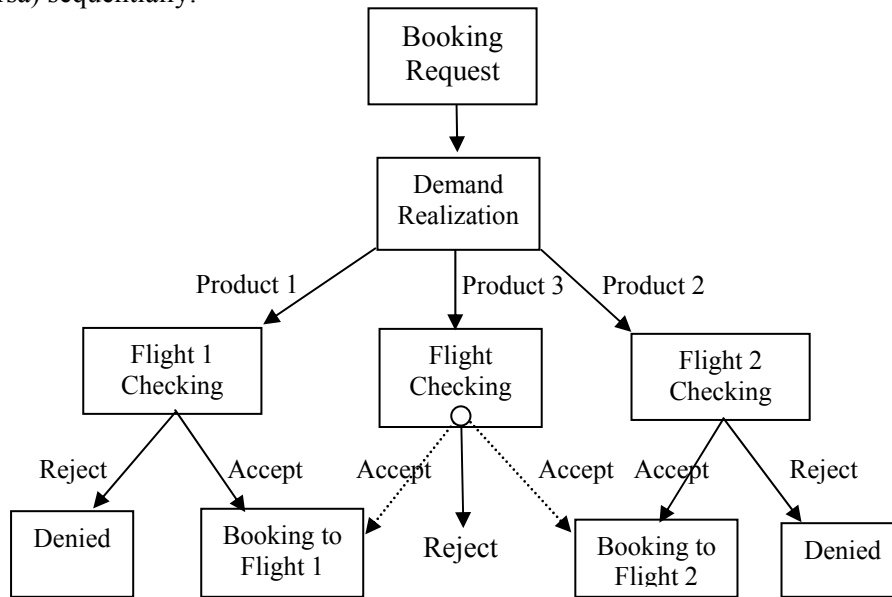


Figure 4.1: Booking Procedure for Passengers

It is worthwhile to highlight that a fare class here contains more information than the ticket price because there are various restrictions or conditions that are part of a fare class that are used to segment markets. This extension of a fare class is important in understanding why a customer is indifferent between two fare classes from two flights even though ticket prices may be different. The objective is to maximize the total expected revenue of the two flights. The basic research question is: What is the optimal booking policy, and more importantly, what special structure, if any, does the optimal policy possess?

There are other papers that deal with the related multi-flight problems. Feng and Xiao (2001) consider the case of multiple origins, one hub and one destination with capacity constraints on origin-hub flights. The problem is analyzed by a multi-dimension dynamic programming equation, which can then be decomposed into one-dimension equations due to independent demands for the origin-hub flights. Talluri (2001) considers the network system with demand routings. He proposes heuristic approximation to solve the problem. Gallego and Phillips (2004) also consider three types of customers. They divided the reservation horizon into two periods. Customer booking requests of

the third type accepted in the first period will be assigned to a flight in the second period or denied with a penalty payment. van Ryzin and Liu (2004) and Zhang and Cooper (2005) concentrate on the customer choice models and various computational methods. Based on previous studies our paper has the following aspects: (a) we will follow the traditional approach for dynamic models, namely, dividing the booking horizon into many “small” booking periods; (b) the random demands for the two flights are independent; (c) each booking request, regardless of which type, will be accepted or denied instantaneously; and (d) we intend to characterize the optimal booking policies by applying techniques developed in Hajek (1984), Glasserman and Yao (1994), Ha (1997), and most recently Chen (2004).

4.2 Model Formulation

Suppose that there are two flights between two cities in a period scheduled by an airline company. The company tries to maximize the total expected revenues of these two flights. As in Lee and Hersh (1993), Subramanian, et. al. (1999), and Lautenbacher and Stidham (1999), we assume that booking horizon can be divided into subintervals such that no more than one event happens in each subinterval, where an event is associated with a booking request for a particular fare class on one of these two flights. A booking request will purchase one of the three products. Two are specific products for one of the two flights respectively and the third is the flexible product that the booking request can be assigned to either the first flight or the second flight after being accepted. We make assumption that any rejected request will be lost to the airline (i.e., spill).

4.2.1 Dynamic Model Formulation

Assume that there are m fare classes, and for $i = 1, \dots, m$, r_i and R_i are the net revenue of fare class i on flight 1 and flight 2 respectively such that, without loss of generality, $r_1 \geq r_2 \geq \dots \geq r_m$ and $R_1 \geq R_2 \geq \dots \geq R_m$.

Let N be the total number of booking periods for the two flights; and periods be numbered consecutively by the number of periods remaining, i.e., period N is the starting period of booking and period 0 is the ending booking period. Let P_{ji}^n be the probability of a booking request of

product j for fare class i in period n , $j = 1, 2, 3$, and P_0^n be the probability that no event occurs in period n . It is clear that in period n , we have

$$P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n + P_{3i}^n) = 1.$$

For $k = 1, 2$, define C_k as the total seating capacity for flight k and x_k as the state variable indicating the number of seats already booked. Hence, $C_k - x_k$ is the remaining capacity for flight k .

Let $V_n(x_1, x_2)$ be the maximal expected revenue from period n to 0 when accepted seats on flight 1 and flight 2 are given x_1 and x_2 respectively. Then the dynamic recursive formulation of this problem is

$$\begin{aligned} V_n(x_1, x_2) = & P_0^n V_{n-1}(x_1, x_2) + \sum_{i=1}^m P_{1i}^n \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\ & + \sum_{i=1}^m P_{2i}^n \max \{R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\ & + \sum_{i=1}^m P_{3i}^n \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \end{aligned} \quad (4-1)$$

$$V_0(x_1, x_2) = -\bar{R}(x_1 - C_1)I_{\{x_1 > C_1\}} - \bar{R}(x_2 - C_2)I_{\{x_2 > C_2\}} \quad (4-2)$$

where $\bar{R} \geq \max_{1 \leq i \leq m} \{r_i, R_i\}$, and

$$I_{\{x_k > C_k\}} = \begin{cases} 1 & \text{if } x_k > C_k \\ 0 & \text{if } x_k \leq C_k \end{cases}$$

is the usual indicator function.

The boundary condition (4-2) is the same as that of the Lautenbacher and Stidham (1999). It is a technical treatment introduced for the convenience of computation. Moreover, it can avoid numerical rounding problem in computation. It is equivalent to the boundary condition that any state is zero in the last period or when it is out of seat capacity. On the other hand, if $x_k > C_k$ at the time of departure, then (4-2) can be interpreted as the penalty for overbooking. Note that, if flight 2 departs before flight 1, then P_{2i}^n and P_{3i}^n are set to 0 for those periods when the booking for flight 2

has been closed. The probabilities of the flexible booking requests are merged into those of the booking requests for the first flight.

Our problem has been formulated as a discrete-time Markov Decision Process (MDP), i.e., a sequential decision process, which is a system that evolves over time, with opportunities to influence its future evolution by taking admission/rejection actions at various discrete points of time. There is revenue incurred at these discrete points of time, that depend on the actions taken and the way in which the process evolves. The objective of the carriers is to maximize the total expected revenue over a certain finite booking horizon. Traditionally, in Revenue Management, the objective function is not associated with the discounting criterion, which is common in queueing control models and inventory control models. Before we proceed on, let us highlight the benefit of considering the flexibility of passengers.

4.2.2 Advantages of the Model with Identified Flexible Demand

It is clear that the fourth term in (4-1) captures the value of flexibilities from the flexible booking requests. As noted earlier, the booking process presented in Figure 4.1 is different from a sequential booking model. In a sequential booking system, when a flexible booking request comes in, Flight 1 (or Flight 2) will be first checked with a decision to accept or reject. Only when rejected, Flight 2 (or Flight 1) will then be checked next. In our model, after confirming that a booking request purchases the flexible product, the airline will dispose the booking to the flight that will yield the highest positive increase in expected revenue based on the booking status of the two flights at the time of request. If neither flight gives a positive return by accepting the booking request, then the request will be denied. Thus, this booking process improves by optimizing globally over the sequential one, which in turn improves over an independent system where each of the two flights has its own independent booking system. From the view of passengers, this flexibility produces a higher chance of acceptance to a flight when booking into a comparative system. Indeed, it can be readily shown mathematically that model (4-1) yields a higher value than independent booking systems.

Here we study the value of flexibility of customers by comparing two model formulations. Firstly, airline does not consider the flexibility of booking requests, the third type booking requests can be naturally divided into two streams of booking requests, one for flight one and the other for flight two. Here we assume that at period n , without loss of generality,

q_{3i}^n = the probability of the third type booking request for class i on flight 1 in period n .

$(P_{3i}^n - q_{3i}^n)$ is the probability of the third type booking request for class i on flight 2 in period n .

The dynamic programming equation is formulated as follow:

$$U_n(x_1, x_2) = P_0^n U_{n-1}(x_1, x_2) + \sum_{i=1}^m (P_{1i}^n + q_{3i}^n) \max \{r_i + U_{n-1}(x_1 + 1, x_2), U_{n-1}(x_1, x_2)\} \\ + \sum_{i=1}^m (P_{2i}^n + P_{3i}^n - q_{3i}^n) \max \{R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\} \quad (4-3)$$

$$U_0(x_1, x_2) = -\bar{R}(x_1 - C_1)I_{x_1 > C_1} - \bar{R}(x_2 - C_2)I_{x_2 > C_2}$$

The boundary condition is the same as (4-2). To demonstrate the value of considering the flexibility of booking requests, we will prove that $U_n(x_1, x_2)$ is less than $V_n(x_1, x_2)$ for any given x_1, x_2 .

Proposition 4.1 $V_n(x_1, x_2) \geq U_n(x_1, x_2)$ for all x_1, x_2 .

Proof. By induction, we prove our result in the finite period n . The initial condition is

$$V_0(x_1, x_2) = U_0(x_1, x_2) = -\bar{R}(x_1 - C_1)I_{x_1 > C_1} - \bar{R}(x_2 - C_2)I_{x_2 > C_2}$$

So the result holds for $n = 0$. Assume that $V_{n-1}(x_1, x_2) \geq U_{n-1}(x_1, x_2)$ for any x_1, x_2 , we need to prove that $V_n(x_1, x_2) \geq U_n(x_1, x_2)$. From (4-1) and (4-3), we have

$$\begin{aligned}
V_n(x_1, x_2) - U_n(x_1, x_2) &= P_0^n \{V_{n-1}(x_1, x_2) - U_{n-1}(x_1, x_2)\} \\
&+ \sum_{i=1}^m P_{1i}^n \{ \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} - \max \{r_i + U_{n-1}(x_1 + 1, x_2), U_{n-1}(x_1, x_2)\} \} \\
&+ \sum_{i=1}^m P_{2i}^n \{ \max \{R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} - \max \{R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\} \} \\
&+ \sum_{i=1}^m P_{3i}^n \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&- \sum_{i=1}^m q_{3i}^n \max \{r_i + U_{n-1}(x_1 + 1, x_2), U_{n-1}(x_1, x_2)\} \\
&- \sum_{i=1}^m (P_{3i}^n - q_{3i}^n) \max \{R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\}
\end{aligned}$$

For the first term of the RHS of the above difference equation, we have

$$P_0^n \{V_{n-1}(x_1, x_2) - U_{n-1}(x_1, x_2)\} \geq 0$$

For the second term of the RHS of the above difference equation,

based on $r_i + V_{n-1}(x_1 + 1, x_2) \geq r_i + U_{n-1}(x_1 + 1, x_2)$ and $V_{n-1}(x_1, x_2) \geq U_{n-1}(x_1, x_2)$

We have $\max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} - \max \{r_i + U_{n-1}(x_1 + 1, x_2), U_{n-1}(x_1, x_2)\} \geq 0$

Similarly, we can prove that the third term of the RHS is no less than 0. We are now ready to prove that the forth, fifth and sixth term of the RHS is no less than 0. Firstly, it is readily to check

$$\begin{aligned}
&\max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&\geq \max \{r_i + U_{n-1}(x_1 + 1, x_2), R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\}
\end{aligned}$$

By this inequality, we have

$$\begin{aligned}
&\sum_{i=1}^m P_{3i}^n \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&- \sum_{i=1}^m q_{3i}^n \max \{r_i + U_{n-1}(x_1 + 1, x_2), U_{n-1}(x_1, x_2)\} \\
&- \sum_{i=1}^m (P_{3i}^n - q_{3i}^n) \max \{R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\}
\end{aligned}$$

$$\begin{aligned}
&\geq \sum_{i=1}^m P_{3i}^n \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&\quad - \sum_{i=1}^m q_{3i}^n \max \{r_i + U_{n-1}(x_1 + 1, x_2), R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\} \\
&\quad - \sum_{i=1}^m (P_{3i}^n - q_{3i}^n) \max \{r_i + U_{n-1}(x_1 + 1, x_2), R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\} \\
&= \sum_{i=1}^m P_{3i}^n \{ \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&\quad - \max \{r_i + U_{n-1}(x_1 + 1, x_2), R_i + U_{n-1}(x_1, x_2 + 1), U_{n-1}(x_1, x_2)\} \} \\
&\geq 0
\end{aligned}$$

Therefore, the RHS of the difference equation is no less than 0 that is $V_n(x_1, x_2) - U_n(x_1, x_2) \geq 0$.

This completes the proof. ■

4.2.3 Structures of the Optimal Booking Policy

Unless it is feasible to decompose the two-dimensional dynamic programming problem (4-1) into a series of one-dimensional problems, tackling it directly for optimal solutions would be difficult. So in this paper, we will take a different approach by exploiting the monotonicity properties, such as componentwise concavity, submodularity and difference monotone, of the value function over a lattice.

First, define \hat{V} as the set of functions on Z_+^2 such that if $V \in \hat{V}$, then:

- (a) $V(x_1, x_2) - V(x_1 + 1, x_2) \quad \uparrow x_1 \uparrow x_2$
- (b) $V(x_1, x_2) - V(x_1, x_2 + 1) \quad \uparrow x_1 \uparrow x_2$
- (c) $V(x_1 + 1, x_2) - V(x_1, x_2 + 1) \quad \downarrow x_1 \uparrow x_2$

where \uparrow and \downarrow means nondecreasing and nonincreasing respectively. Properties (a) and (b) actually address that $V(x_1, x_2)$ is componentwise concave and submodular. Property (c) is called subconcave similar to the concept of “ μ -difference monotone” in Chen (2004). For simplicity, we will use “increase” to mean “nondecrease” and “decrease” to mean “nonincrease” throughout the paper.

We now define the following operators T_1, T_2, T_3 and T :

$$T_1(V, i) = \max \{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\}$$

$$T_2(V, i) = \max \{R_i + V(x_1, x_2 + 1), V(x_1, x_2)\}$$

$$T_3(V, i) = \max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\}$$

$$T(V, n) = P_0^n V(x_1, x_2) + \sum_{i=1}^m P_{1i}^n T_1(V, i) + \sum_{i=1}^m P_{2i}^n T_2(V, i) + \sum_{i=1}^m P_{3i}^n T_3(V, i)$$

Lemma 4.1 If $V \in \hat{V}$, then $T_1(V, i), T_2(V, i), T_3(V, i)$, and $T(V, n) \in \hat{V}$.

Proof. See the Appendix. ■

Remark: To prove lemma 4.1, we could use the technique in Hajek (1984) and Ha (1997) which simplifies the case-by-case checking but cannot avoid it completely. Alternatively, we can construct our problem by using the event-based operators in Koole (1998). Here, we choose to check each case, similar to Song and Sun (1999). The checking cases will be greatly reduced by noticing the fact that results are evident under taking same control actions. To be noted, some cases may not exist even though we check them all. For instance, when overbooking occurs, e.g. state variable x_1, x_2 exceed the seat capacities on two flights, we check it as if overbooking is allowed regardless of the capacity constraints. Our final boundary condition which has a high penalty on overbooking ensures that it is always optimal to reject the overbooking requests. This treatment will remove discussion on the boundary cases of the state variables in our proofs. It can be regarded as another benefit from the technical treatment of boundary condition (4-2).

Lemma 4.2 $V_n(x_1, x_2)$ satisfies properties (a), (b), and (c), i.e., $V_n(x_1, x_2) \in \hat{V}$, $n = 0, 1, \dots, N$.

Proof. It can be readily checked that $V_0(x_1, x_2) \in \hat{V}$. Thus, the theorem can be proved inductively by applying Lemma 4.1 to the equation $V_n(x_1, x_2) = T(V_{n-1}, n)$. ■

To continue our discussion, define two difference operators as follow:

$$D_1 V(x_1, x_2) = V(x_1 + 1, x_2) - V(x_1, x_2) \text{ and } D_2 V(x_1, x_2) = V(x_1, x_2 + 1) - V(x_1, x_2).$$

The next two propositions demonstrate that there are upper bound and lower bound for the above two difference operators.

Proposition 4.2 $D_1V_n(x_1, x_2) \leq -r_1$, when $x_1 \geq C_1$; $D_2V_n(x_1, x_2) \leq -R_1$, when $x_2 \geq C_2$

Proof. The proof is by induction on n . It is true for $V_0(x_1, x_2)$. Then

$$\begin{aligned} D_1V_n(x_1, x_2) &= V_n(x_1 + 1, x_2) - V_n(x_1, x_2) = P_0^n D_1V_{n-1}(x_1, x_2) \\ &+ \sum_{i=1}^m P_{1i}^n \left\{ \max \{r_i + V_{n-1}(x_1 + 2, x_2), V_{n-1}(x_1 + 1, x_2)\} - \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \right\} \\ &+ \sum_{i=1}^m P_{2i}^n \left\{ \max \{R_i + V_{n-1}(x_1 + 1, x_2 + 1), V_{n-1}(x_1 + 1, x_2)\} - \max \{R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \right\} \\ &+ \sum_{i=1}^m P_{3i}^n \left\{ \max \{r_i + V_{n-1}(x_1 + 2, x_2), R_i + V_{n-1}(x_1 + 1, x_2 + 1), V_{n-1}(x_1 + 1, x_2)\} \right. \\ &\left. - \max \{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \right\} \end{aligned}$$

For the second term of the above equation, we consider two cases,

$$\begin{aligned} &r_i + V_{n-1}(x_1 + 2, x_2) - \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\ &\leq r_i + V_{n-1}(x_1 + 2, x_2) - r_i - V_{n-1}(x_1 + 1, x_2) = D_1V_{n-1}(x_1 + 1, x_2) \leq -r_1 \\ &V_{n-1}(x_1 + 1, x_2) - \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\ &\leq V_{n-1}(x_1 + 1, x_2) - V_{n-1}(x_1, x_2) = D_1V_{n-1}(x_1, x_2) \leq -r_1 \end{aligned}$$

yielding $\max \{r_i + V_{n-1}(x_1 + 2, x_2), V_{n-1}(x_1 + 1, x_2)\} - \max \{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \leq -r_1$.

Further, by the same argument, both the third term and fourth term in the big bracket are less than or equal to $-r_1$. Thus, our result $D_1V_n(x_1, x_2) \leq -r_1$ follows. Similarly, we can prove $D_2V_n(x_1, x_2) \leq -R_1$. ■

Proposition 4.2 implies that it is always to reject an overbooking request since $D_1V_n(x_1, x_2) \leq -r_1 \leq -r_2 \leq \dots \leq -r_m$ and $D_2V_n(x_1, x_2) \leq -R_1 \leq -R_2 \leq \dots \leq -R_m$.

Proposition 4.3 $D_1V_n(x_1, x_2) \geq -r_1$, when $x_1 < C_1$; $D_2V_n(x_1, x_2) \geq -R_1$, when $x_2 < C_2$

Proof. We prove it by induction on n . It is true for $V_0(x_1, x_2)$ and we assume the same for $V_{n-1}(x_1, x_2)$. First, from equation (4-1),

$$\begin{aligned}
r_1 + V_n(x_1 + 1, x_2) &= P_0^n \{r_1 + V_{n-1}(x_1 + 1, x_2)\} + \sum_{i=1}^m P_{1i}^n \{r_1 + \max\{r_i + V_{n-1}(x_1 + 2, x_2), V_{n-1}(x_1 + 1, x_2)\}\} \\
&+ \sum_{i=1}^m P_{2i}^n \{r_1 + \max\{R_i + V_{n-1}(x_1 + 1, x_2 + 1), V_{n-1}(x_1 + 1, x_2)\}\} \\
&+ \sum_{i=1}^m P_{3i}^n \{r_1 + \max\{r_i + V_{n-1}(x_1 + 2, x_2), R_i + V_{n-1}(x_1 + 1, x_2 + 1), V_{n-1}(x_1 + 1, x_2)\}\} \\
&\geq P_0^n V_{n-1}(x_1, x_2) + \sum_{i=1}^m P_{1i}^n \max\{r_i + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\
&+ \sum_{i=1}^m P_{2i}^n \max\{R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&+ \sum_{i=1}^m P_{3i}^n \max\{r_i + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\
&= V_n(x_1, x_2)
\end{aligned}$$

To be noted that at the boundary if $x_1 + 2 = C_1 + 1$, then $r_i + V_{n-1}(x_1 + 2, x_2) \leq V_{n-1}(x_1 + 1, x_2)$ by Proposition 4.2. For $D_2V(x_1, x_2) \geq -R_1$, the same proof process applies, resulting in our result. ■

Proposition 4.3 is quite intuitive, which implies that it is always to accept the first fare class whenever there are seats available on the two flights since the booking in the first fare class can bring in the largest revenue. So, it is unreasonable to reject it.

In order to examine the structure of the optimal policy, we define four functions as follows:

$$S_{1n}(x_2, i) = \min\{x_1 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \text{ given } x_2\}$$

$$S_{2n}(x_1, i) = \min\{x_2 \mid V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - R_i \geq 0, \text{ given } x_1\}$$

$$S_{3n}(x_1, i) = \min\{x_2 \mid \begin{array}{l} V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \\ V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - R_i \geq 0, \text{ given } x_1 \end{array}\}$$

$$S_{4n}(x_1, i) = \min\{x_2 \mid V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1) + r_i - R_i \geq 0, \text{ given } x_1\}$$

where $n = 1, 2, \dots, N$.

Theorem 4.1

- (i) $S_{1n}(x_2, i) \downarrow x_2$; $S_{2n}(x_1, i)$, $S_{3n}(x_1, i) \downarrow x_1$; $S_{4n}(x_1, i) \uparrow x_1$;

- (ii) The optimal booking policy is characterized by four monotone switching curves $S_{1n}(x_2, i)$, $S_{2n}(x_1, i)$, $S_{3n}(x_1, i)$ and $S_{4n}(x_1, i)$ upon receiving a booking request, given the current reservations on hand on two flights (x_1, x_2) .
- (iii) The ‘slope’ of $S_{1n}(x_2, i)$ is less than or equal to -1; while that of $S_{2n}(x_1, i)$ is greater or equal to -1.

Proof. The proof is very much the same as the one used in proving Proposition 3 in Chen (2004). For completeness, we briefly discuss it here.

For (i), take $S_{1n}(x_2, i)$ first. First, it is clear that the set

$$\{x_1 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \text{ given } x_2\}$$

is nonempty since $V_n(C_1, x_2) - V_n(C_1 + 1, x_2) - r_i \geq 0$ by Proposition 4.2. By the definition of $S_{1n}(x_2, i)$ and property (a), we have

$$\begin{aligned} & V_n(S_{1n}(x_2, i), x_2 + 1) - V_n(S_{1n}(x_2, i) + 1, x_2 + 1) - r_i \\ & \geq V_n(S_{1n}(x_2, i), x_2) - V_n(S_{1n}(x_2, i) + 1, x_2) - r_i \geq 0, \end{aligned}$$

which implies that $S_{1n}(x_2, i) \in \{x_1 \mid V_n(x_1, x_2 + 1) - V_n(x_1 + 1, x_2 + 1) - r_i \geq 0, \text{ given } x_2\}$.

Therefore, by the definition of $S_{1n}(x_2 + 1, i)$, it follows that $S_{1n}(x_2 + 1, i) \leq S_{1n}(x_2, i)$.

Similarly, one can prove that $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$ are decreasing in x_1 .

By the definition of $S_{4n}(x_1, i)$ and property (c),

$$\begin{aligned} & V_n(x_1, S_{4n}(x_1, i)) - V_n(x_1 - 1, S_{4n}(x_1, i) + 1) + r_i - R_i \\ & \geq V_n(x_1 + 1, S_{4n}(x_1, i)) - V_n(x_1, S_{4n}(x_1, i) + 1) + r_i - R_i \geq 0. \end{aligned}$$

Thus, $S_{4n}(x_1, i) \in \{x_2 \mid V_n(x_1, x_2) - V_n(x_1 - 1, x_2 + 1) + r_i - R_i \geq 0, \text{ given } x_1\}$, and therefore $S_{4n}(x_1 - 1, i) \leq S_{4n}(x_1, i)$.

For (ii), note that by the monotone property of $[V_n(x_1, x_2) - V_n(x_1 + 1, x_2)]$ and the definition of $S_{1n}(x_2, i)$, we have for any given x_2 ,

$$V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \text{ for all } x_1 \geq S_{1n}(x_2, i); \text{ and}$$

$$V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i < 0, \text{ for all } x_1 < S_{1n}(x_2, i).$$

That is, for a given x_2 value (i.e., the number of seats on Flight 2 already booked), when $x_1 \geq S_{1n}(x_2, i)$ (i.e., on the right side of $S_{1n}(x_2, i)$), it is optimal to reject a booking for the first flight if requested. Conversely, it is optimal to accept a booking for the first flight if requested when $x_1 < S_{1n}(x_2, i)$ (on the left side of curve $S_{1n}(x_2, i)$). Thus, the curve $S_{1n}(x_2, i)$ separates the acceptance and rejection region for flight 1 requests, as illustrated in Figure 4.2.

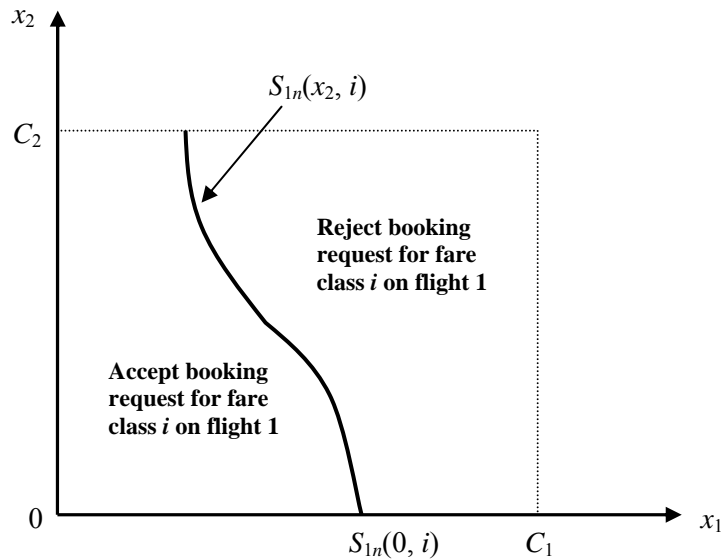


Figure 4.2: Switching Curve for the Booking Limits of Fare Class i on Flight 1

Note that the critical number $S_{1n}(x_2, i)$ is lower with a bigger x_2 value. This confirms with the intuition: when many of the Flight 2 seats are booked out, more flexible requests in the future can thus be allocated to Flight 1, and hence it can afford to wait instead of rushing now for a current request booking.

It is clear then that $S_{2n}(x_1, i)$ separates the acceptance and rejection region for flight 2 requests, as illustrated in Figure 4.3.

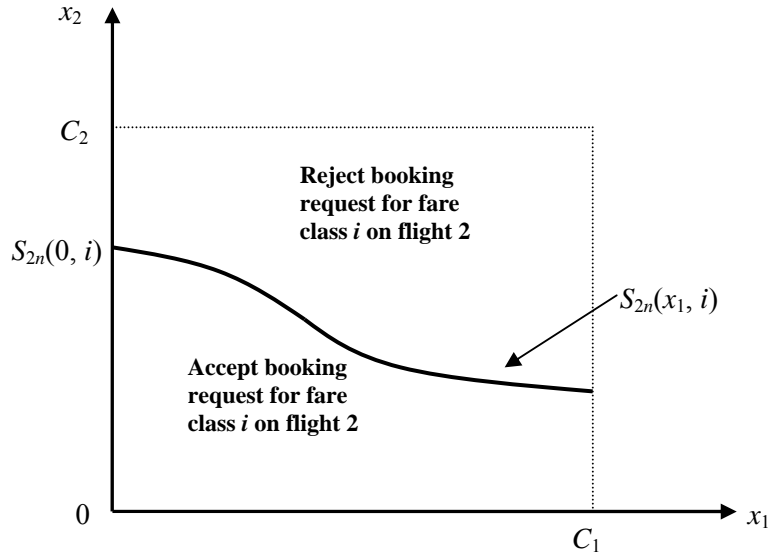


Figure 4.3: Switching Curve for the Booking Limits of Fare Class i on Flight 2

Finally, $S_{3n}(x_1, i)$ is the switching curve separating the acceptance and rejection regions for flexible requests, and when accepted, switching curve $S_{4n}(x_1, i)$ determines which of the two flights the request is to be booked to, as presented in Figure 4.4.

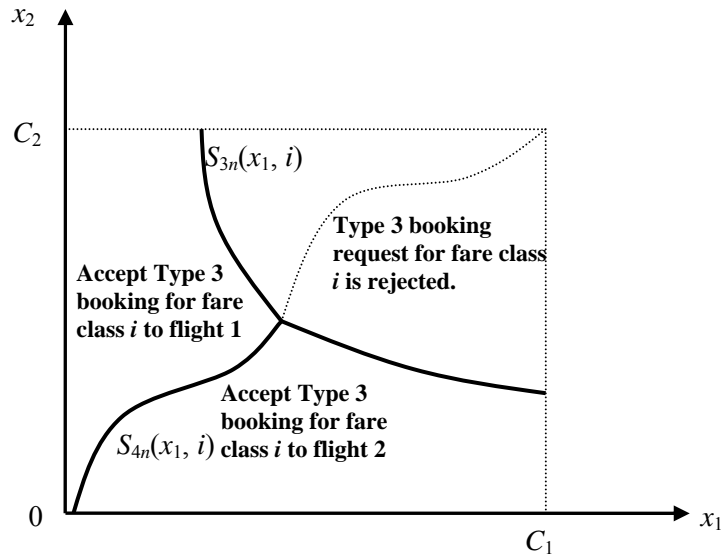


Figure 4.4: Switching Curve for the Booking Limits of Class i Flexible Passengers

The reasoning goes as follows. First of all, the rejection region for flexible requests is the joint rejection regions for bookings on flight 1 and flight 2, i.e., the set in the definition of

curve $S_{3n}(x_1, i)$. By the definition of $S_{4n}(x_1, i)$ and property (c), at a given x_1 value, for all $x_2 \geq S_{4n}(x_1, i)$ (i.e., on the above of curve $S_{4n}(x_1, i)$), we have

$$V_n(x_1 + 1, x_2) + r_i - (V_n(x_1, x_2 + 1) + R_i) = V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1) + r_i - R_i \geq 0,$$

which means that the revenue increase from a request booking to Flight 1 is at least as high as that from a Flight 2 request booking. Conversely, when $x_2 < S_{4n}(x_1, i)$ (below curve $S_{4n}(x_1, i)$), the marginal revenue from a Flight 2 booking is higher than that from Flight 1.

Finally for (iii), we introduce the ‘slope’ of a curve to determine its shape precisely. To show that the ‘slope’ of the curve $S_{1n}(x_2, i)$ is less than or equal to -1, let (x_1, x_2) be any point on the curve $S_{1n}(x_2, i)$ (see Figure 4.5).

Then, $V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0$, and $(x_1 - 1, x_2)$, $V_n(x_1 - 1, x_2) - V_n(x_1, x_2) - r_i < 0$.

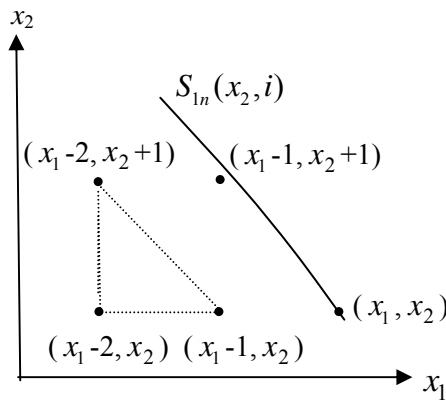


Figure 4.5: Properties of the Switching Curve

At point $(x_1 - 2, x_2 + 1)$, by property (c), we have

$$\begin{aligned} & V_n(x_1 - 2, x_2 + 1) - V_n(x_1 - 1, x_2 + 1) - r_i \\ &= V_n(x_1 - 2, x_2 + 1) - V_n(x_1 - 1, x_2) + V_n(x_1 - 1, x_2) - V_n(x_1 - 1, x_2 + 1) - r_i \\ &\leq V_n(x_1 - 1, x_2 + 1) - V_n(x_1, x_2) + V_n(x_1 - 1, x_2) - V_n(x_1 - 1, x_2 + 1) - r_i \\ &= V_n(x_1 - 1, x_2) - V_n(x_1, x_2) - r_i < 0. \end{aligned}$$

Thus, point $(x_1 - 2, x_2 + 1)$ is strictly on the left side of curve $S_{1n}(x_2, i)$, then point $(x_1 - 1, x_2 + 1)$ should be on the left of or on the curve $S_{1n}(x_2, i)$. Therefore, the slope of $S_{1n}(x_2, i)$ cannot be greater (flatter) than -1. By symmetry, the slope of $S_{2n}(x_1, i)$ cannot be less (steeper) than -1.

$S_{1n}(x_2, i)$ and $S_{2n}(x_1, i)$ are portrayed in Figures 4.2 and 4.3, respectively, based on these properties. ■

For (i), in figure 4.2, as x_2 increases, the switching function $S_{1n}(x_2, i)$ decreases in x_2 which means it is less likely to accept the flight 1 booking request when there are already many seats booked out in flight 2. Intuitively speaking, it is more likely to route the flexible booking requests to flight 1 in case more seats are occupied on flight 2. As a result, to reserve seats for flexible passengers, carriers tend to reject flight 1 booking requests. Same interpretation applies to $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$. For (ii), $S_{3n}(x_1, i)$ is the upper envelop of $S_{1n}(x_2, i)$ and $S_{2n}(x_1, i)$. It is evident that the flexibility of booking expands the acceptance region, which creates a win-win situation for the airline and the customers. For $S_{4n}(x_1, i)$, above it, it is always better to accept the booking for flight 1 since the number of accepted to flight 2 is relatively greater than that of accepted to flight 1. Same interpretation applied to the area below the curve. For (iii), we have determined that the ‘slope’ of the curve $S_{1n}(x_2, i)$ is less than or equal to -1 and $S_{2n}(x_1, i)$ is greater than or equal to -1. Furthermore, it means that the two switching curves have at most one intersection point. That is, if they intersect at some point they will go away and never meet again. To be noted, actually, at the intersection point, the slopes of both switching curves are 0.

It is important to highlight that the switching curves are different from the booking curves or protection levels commonly known in static models. The key difference is that there are a lot of switching curves: four switching curves for each pair (r_i, R_i) at each decision period for each fare class $i, n = 1, 2, \dots, N$. Since there is at most one booking request in each decision period, only the relevant switching curve(s) will be used in the decision-making process.

4.2.4 Properties of the Optimal Value Function and Switching Curves

Parallel to that of the single-flight problem, the optimal value function of the two-flight problem has many common properties. In this subsection, we will further discuss properties of the optimal value function and the underlying switching curves, which will help to improve our understanding on the optimal booking policy in a dynamic setting. The next result shows that the

value function has a special submodularity property. It implies the existence of time threshold for each class of flight 1 and flight 2 passengers.

Proposition 4.4 $V_n(x_1, x_2)$ is submodular in (x_1, n) and (x_2, n) respectively.

Proof. To prove V_n is submodular in (x_1, n) , it suffices to show that $D_1V_n(x_1, x_2)$ is decreasing in n .

Since $P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n + P_{3i}^n) = 1$, it follows that

$$\begin{aligned}
 & D_1V_n(x_1, x_2) - D_1V_{n-1}(x_1, x_2) \\
 &= (V_n(x_1 + 1, x_2) - V_{n-1}(x_1 + 1, x_2)) - (V_n(x_1, x_2) - V_{n-1}(x_1, x_2)) \\
 &= \sum_{i=1}^m P_{1i}^n \left[\max \{r_i + D_1V_{n-1}(x_1 + 1, x_2), 0\} - \max \{r_i + D_1V_{n-1}(x_1, x_2), 0\} \right] \\
 &\quad + \sum_{i=1}^m P_{2i}^n \left[\max \{R_i + D_2V_{n-1}(x_1 + 1, x_2), 0\} - \max \{R_i + D_2V_{n-1}(x_1, x_2), 0\} \right] \\
 &\quad + \sum_{i=1}^m P_{3i}^n \left[\max \{r_i + D_1V_{n-1}(x_1 + 1, x_2), R_i + D_2V_{n-1}(x_1 + 1, x_2), 0\} \right. \\
 &\quad \quad \left. - \max \{r_i + D_1V_{n-1}(x_1, x_2), R_i + D_2V_{n-1}(x_1, x_2), 0\} \right]
 \end{aligned} \tag{4-4}$$

By Lemma 4.2, $V_n(x_1, x_2)$ is componentwise concave and submodular in (x_1, x_2) . Hence we have

$$D_1V_{n-1}(x_1 + 1, x_2) \leq D_1V_{n-1}(x_1, x_2) \text{ and } D_2V_{n-1}(x_1 + 1, x_2) \leq D_2V_{n-1}(x_1, x_2).$$

Applying the above two inequalities to each of the three terms in the right hand side of (4.4) will lead to the conclusion that $D_1V_n(x_1, x_2) - D_1V_{n-1}(x_1, x_2) \leq 0$, i.e., $D_1V_n(x_1, x_2)$ is decreasing in n .

Similarly, it can be shown that $D_2V_n(x_1, x_2)$ is decreasing in n . ■

From Proposition 4.4, the monotonicity in period of the marginal increment of value function implies that there exists time threshold for each capacity pairs and each fare class of flight 1 and flight 2 bookings. As a result, we can define critical period

$$n_1^*(x_1, x_2, i) = \min \{n \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0\}$$

For $n \geq n_1^*$, at current state (x_1, x_2) , it is optimal to reject the booking of fare class i of flight 1 since $V_n(x_1, x_2) \geq r_i + V_n(x_1 + 1, x_2)$; otherwise, when $n < n_1^*$, it is optimal to accept the booking.

Similarly, we can define

$$n_2^*(x_1, x_2, i) = \min \{n \mid V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - R_i \geq 0\}$$

For $n \geq n_2^*$, at (x_1, x_2) , it is optimal to reject the booking of fare class i of flight 2; otherwise, it is optimal to accept the booking.

For the flexible bookings, it needs both accept/reject and the switching control between flight 1 and flight 2. For the accept/reject control on flexible bookings, we can simply define the critical period by

$$n_3^* = \max \{n_1^*, n_2^*\}$$

Then we can use n_3^* as time threshold for the accept/reject control of the flexible bookings. It is optimal to reject a flexible booking when $n \geq n_3^*$. Otherwise, accept it.

In terms of the switching control of the flexible bookings, unfortunately, we cannot derive the monotone property in period of the difference $V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1)$. Thus, we have to resort to the switching function $S_{4n}(x_1, i)$ defined previously to decide which flight to accommodate the flexible bookings if accepted in each period.

Proposition 4.5 $V_n(x_1, x_2) \uparrow r_i \uparrow R_i \downarrow x_1 \downarrow x_2 \uparrow P_{3i}^n$

Proof. By induction, we can easily check the monotone properties in r_i, R_i, x_1 and x_2 . For $\uparrow P_{3i}^n$, the value function increases with P_{3i}^n which demonstrate the value of flexible demand. Notice the facts that

$$\max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \geq V(x_1, x_2)$$

$$\max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \geq \max \{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\}$$

$$\max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \geq \max \{R_i + V(x_1, x_2 + 1), V(x_1, x_2)\}$$

By induction, our result follows naturally. Details are left to readers. ■

The value function increases with P_{3i}^n , which demonstrates the value of flexible demand.

Next, we are going to discuss the monotonic properties of switching curves. The following proposition says that switching curves are monotone in fare classes. It corresponds to the nested structure of fare classes in previous research.

Proposition 4.6 $S_{1n}(x_2, i) \uparrow r_i$; $S_{2n}(x_1, i) \uparrow R_i$; $S_{3n}(x_1, i) \uparrow r_i$ and $\uparrow R_i$; $S_{4n}(x_1, i) \uparrow (R_i - r_i)$.

Proof: The proofs follow easily from the definition. ■

Proposition 4.7 The switching curves $S_{1n}(x_2, i)$, $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$ are decreasing in n .

Proof: By the definition of $S_{1n-1}(x_2, i)$ and Proposition 4.4, we have

$$\begin{aligned} & V_n(S_{1n-1}(x_2, i), x_2) - V_n(S_{1n-1}(x_2, i) + 1, x_2) - r_i \\ & \geq V_{n-1}(S_{1n-1}(x_2, i), x_2) - V_{n-1}(S_{1n-1}(x_2, i) + 1, x_2) - r_i \geq 0, \end{aligned}$$

hence, $S_{1n-1}(x_2, i) \in \{x_1 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \text{ given } x_2\}$, and therefore,

$$S_{1n}(x_2, i) \leq S_{1n-1}(x_2, i).$$

Similarly, $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$ decrease in n . ■

Proposition 4.7 implies that the rejection region is getting smaller as time goes toward the end of booking horizon. It ensures that the critical numbers are consistent over time in the sense that it would not run into a contradictory case where some accepted requests are to be denied later – if the critical number for a later period (smaller n) is smaller than an earlier (bigger n) one.

In the following sections, we will examine three extensions to our basic model as developed in the previous section. The first extension considers the case that customers are flexible because of a price discount; while the second extension considers the case that some customers can be induced to buy a ticket from a less desired flight by a discount, a kind of *buying down* behavior across two flights. This is a case that a *flexible segment is implicit*. The third extension covers the issue on customer choice.

4.3 Flexible Segment induced by Discount

In this subsection, we will look into the situation where the flexible segment is mainly driven by price factors, namely, they remain to be flexible as long as there is some discount on the fare class that they intend to book. It is clear from earlier analysis that flexible requests have the advantage of enabling better utilization of flight capacities through risk pooling in the booking process. If necessary, a lower price $p_{1i} < p_{2i} < \min\{r_i, R_i\}$ may be offered to attract customers to

opt for ‘flexible flight’ similar to that in Gallego and Phillips (2004). By replacing r_i and R_i with p_{1i} and p_{2i} in the last term in (4-1), all of our earlier analysis and hence results apply. However, we have to redefine the switching functions of $S_{3n}(x_1, i)$ and $S_{4n}(x_1, i)$ as follow

$$S_{3n}(x_1, i) = \min \{x_2 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - p_{1i} \geq 0, \\ V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - p_{2i} \geq 0\}$$

$$S_{4n}(x_1, i) = \min \{x_2 \mid V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1) + p_{1i} - p_{2i} \geq 0\}$$

where $n = 1, 2, \dots, N$.

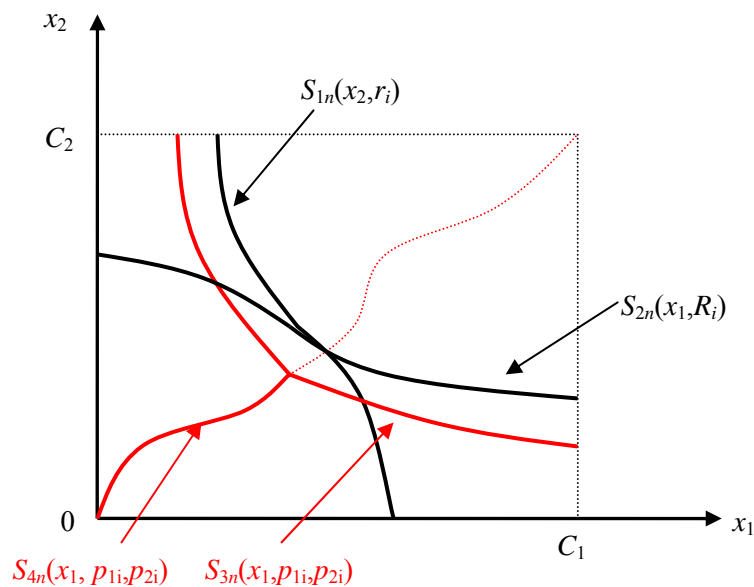


Figure 4.6: Switching Curves for the Price Discount Case

The slight change in $S_{3n}(x_1, i)$ and $S_{4n}(x_1, i)$ in the discount case has been shown in the above figure. One point should be mentioned is that the offer of discount fares may generate higher flexible demand but at the same time cannibalize some demand from the first and second type of demand that are for specific flights. Simulation results in Gallego and Phillip (2004) show that offering discount fares above certain level can increase total revenue. However, too much discount for the fares of the flexible demand may lead to loss in total revenue due to the cannibalizing effect.

4.4 An Implicit Flexible Segment Induced by Cross-Flight Buying Down

In some cases, it is difficult to identify the flexible demand, i.e., flexible customers in this paper. We can only identify the demand for specific flight. However, as it is often the case, some passengers for specific flight could be persuaded to switch to other flights with some discounts or benefits. Hence, management could use the discount strategy or any other benefits to encourage passengers to be flexible.

We can reformulate our problem with only two types of demand which are dedicated to one of the two flights. Furthermore, we assume that these two types of demand can be induced to switch over with each other at some discount price. In this scenario, a firm booking for the fare class i on flight 1 can be persuaded to switch to flight 2 for the same fare class at a discount price $R_i^d < R_i$. Similarly, a firm booking for the fare class i on flight 2 can switch to flight 1 for the same fare class at a discount price $r_i^d < r_i$. Then the recursive dynamic programming equation becomes

$$\begin{aligned}
 V_n(x_1, x_2) = & P_0^n V_{n-1}(x_1, x_2) \\
 & + \sum_{i=1}^m P_{1i}^n \max \{ r_i + V_{n-1}(x_1 + 1, x_2), R_i^d + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2) \} \\
 & + \sum_{i=1}^m P_{2i}^n \max \{ r_i^d + V_{n-1}(x_1 + 1, x_2), R_i + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2) \}
 \end{aligned} \tag{4-5}$$

$$P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n) = 1$$

The boundary condition is the same as (4-2). For model (4-5), it is clear that the structural properties of optimal value function are (a), (b) and (c) in section 4.2.3. All of our earlier analysis and results apply. Here, each of the two types of booking requests has three switching curves for each class i at some time period. The switching curves have small alteration from previous ones, owing to the discount price R_i^w and r_i^w in model (4-5). Here, we can define four switching curves that are similar to our previous models:

$$S_{3n}(x_1, r_i, R_i^d) = \min\{x_2 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq 0, \\ V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - R_i^d \geq 0; \text{ given } x_1\};$$

$$S_{3n}(x_1, r_i^d, R_i) = \min\{x_2 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i^d \geq 0, \\ V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - R_i \geq 0; \text{ given } x_1\}.$$

$$S_{4n}(x_1, r_i, R_i^d) = \min\{x_2 \mid V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1) + r_i - R_i^d \geq 0, \text{ given } x_1\}$$

$$S_{4n}(x_1, r_i^d, R_i) = \min\{x_2 \mid V_n(x_1 + 1, x_2) - V_n(x_1, x_2 + 1) + r_i^d - R_i \geq 0, \text{ given } x_1\}$$

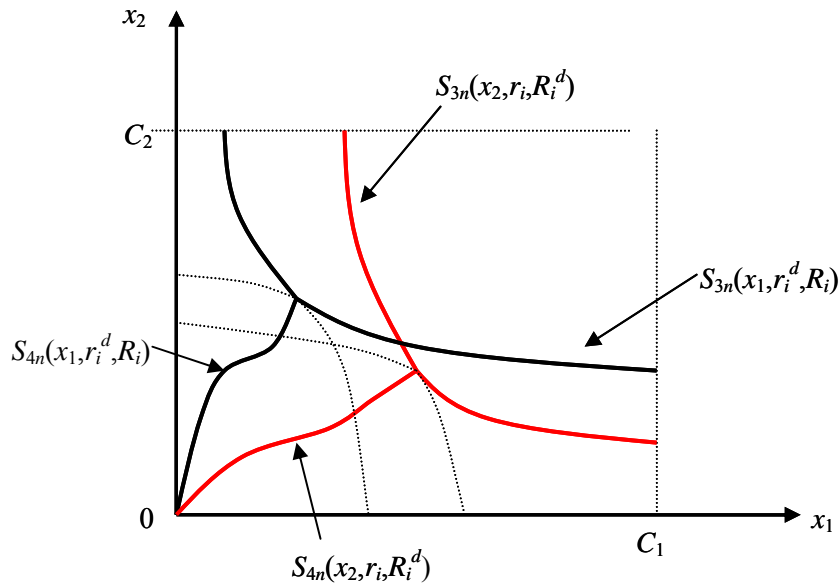


Figure 4.7: Switching Curves for the Cross-Flight Buying Down Case

Here , $S_{3n}(x_1, r_i, R_i^d)$ is a switching curve separating the acceptance and rejection region of fare class i on flight 1 and $S_{3n}(x_1, r_i^d, R_i)$ is the switching curve separating the acceptance and rejection region of fare class i on flight 2. $S_{4n}(x_1, r_i, R_i^d)$ and $S_{4n}(x_1, r_i^d, R_i)$ are used to determine which flight a booking request should be disposed to if it is accepted.

It is evident that the induced flexibility for each type of booking expands the acceptance region, which creates a win-win situation for the airline and the passenger.

4.5 Model with Passenger Choice

It is more interesting and realistic to include customer choice in the model. The model formulation has been introduced as follow.

$$\begin{aligned}
 V_n(x_1, x_2) = & P_0^n V_{n-1}(x_1, x_2) + \sum_{i=1}^m P_i^n q_{1i} \max \{ r_{1i} + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2) \} \\
 & + \sum_{i=1}^m P_i^n q_{2i} \max \{ r_{2i} + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2) \} \\
 & + \sum_{i=1}^m P_i^n q_{3i} \max \{ r_{3i} + V_{n-1}(x_1 + 1, x_2), r_{3i} + V_{n-1}(x_1, x_2 + 1) \} \\
 & + \sum_{i=1}^m P_i^n (1 - q_{1i} - q_{2i} - q_{3i}) V_{n-1}(x_1, x_2)
 \end{aligned}$$

For this customer choice model, passengers divide themselves into three types upon arrival. The arrival probabilities are revised to P_0^n and P_i^n with $P_0^n + \sum_{i=1}^m P_i^n = 1$. Here, q_{1i} , q_{2i} and q_{3i} are the passenger choice probabilities for flight 1, flight 2 and the flexible flight respectively. An arrival may leave with the probability $(1 - q_{1i} - q_{2i} - q_{3i})$ with no choice. The revenues are r_{1i} , r_{2i} and r_{3i} for flight 1, flight 2 and the flexible product respectively.

It is clear that the structure of the optimal policy is characterized as switching curves S_1 , S_2 and S_4 which can be defined in the same way as previous ones. Switching curve S_3 has been removed since we have no rejection control for booking requests committed to buying the flexible product.

Different from many previous researches on customer choice, we do not consider the interactions between prices and the customer choice probabilities. Instead, we impose the admission controls on booking requests for flight 1 and flight 2 respectively.

4.6 Numerical Example

In the numerical examples presented in this section, we use 0, 1, 2 to indicate three decisions, namely, rejecting, accepting to Flight 1 and accepting to Flight 2, respectively. The booking horizon is divided into 15 periods with $N = 15$ to indicate the beginning period.

The probability for each booking request is taken from Example 1 of Lee and Hersh (1993) and Example 3 of Subramanian, Stidham and Lautenbacher (1999).

The capacity for the two flights are $C_1 = 10, C_2 = 10$. The fares for each flight are only classified as two classes: low and high fares. We consider four fares for the two flights with $r_{11} = 150, r_{12} = 100, r_{21} = 120, r_{22} = 80$, where r_{ij} = fare for class j of flight i .

Table 4.1: Probabilities for the Booking Requests

Probability	Period (from 15 to 1)				
	15--12	11--9	8--5	4--3	2--1
p_{11}	0.06	0.1	0.14	0.15	0.15
p_{12}	0.14	0.1	0.16	0.15	0
p_{21}	0.06	0.1	0.14	0.15	0
p_{22}	0.14	0.1	0.16	0	0
p_{31}	0.086	0.143	0.182	0.25	0
p_{32}	0.2	0.143	0.143	0.2	0
p_0	0.314	0.314	0.075	0.1	0.85

To implement the booking control in practice, management needs to use the recursive dynamic programming equation to compute offline state-dependent booking and routing control limits characterized as switching curves in the thesis. All the information of these booking limits is stored in airline seat reservation system. The seat reservation system will decide whether to accept a booking request or not upon its arrival time and fare class. We have a switching curve for each fare class of the three types of passengers in each booking period. To have a snap view of the optimal policy structure, we give the switching curve for the booking requests of the low-fare class at $n = 15$ (the starting period) and $n = 10$. They are presented in Figures 4.8, 4.9, 4.10, and 4.11.

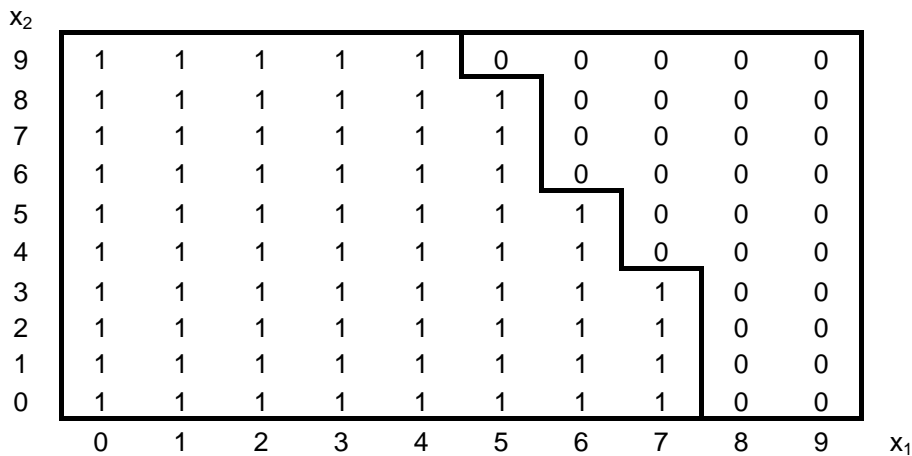


Figure 4.8: Optimal Decisions for Flight 1 Booking Requests at $n = 15$.

From the above figure, $S_1(x_2, 2)$ is the decision curve for booking request specially dedicated to flight 1. When a booking request of type 1 for low fare arrives, airline seat reservation system will check the current seat booking status on flight 1 and flight 2. It tries to compare current booking status with the booking control information for low fare class of type 1 in Figure 4.8 which is computed before the start of the whole booking process. If current booking status lies on the right side of the switching curve $S_1(x_2, 2)$, which is the area with decision 0, we reject the booking request for low-fare class seat on flight 1 at period 15. On the left side of the switching curve, which is the area with decision 1, accept the booking request for low-fare class seat of flight 1. For instance, if the seat reservation system finds out that seven seats and two seats are booked on flight 1 and flight 2 respectively, which corresponds to decision 1 in figure 4.8, it will accept this booking request. If six seats are currently booked out on both flights which corresponds to decision 0 in figure 4.8, it will reject the booking request.

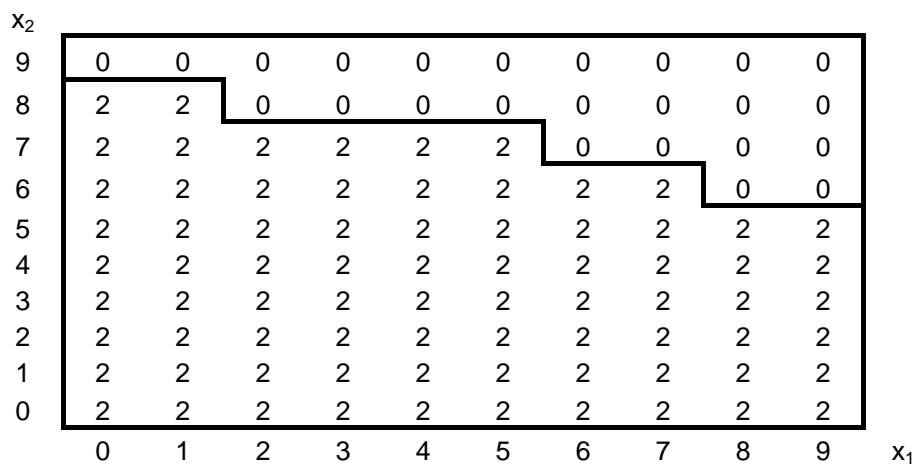


Figure 4.9: Optimal Decisions for Flight 2 Booking Requests at $n = 15$.

From the above figure, above of the switching curve $S_2(x_1)$, we reject the booking request for low-fare class seat on flight 2 at period 15. Below the switching curve, we accept the booking request.

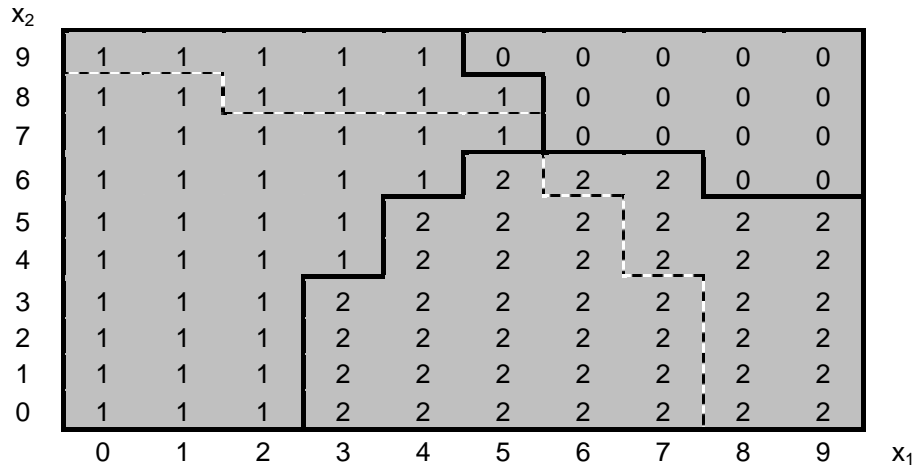


Figure 4.10: Optimal Decisions for Flexible Booking Requests at $n = 15$.

In the above figure, there are two decision curves (solid curves) $S_3(x_1, 2)$ and $S_4(x_1, 2)$ for flexible passengers for low fare at the start of the booking periods. $S_3(x_1, 2)$ is monotone decreasing and $S_4(x_1, 2)$ is monotone increasing. After a flexible booking request for low fare class arrives, airline seat reservation system will check current seat booking status on both flights. If status lies above $S_3(x_1, 2)$, where decisions are indicated by 0, i.e., we reject the booking request. Otherwise, if status lies below $S_3(x_1, 2)$, accept the booking request. If a flexible booking request has been accepted, thereafter, we use $S_4(x_1, 2)$ to determine which flight to accommodate it. If current booking status of the two flights lies on the left of the $S_4(x_1, 2)$, which is the decision zone of 1, it is better to accept the booking request to flight 1. Otherwise, accept the booking request to flight 2.

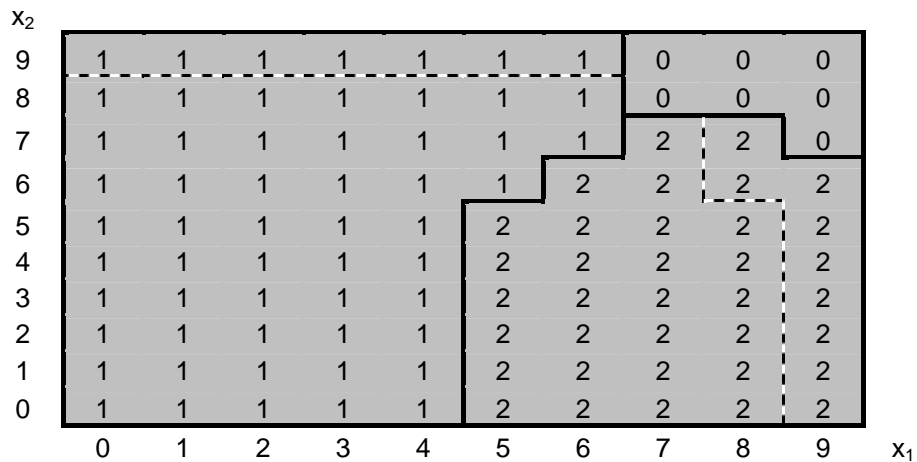


Figure 4.11: Optimal Decisions for Flexible Booking Requests at $n = 10$.

From the above figure, we can infer that the rejection regions are getting smaller with smaller n . After comparing figure 4.10 and figure 4.11, as the end of the booking period impending, it is more likely for airline to accept all fare class of booking requests.

For the case of flexible demand with discount fares, we consider four discount fares of the flexible demand for two flights with $r_{11}^d = 140$, $r_{12}^d = 90$, $r_{21}^d = 110$, $r_{22}^d = 70$ where $r_{ij}^d =$ discount fares for flexible demand for class j of flight i .

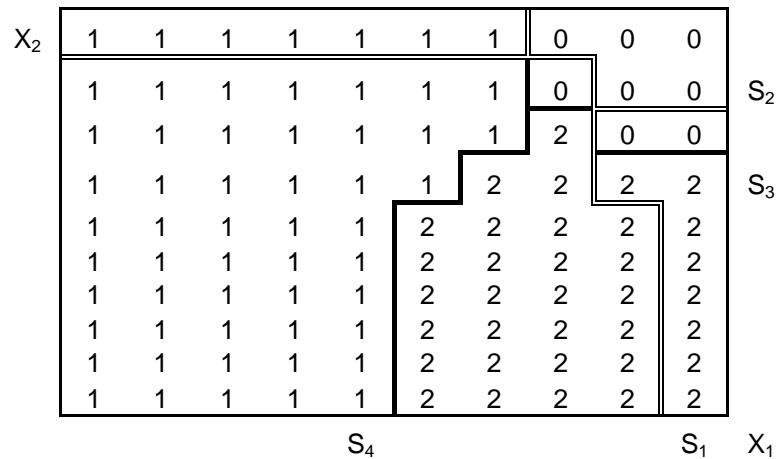


Figure 4.12: Switching Curves with Discount Fares at $n = 10$.

Switching curves S_1 and S_2 are the decision curves for flight 1 and flight 2 passengers respectively. The solid decreasing curve S_3 (with some common parts with S_1) and the increasing S_4 are for the flexible passengers. For figure 4.12, different from the case without discount fares, S_3 is not necessary the combined upper part of S_1 and lower part of S_2 . Here, it only has two separate common parts with S_1 as shown in the figure.

Following figure illustrates the monotone increasing pattern of $V(x_1, 6) - V(x_1+1, 6)$ at the 15th period, e.g. the starting period.

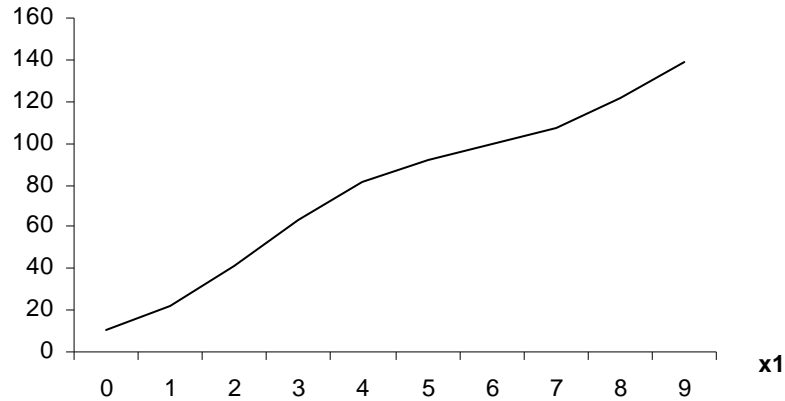


Figure 4.13: Marginal Expected Value of $V(x_1, 6) - V(x_1+1, 6)$ at period 15.

Following figure illustrates the monotone increasing pattern of $V(3, x_2) - V(3, x_2+1)$ at the 15th period, e.g. the starting period.

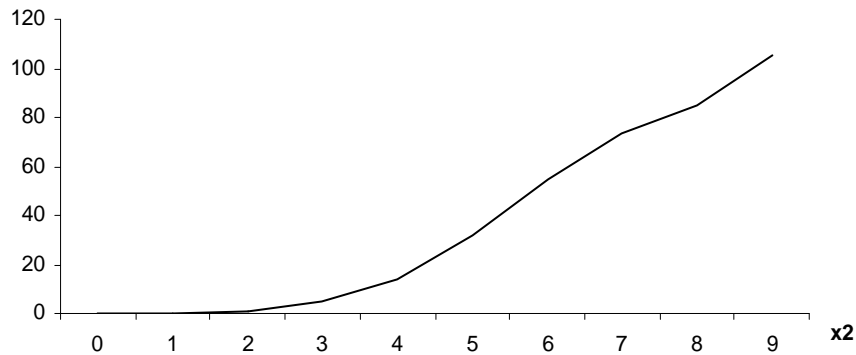


Figure 4.14: Marginal Expected value of $V(3, x_2) - V(3, x_2+1)$ at Period 15.

Following two figures illustrate, respectively, the monotone decreasing property of $D_1V_n(x_1, x_2)$ and $D_2V_n(x_1, x_2)$ with respect to period n . They imply the existence of time thresholds for admission/rejection control of type 1 and type 2 passengers. The monotone decreasing patterns with respect to argument x_1 and x_2 in two figures, respectively, indicate the submodularity of $V(x_1, x_2)$.

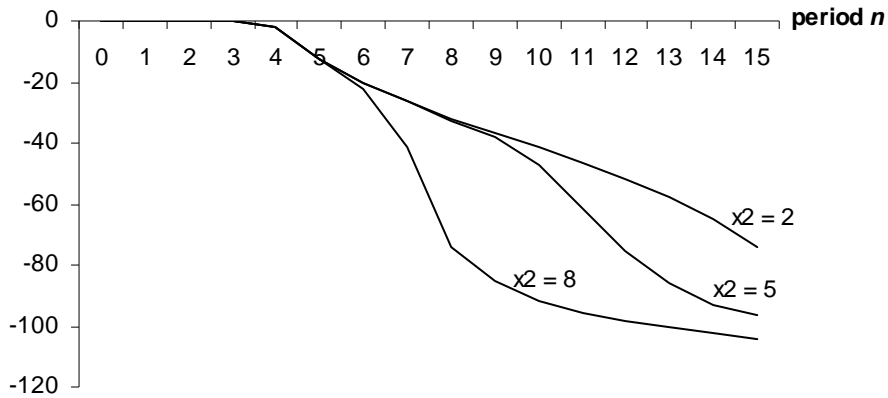


Figure 4.15: Marginal Expected Seat Revenue of $V(7, x_2) - V(6, x_2)$.

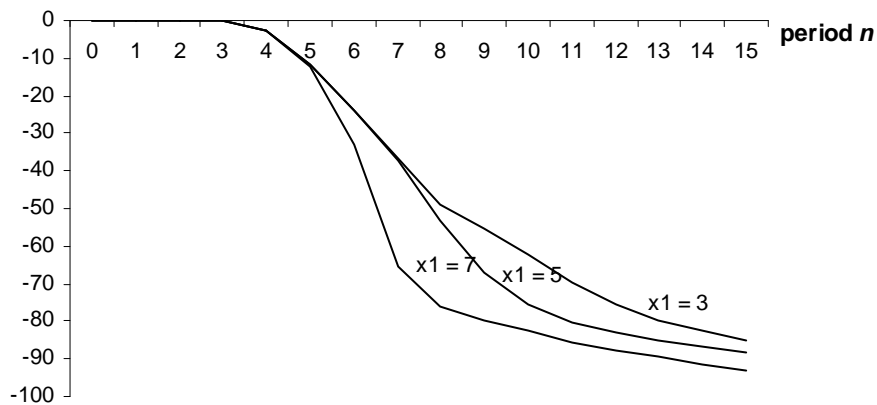


Figure 4.16: Marginal Expected Seat Revenue of $V(x_1, 9) - V(x_1, 8)$.

4.7 Conclusions

In this section, the optimal policy for the two-flight problem with flexible demand segment is thoroughly analyzed, and interesting structural results are obtained. The main result here is to prove that the structure of the optimal booking policy can be characterized by monotone switching curves.

It is clear that the model and results presented in this section can be applied to a more general situation involving two products that are substitutable with each other in use for certain passengers. For instance, we expect our model can be used in many retailing sectors. In particular,

Internet retailing opens up a new distribution channel for many products, such as airline tickets, hotel rooms and other hospitality-related products. For example, Hotwire.com and Priceline.com have designed the mechanism in the same way as discussed in this paper. Once a customer purchases the flexible product, the online reservation system will immediately assign the request to a flight or a hotel if accepted. As mentioned in Talluri and van Ryzin (2004), this mechanism is appealing to price-sensitive customers who are flexible about their choice on time and routings. For a more elaborate discussion, refer to Gallego and Phillips (2004).

Further research may focus on how to extend the results of two-flight problem to the general

4.8 Appendix:

Proof of Lemma 4.1

First, we will show that if $V \in \hat{V}$, then $T_1(V, i) \in \hat{V}$. Define

$$f(x_1, x_2) = T_1(V, i) = \max\{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\}$$

Proof of Property (a): $f(x_1, x_2) - f(x_1 + 1, x_2) \uparrow_{x_1} \uparrow_{x_2}$

Note that

$$\begin{aligned} & f(x_1, x_2) - f(x_1 + 1, x_2) \\ &= \max\{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \end{aligned} \quad (\text{A1})$$

and

$$\begin{aligned} & f(x_1 + 1, x_2) - f(x_1 + 2, x_2) \\ &= \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max\{r_i + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A2})$$

In order to show $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$, i.e., $f(x_1, x_2)$ is concave in x_1 , we have to check all cases.

Since $V \in \hat{V}$, from (A1) and (A2), we have

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 2, x_2) \\ & \leq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 3, x_2) \\ & \leq \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 3, x_2) \\ & r_i + V(x_1 + 1, x_2) - \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 2, x_2) \\ & \leq \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2) \end{aligned}$$

Combining the above two inequalities, yields

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max\{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max\{r_i + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A3})$$

Similarly, we can obtain

$$\begin{aligned} & V(x_1, x_2) - \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A4})$$

Combining (A3) and (A4), we have

$$\begin{aligned} & \max \{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$, i.e., $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_1 .

We now proceed to show that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_2 , i.e.,

$$f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1),$$

or by rearranging terms, $f(x_1, x_2) - f(x_1, x_2 + 1) \leq f(x_1 + 1, x_2) - f(x_1 + 1, x_2 + 1)$.

Note that

$$\begin{aligned} & f(x_1, x_2) - f(x_1, x_2 + 1) \\ & = \max \{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \end{aligned} \quad (\text{A5})$$

and

$$\begin{aligned} & f(x_1 + 1, x_2) - f(x_1 + 1, x_2 + 1) \\ & = \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A6})$$

By the structural properties of $V(x_1, x_2)$, we have

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 1, x_2 + 1) \\ & \leq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 2, x_2 + 1) \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 2, x_2 + 1) \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq V(x_1 + 1, x_2) - V(x_1 + 1, x_2 + 1) \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - V(x_1 + 1, x_2 + 1) \end{aligned} \quad (\text{A8})$$

Combining (A7) and (A8), yields

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A9})$$

Similarly, we can obtain that

$$\begin{aligned} & V(x_1, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A10})$$

Combining (A9) and (A10), yields

$$\begin{aligned} & \max \{r_i + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_2 .

Proof of Property (b): $f(x_1, x_2) - f(x_1, x_2 + 1) \uparrow_{x_1} \uparrow_{x_2}$

It is clear that Property (b) is symmetrical to Property (a) with respect to x_1 and x_2 , and hence our result applies.

Proof of Property (c): $f(x_1 + 1, x_2) - f(x_1, x_2 + 1) \downarrow_{x_1} \uparrow_{x_2}$

First, we prove that $f(x_1 + 1, x_2) - f(x_1, x_2 + 1)$ is increasing in x_2 , which is equivalent to requiring $f(x_1 + 1, x_2) - f(x_1, x_2 + 1) \leq f(x_1 + 1, x_2 + 1) - f(x_1, x_2 + 2)$. Note that

$$\begin{aligned} & f(x_1 + 1, x_2) - f(x_1, x_2 + 1) \\ & = \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & f(x_1 + 1, x_2 + 1) - f(x_1, x_2 + 2) \\ & = \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_i + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 2)\} \end{aligned}$$

We check different cases. First we have

$$\begin{aligned} & r_i + V(x_1 + 2, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 1, x_2 + 1) \\ & \leq r_i + V(x_1 + 2, x_2 + 1) - r_i - V(x_1 + 1, x_2 + 2) \\ & \leq \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - r_i - V(x_1 + 1, x_2 + 2) \end{aligned} \quad (\text{A11})$$

$$\begin{aligned} & r_i + V(x_1 + 2, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 1, x_2 + 1) \\ & \leq V(x_1 + 1, x_2) - V(x_1, x_2 + 1) \\ & \leq V(x_1 + 1, x_2 + 1) - V(x_1, x_2 + 2) \\ & \leq \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - V(x_1, x_2 + 2) \end{aligned} \quad (\text{A12})$$

Thus, we have

$$\begin{aligned} & r_i + V(x_1 + 2, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_i + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 2)\} \end{aligned} \quad (\text{A13})$$

By the same approach as the above, we can obtain

$$\begin{aligned} & V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_i + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 2)\} \end{aligned} \quad (\text{A14})$$

Combining (A13) and (A14), we obtain that

$$\begin{aligned} & \max \{r_i + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_i + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 2)\} \end{aligned}$$

Therefore, $f(x_1 + 1, x_2) - f(x_1, x_2 + 1)$ is increasing in x_2 . The rest half of Property (c) can be checked in a same way.

$$\text{Thus, } f(x_1, x_2) = T_1(V, i) \in \hat{V}, \text{ if } V \in \hat{V}.$$

The verification for $T_2(V, i) \in \hat{V}$ if $V \in \hat{V}$ is the same as the above proof, and is thus omitted.

We now proceed to prove that $T_3(V, i) \in \hat{V}$ if $V \in \hat{V}$. The proving procedure goes very much the same way as that for $T_1(V, i) \in \hat{V}$ and $T_2(V, i) \in \hat{V}$, but is more tedious due to the fact that T_3 is the operator with three terms under maximization. Let

$$f(x_1, x_2) = T_3(V, i) = \max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\}$$

We need to prove that $f(x_1, x_2)$ satisfies the properties (a), (b) and (c).

For Property (a), we first show that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_1 or $f(x_1, x_2)$ is concave in x_1 , i.e., $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$.

$$\begin{aligned} & f(x_1, x_2) - f(x_1 + 1, x_2) = \max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \\ & - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \end{aligned} \quad (\text{A15})$$

$$f(x_1 + 1, x_2) - f(x_1 + 2, x_2) = \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \quad (\text{A16})$$

Checking through different cases, we have

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 2, x_2) \\ & \leq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 3, x_2) \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 3, x_2) \end{aligned} \quad (\text{A17})$$

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 2, x_2) \\ & \leq R_i + V(x_1 + 1, x_2 + 1) - R_i - V(x_1 + 2, x_2 + 1) \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - R_i - V(x_1 + 2, x_2 + 1) \end{aligned} \quad (\text{A18})$$

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq r_i + V(x_1 + 1, x_2) - V(x_1 + 1, x_2) \\ & = r_i + V(x_1 + 2, x_2) - V(x_1 + 2, x_2) \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2) \end{aligned} \quad (\text{A19})$$

Combining (A17), (A18) and (A19), yields

$$\begin{aligned} & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \quad - \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A20})$$

Similarly we can obtain

$$\begin{aligned} & R_i + V(x_1, x_2 + 1) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \quad - \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A21})$$

and

$$\begin{aligned} & V(x_1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & \quad - \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A22})$$

Combining (A20), (A21) and (A22) together we have

$$\begin{aligned}
 & \max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & - \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\}
 \end{aligned}$$

To check for the second half of Property (a), we need to show that for any integers (x_1, x_2) within feasible region, $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1)$. Since

$$\begin{aligned}
 f(x_1, x_2) - f(x_1 + 1, x_2) &= \max \{r_i + V(x_1 + 1, x_2), R_i + V(x_1, x_2 + 1), V(x_1, x_2)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\}
 \end{aligned} \tag{A23}$$

$$\begin{aligned}
 & f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1) \\
 &= \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned} \tag{A24}$$

Checking through different cases, we have

$$\begin{aligned}
 & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq r_i + V(x_1 + 1, x_2) - r_i - V(x_1 + 2, x_2) \\
 & \leq r_i + V(x_1 + 1, x_2 + 1) - r_i - V(x_1 + 2, x_2 + 1) \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} - r_i - V(x_1 + 2, x_2 + 1) \\
 & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} - R_i - V(x_1 + 1, x_2 + 2) \\
 & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} - V(x_1 + 1, x_2 + 1)
 \end{aligned}$$

Combining the above three inequalities, we have

$$\begin{aligned}
 & r_i + V(x_1 + 1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned} \tag{A25}$$

Similarly, we can obtain

$$\begin{aligned}
 & R_i + V(x_1, x_2 + 1) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned} \tag{A26}$$

$$\begin{aligned}
 & V(x_1, x_2) - \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned} \tag{A27}$$

Combining (A25), (A26) and (A27) concludes the second half of Property (a). It follows that Property (b) also applies due to the symmetry between x_1 and x_2 in the properties.

Finally for Property (c), we first show that $f(x_1 + 1, x_2) - f(x_1, x_2 + 1)$ is decreasing in x_1 , that is, $f(x_1 + 1, x_2) - f(x_1, x_2 + 1) \geq f(x_1 + 2, x_2) - f(x_1 + 1, x_2 + 1)$. Note that

$$f(x_1 + 1, x_2) = \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \tag{A28}$$

$$f(x_1, x_2 + 1) = \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \tag{A29}$$

$$f(x_1 + 2, x_2) = \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \tag{A30}$$

$$f(x_1 + 1, x_2 + 1) = \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \tag{A31}$$

Checking for different cases, we have

$$\begin{aligned}
 & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 1, x_2 + 1) \\
 & \geq r_i + V(x_1 + 2, x_2) - r_i - V(x_1 + 1, x_2 + 1) \geq r_i + V(x_1 + 3, x_2) - r_i - V(x_1 + 2, x_2 + 1) \\
 & \geq r_i + V(x_1 + 3, x_2) - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned}$$

and

$$\begin{aligned}
 & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 1, x_2 + 1) \\
 & \geq R_i + V(x_1 + 2, x_2 + 1) - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned}$$

and

$$\begin{aligned}
 & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 1, x_2 + 1) \\
 & \geq V(x_1 + 2, x_2) - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned}$$

Combining the above three inequalities we have

$$\begin{aligned}
 & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - r_i - V(x_1 + 1, x_2 + 1) \\
 & \geq \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \\
 & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\}
 \end{aligned}$$

Using the same approach to the second and third term of (A29) respectively, we can obtain

$$\begin{aligned} & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - R_i - V(x_1, x_2 + 2) \\ & \geq \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \\ & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \end{aligned}$$

and

$$\begin{aligned} & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} - V(x_1, x_2 + 1) \\ & \geq \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \\ & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \end{aligned}$$

Finally combining the above three inequality leads to

$$\begin{aligned} & \max \{r_i + V(x_1 + 2, x_2), R_i + V(x_1 + 1, x_2 + 1), V(x_1 + 1, x_2)\} \\ & - \max \{r_i + V(x_1 + 1, x_2 + 1), R_i + V(x_1, x_2 + 2), V(x_1, x_2 + 1)\} \\ & \geq \max \{r_i + V(x_1 + 3, x_2), R_i + V(x_1 + 2, x_2 + 1), V(x_1 + 2, x_2)\} \\ & - \max \{r_i + V(x_1 + 2, x_2 + 1), R_i + V(x_1 + 1, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \end{aligned}$$

Thus, $f(x_1 + 1, x_2) - f(x_1, x_2 + 1)$ is decreasing in x_1 . Similarly, it can be shown that,

$f(x_1 + 1, x_2) - f(x_1, x_2 + 1)$ is increasing in x_2 , and hence Property (c) applies. Therefore,

$$T_3(V, i) \in \hat{V} \text{ if } V \in \hat{V}.$$

We have shown that properties (a), (b) and (c) are preserved under the operators T_1 , T_2 and T_3 . Since operator T is a linear and positive combination of T_1 , T_2 and T_3 , and thus $T(V, n) \in \hat{V}$ if $V \in \hat{V}$. This completes the proof. ■

5. Optimal Seat Allocation for Round-Trip/Two-Leg Problems

5.1 Introduction

In airline seat inventory management, the pool of identical seats is often divided into several fare classes. Booking requests are classified according to these fare classes. The fundamental issue of airline seat inventory control is concerned with the decision of whether accept or deny booking requests upon their arrivals at some point in time during the booking horizon. Previous research suggests that low-fare class passengers prefer booking earlier than high-fare class passengers. However, if too many seats are sold to low-fare class passengers, airline may lose the potential revenue from later high-fare class passengers. On the other hand, if many booking requests of the low-fare classes are denied, airlines may undertake the risk of having some vacant seats in the end. Thus, management tends to control the admission of various fare classes of passengers to maximize revenue.

In this chapter, firstly, we study the round-trip problem which has a go-flight and a return-flight. The management tries to maximize total revenue of these two flights. Booking requests are classified as three types. One is only for the go-flight. Another is only for the return flight. The third one is the round-trip booking requests for both of the two flights. Secondly, in case of two-leg problem, a flight will firstly fly from origin to an intermediate stop and then from the intermediate stop to the destination. The first type passengers depart from the origin and get off at the intermediate stop. The second type passengers are on board at the intermediate stop and travel to the destination. The third type passengers travel directly from origin to destination. Essentially, for model formulation, the round-trip problem and two-leg problem are same except for the booking periods. In round-trip problem, the booking period of the return-flight is longer than the go-flight. As for two-leg problem, the booking periods for the two legs are nearly the same.

From the above description, we formulate the round-trip/two-leg problem using discrete-time Markov Decision Process (MDP). Consistent with other dynamic models, we assume that the booking period can be divided into many small intervals. Within each interval, at most one

booking request arrives. We consider both the single-seat booking case and the multiple-seat booking case. In the single-seat case, each booking request requires only one seat whereas it requires multiple seats in the multiple-seat booking case. In our research, we concentrate on characterization of the structure of the optimal booking policy.

5.2 The Round-Trip Problem

Suppose that there are two flights between two cities scheduled by an airline company. Flight 1 is from city A to city B, termed as go-flight and Flight 2 is from city B to city A, termed as return-flight. It should be mentioned that the round-trip problem actually becomes single flight problem after the go-flight departure. But it does not affect our mathematical formulation and treatment. The airline tries to maximize the total revenue of these two flights. We assume that booking period horizon can be divided into subintervals so small that no more than one event happens in each of them. Booking requests can be classified into three types. The booking requests in the first two types are specifically for one of the two flights. The third type of booking requests is for both the go-flight and the return-flight. Passengers leave away and potential revenue is lost if their booking requests are denied.

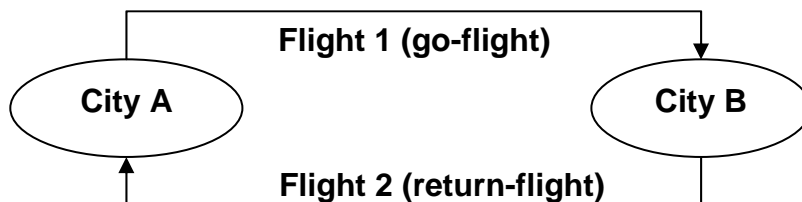


Figure 5.1: The Scenario of the Round-Trip Problem

5.2.1 Dynamic Model Formulation

n : the total decision periods for the two flights, $n = N, N-1, \dots, 0$. where N is the starting period of booking and 0 is the ending period.

m : the fare classes

r_{1i} : revenue incurred for accepting a seat booking of class i of the go-flight

r_{2i} : revenue incurred for accepting a seat booking of class i of the return-flight

r_{3i} : revenue incurred for a booking request of class i for both the go-flight and the return-flight

Without loss of generality, $r_{11} \geq r_{12} \geq \dots \geq r_{1m}$, $r_{21} \geq r_{22} \geq \dots \geq r_{2m}$, $r_{31} \geq r_{32} \geq \dots \geq r_{3m}$,

$r_{3i} \leq r_{1i} + r_{2i}$ and $r_{3i} \geq \max\{r_{1i}, r_{2i}\}$, $i = 1, 2, \dots, m$.

P_{ji}^n = probability of a booking request of type j and fare class i in period n . $j = 1, 2, 3$.

P_0^n = probability that no event happens in period n .

Then in period n , $n = N, N-1, \dots, 0$, we have

$$P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n + P_{3i}^n) = 1$$

C_i : seat capacity of each flight for $i = 1, 2$.

x_i : state variable indicating the seats already booked. Hence, $C_i - x_i$ is the remaining booking capacity on each flight

$V(x_1, x_2)$: the expected revenue function from periods n to 0 given current booking state of x_1, x_2 .

The dynamic programming equation is formulated as

$$\begin{aligned} V_n(x_1, x_2) = & P_0^n V_{n-1}(x_1, x_2) + \sum_{i=1}^m P_{1i}^n \max\{r_{1i} + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\ & + \sum_{i=1}^m P_{2i}^n \max\{r_{2i} + V_{n-1}(x_1, x_2 + 1), V_{n-1}(x_1, x_2)\} \\ & + \sum_{i=1}^m P_{3i}^n \max\{r_{3i} + V_{n-1}(x_1 + 1, x_2 + 1), V_{n-1}(x_1, x_2)\} \end{aligned} \quad (5-1)$$

$$V_0(x_1, x_2) = -R(x_1 - C_1)I_{x_1 > C_1} - R(x_2 - C_2)I_{x_2 > C_2} \quad (5-2)$$

where $R \geq \max_i \{r_{1i}, r_{2i}, r_{3i}\}$

$$I_{x_i > C_i} = \begin{cases} 1 & \text{if } x_i > C_i \\ 0 & \text{if } x_i \leq C_i \end{cases} \quad i = 1, 2 \text{ is the indicator function}$$

Remark: The form of the boundary (5-2) is a technical treatment for proof and computation. It can be regarded as a two-dimension counterpart of that in Lautenbacher and Stidham (1999).

5.2.2 Structures of the Optimal Booking Policy

Define \hat{V} to be the set of functions on Z_+^2 such that if $V \in \hat{V}$, then:

$$\text{a) } V(x_1, x_2) - V(x_1 + 1, x_2) \quad \uparrow x_1 \downarrow x_2$$

$$\text{b) } V(x_1, x_2) - V(x_1, x_2 + 1) \quad \downarrow x_1 \uparrow x_2$$

$$\text{c) } V(x_1, x_2) - V(x_1 + 1, x_2 + 1) \quad \uparrow x_1 \uparrow x_2$$

$$\text{d) } V(x_1, x_2) - V(x_1 + 1, x_2 + 1) \quad \uparrow (x_1 + x_2)$$

Here, we use \uparrow and \downarrow to denote nondecreasing and nonincreasing respectively. We know that properties a) and b) are the concavity and supermodularity of the value function $V(x_1, x_2)$. For supermodularity $V_n(x_1 + 1, x_2 + 1) - V_n(x_1 + 1, x_2) \geq V_n(x_1, x_2 + 1) - V_n(x_1, x_2)$, we could interpret the inequality as follows. If at the state (x_1, x_2) we accept a booking request of type 2, then we accept it if current state is $(x_1 + 1, x_2)$, i.e, if we accept a booking request of type 2 when x_1 and x_2 number of seats are booked on go-trip/leg 1 and return-trip/leg 2 respectively, we will accept it when $x_1 + 1$ and x_2 number of seats are booked on go-trip/leg 1 and return-trip/leg 2 respectively.

Furthermore, property c) is the so-called superconcave. Property d) indicates the concave of $V(x_1, x_2)$ in direction $(e_1 + e_2)$ where e_1 and e_2 are unit vectors. The concavity of $V(x_1, x_2)$ in e_1 , e_2 and $(e_1 + e_2)$ respectively, implies that there are thresholds for admission/rejection decisions for each type of the passengers.

Furthermore, it is trivial to show that $V(x_1, x_2)$ has following properties. For any integers $\alpha > 0$ and $\beta > 0$,

$$\text{P.a) } V(x_1, x_2) - V(x_1 + \alpha, x_2) \quad \uparrow x_1 \downarrow x_2$$

$$\text{P.b) } V(x_1, x_2) - V(x_1, x_2 + \beta) \quad \downarrow x_1 \uparrow x_2$$

$$\text{P.c) } V(x_1, x_2) - V(x_1 + \alpha, x_2 + \beta) \quad \uparrow x_1 \uparrow x_2 \uparrow (x_1 + x_2)$$

Next, we are proceeding to analyze the structures of the optimal booking policy.

Define operators $T_{1,i}$, $T_{2,i}$, $T_{3,i}$ and T^n as:

$$T_{1,i}V(x_1, x_2) = \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\}$$

$$T_{2,i}V(x_1, x_2) = \max \{r_{2i} + V(x_1, x_2 + 1), V(x_1, x_2)\}$$

$$T_{3,i}V(x_1, x_2) = \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\}$$

$$T^n V(x_1, x_2) = P_0^n V(x_1, x_2) + \sum_{i=1}^m P_{1i}^n T_{1i} V(x_1, x_2) + \sum_{i=1}^m P_{2i}^n T_{2i} V(x_1, x_2) + \sum_{i=1}^m P_{3i}^n T_{3i} V(x_1, x_2)$$

$$n = N, N-1, \dots, 1$$

Lemma 5.1 If $V \in \hat{V}$, then $T_{1,i}V, T_{2,i}V, T_{3,i}V$ and $T^n V \in \hat{V}$.

Proof: The proof is in appendix. ■

The proof for Lemma 5.1 is through case-by-case checking which is quite straightforward. It considers all cases including those overbooking cases that do not exist. We treat it as if overbooking is allowed regardless of capacity constraints on two flights. The technical treatment of the boundary (5-2) which has a high penalty on overbooking ensures that it is always optimal to reject the overbooking requests. Such treatment can simplify the proofs by removing discussion on the boundary cases of the state variable.

Based on lemma 5.1, we have the following lemma:

Lemma 5.2 The optimal value function $V_n(x_1, x_2)$ is componentwise concave and supermodular.

Proof: This lemma is proved by induction. Firstly, it can be readily checked that $V_0(x_1, x_2)$ belongs to \hat{V} according to its definition. Then suppose that it still hold for $V_{n-1}(x_1, x_2)$. Based on value iteration $V_n(x_1, x_2) = T^n V_{n-1}(x_1, x_2)$, $n = 1, 2, \dots, N$, we obtain $V_n(x_1, x_2) \in \hat{V}$. Therefore, $V_n(x_1, x_2)$ is supermodular and componentwise concave for $n = 0, 1, \dots, N$. This completes the proof. ■

Based on a), b), c) and d), we continue our discussion on the structures of the optimal policy.

Define three functions for each fare class i as:

$$S_{1n}(x_2, i) = \min \{x_1 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_{1i} \geq 0, \text{ given } x_2\}$$

$$S_{2n}(x_1, i) = \min \{x_2 \mid V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - r_{2i} \geq 0, \text{ given } x_1\}$$

$$S_{3n}(x_1, i) = \min \{x_2 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) - r_{3i} \geq 0, \text{ given } x_1\}$$

$$n = 1, 2, \dots, N$$

Remark: We will suppress i in $S_n(\cdot, \cdot)$ for the simplification of notations. The variations of constants r_{1i}, r_{2i}, r_{3i} for fare class i do not change the monotone properties of the above defined functions. Hence, for each fare class i we have such kind of three functions.

Then the following lemma illustrates the monotonicity pattern of the defined functions.

Lemma 5.3 The curve $S_{1n}(x_2)$ is nondecreasing in x_2 ; $S_{2n}(x_1)$ is nondecreasing in x_1 and $S_{3n}(x_1)$ is nonincreasing in x_1 .

Proof: First we prove the existence of $S_{1n}(x_2)$, i.e.,

$$\{x_1 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_{1i} \geq 0, \text{ given } x_2\}$$

is a nonempty set. The boundary condition implies that

$$V_0(C_1, x_2) - V_0(C_1 + 1, x_2) - r_{1i} = R - r_{1i} \geq 0. \text{ Furthermore, inductively, we can prove that}$$

$$V_n(C_1, x_2) - V_n(C_1 + 1, x_2) - r_{1i} = R - r_{1i} \geq 0 \text{ for any } n. \text{ This implies the existence of } S_{1n}(x_2).$$

Similarly, we can prove the existence of $S_{2n}(x_1)$ and $S_{3n}(x_1)$.

From the definitions, we know that for a given x_2 , $S_{1n}(x_2)$ is the least value of x_1 satisfying

$$V_n(S_{1n}(x_2), x_2) - V_n(S_{1n}(x_2) + 1, x_2) - r_{1i} \geq 0, \text{ and}$$

$$S_{1n}(x_2 + 1) \text{ is the least value satisfying } V_n(S_{1n}(x_2 + 1), x_2 + 1) - V_n(S_{1n}(x_2 + 1) + 1, x_2 + 1) - r_{1i} \geq 0$$

From the supermodularity of $V_n(x_1, x_2)$, we have

$$\begin{aligned} & V_n(S_{1n}(x_2 + 1), x_2) - V_n(S_{1n}(x_2 + 1) + 1, x_2) - r_{1i} \\ & \geq V_n(S_{1n}(x_2 + 1), x_2 + 1) - V_n(S_{1n}(x_2 + 1) + 1, x_2 + 1) - r_{1i} \geq 0 \end{aligned}$$

Since $S_{1n}(x_2)$ is the least value of x_1 satisfying

$$V_n(S_{1n}(x_2), x_2) - V_n(S_{1n}(x_2) + 1, x_2) - r_{1i} \geq 0, \text{ we have } S_{1n}(x_2 + 1) \geq S_{1n}(x_2).$$

Therefore, $S_{1n}(x_2)$ is nondecreasing in x_2 .

Similarly, we can prove that $S_{2n}(x_1)$ is nondecreasing in x_1 .

For $S_{3n}(x_1)$, we have $V_n(x_1, S_{3n}(x_1)) - V_n(x_1 + 1, S_{3n}(x_1) + 1) - r_{3i} \geq 0$

Then from property c) we obtain $V_n(x_1 + 1, S_{3n}(x_1)) - V_n(x_1 + 2, S_{3n}(x_1) + 1) - r_{3i} \geq 0$.

isfying

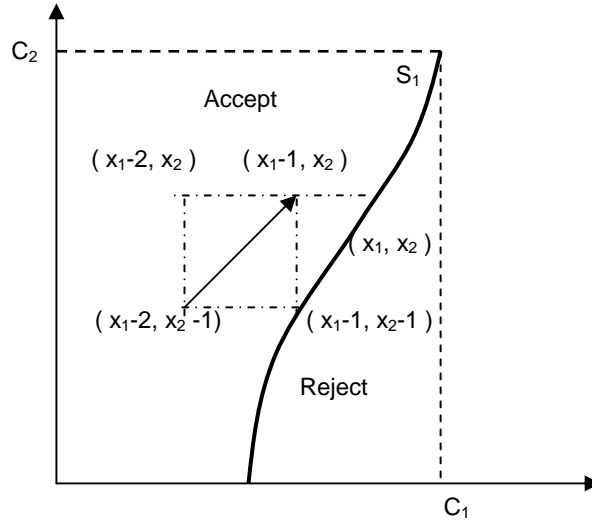


Figure 5.2: The ‘Slope’ of the Switching Curve $S_{1n}(x_2)$

From the definition of $S_{1n}(x_2)$, we can infer that $(x_1 - 2, x_2 - 1)$ is strictly on the left of $S_{1n}(x_2)$ and $(x_1 - 1, x_2 - 1)$ is on the left or on the $S_{1n}(x_2)$. This means that the ‘slope’ of switching curve $S_{1n}(x_2)$ is greater than or equal to 1. Similarly, we can prove the result of $S_{2n}(x_1)$. ■

From the ‘slope’ of $S_{1n}(x_2)$ and $S_{2n}(x_1)$, we know that once the two curves may intersect at some point, then they go away and never meet again, e.g., there is at most one intersection point for S_1 and S_2 .

Theorem 5.1 The structure of the optimal booking policy is characterized by three state-dependent monotone switching curves.

Proof: On the right side of $S_{1n}(x_2)$, for a fixed x_2 , from $x_1 \geq S_{1n}(x_2)$ we have

$$V_n(x_1, x_2) - V_n(x_1 + 1, x_2) - r_i \geq V_n(S_{1n}(x_2), x_2) - V_n(S_{1n}(x_2) + 1, x_2) - r_i \geq 0.$$

This inequality reads that it is optimal to reject the booking request. Otherwise, for a fixed x_2 , accept the booking request when x_1 is on the left side of the curve $S_{1n}(x_1)$. Based on lemma 5.3, we know that $S_{1n}(x_2)$ is monotone and nondecreasing in x_2 . Hence, such monotone switching curve divides the state space into two regions which are the rejection area and the acceptance area respectively.

For $S_{2n}(x_1)$, above the curve, for a fixed x_1 , from $x_2 \geq S_{2n}(x_1)$ we have

$$V_n(x_1, x_2) - V_n(x_1, x_2 + 1) - r_{2i} \geq V_n(x_1, S_{2n}(x_1)) - V_n(x_1, S_{2n}(x_1) + 1) - r_{2i} \geq 0.$$

Therefore, it is optimal to reject a type 2 booking request when current state is above $S_{2n}(x_1)$.

Otherwise accept it when x_2 is below $S_{2n}(x_1)$. For each fixed x_1 , $S_{2n}(x_1)$ can be regarded as the critical points for the acceptance/denial decision. This monotone switching curve divides the state space into two decision regions.

As for $S_{3n}(x_1)$, for fixed x_1 , when $x_2 \geq S_{3n}(x_1)$, we have

$$V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) - r_{3i} \geq V_n(x_1, S_{3n}(x_1)) - V_n(x_1 + 1, S_{3n}(x_1) + 1) - r_{3i} \geq 0.$$

Therefore it is optimal to reject the booking request. Below $S_{3n}(x_1)$, it is optimal to accept the booking request for both go-flight and return-flight. ■

From definitions, we know that for each booking period and each fare class, we have such kind of three switching curves.

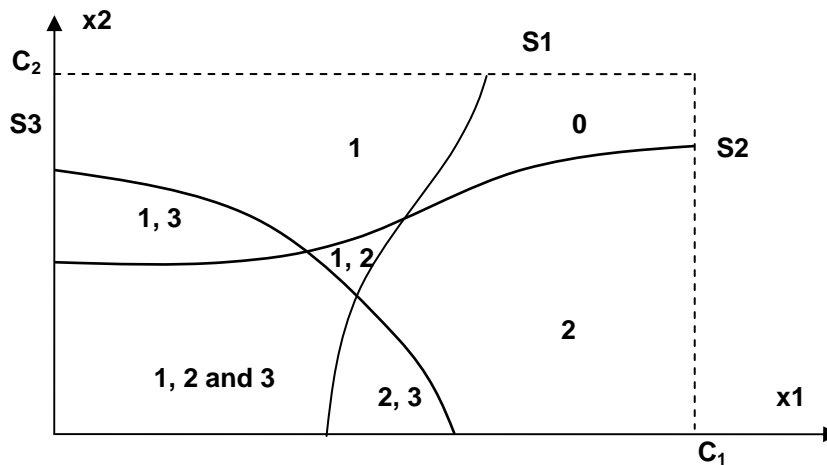


Figure 5.3: The Switching Curves for Three Types of Booking Requests

Figure 5.3 is composed of three graphs which are for type 1, type 2 and type 3 booking requests, respectively. In each graph, there is a switching curve. Here, we combine these three graphs together in a single graph for convenience since there is at most one booking request during each booking period according to our model assumption. Thus, three switching curves divide the feasible region into several decision areas in which decisions are denoted by 0, 1, 2 and 3

respectively. Here, 0, 1, 2 and 3 denote reject, accept to go-flight, accept to return-flight and accept to both flights, respectively. For instance, in the area with decisions 1, 2 and 3, it means that when current state of two flights lies in this region, it is optimal to accept all types of booking requests.

5.2.3 Further Properties of the Optima Value Function

In addition to the structure properties derived in the preceding section, the optimal value function has some other monotone properties.

Proposition 5.1 The optimal value function $V_n(x_1, x_2)$ is nonincreasing in x_1 and x_2 .

Proof: The proof is by induction. It is evident that $V_0(x_1, x_2)$ is nonincreasing in x_1 and x_2 , we assume that $V_{n-1}(x_1, x_2)$ is nonincreasing in x_1 and x_2 . We have

$$\begin{aligned} TV_{1,i}(x_1 + 1, x_2) - TV_{1,i}(x_1, x_2) &= \max \{r_{1i} + V_{n-1}(x_1 + 2, x_2), V_{n-1}(x_1 + 1, x_2)\} \\ &\quad - \max \{r_{1i} + V_{n-1}(x_1 + 1, x_2), V_{n-1}(x_1, x_2)\} \\ &\leq 0 \end{aligned}$$

Hence, $TV_{1,i}$ is nonincreasing in x_1 . Similarly, we can prove that it is nonincreasing in x_2 . By the same argument, we can prove that $TV_{2,i}$ and $TV_{3,i}$ are all nonincreasing in x_1 and x_2 . Finally, from (5-1), we see that the nonnegative linear combination of these nonincreasing functions is still a nonincreasing function. Therefore, $V_n(x_1, x_2)$ is nonincreasing in x_1 and x_2 . ■

Define following operators:

$$D_1V(x_1, x_2) = V(x_1 + 1, x_2) - V(x_1, x_2),$$

$$D_2V(x_1, x_2) = V(x_1, x_2 + 1) - V(x_1, x_2),$$

$$DV(x_1, x_2) = V(x_1 + 1, x_2 + 1) - V(x_1, x_2),$$

Then we have the following proposition.

Proposition 5.2 Optimal value function $V_n(x_1, x_2)$ is nondecreasing in period n for given x_1 and x_2 .

Proof: From (5-1) and $P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n + P_{3i}^n) = 1$, we have

$$\begin{aligned}
 V_n(x_1, x_2) - V_{n-1}(x_1, x_2) &= V_n(x_1, x_2) - V_{n-1}(x_1, x_2) \{P_0^n + \sum_{i=1}^m (P_{1i}^n + P_{2i}^n + P_{3i}^n)\} \\
 &= P_0^n \{V_{n-1}(x_1, x_2) - V_{n-1}(x_1, x_2)\} \\
 &+ \sum_{i=1}^m P_{1i}^n \max\{r_i + D_1 V_{n-1}(x_1, x_2), 0\} + \sum_{i=1}^m P_{2i}^n \max\{r_{2i} + D_2 V_{n-1}(x_1, x_2), 0\} \\
 &+ \sum_{i=1}^m P_{3i}^n \max\{r_{3i} + D V_{n-1}(x_1, x_2), 0\}
 \end{aligned} \tag{5-3}$$

It is evident that each of four terms on the RHS of (5-3) is no less than 0. Hence,

$V_n(x_1, x_2) - V_{n-1}(x_1, x_2) \geq 0$ and $V_n(x_1, x_2)$ is nondecreasing in period n . ■

Proposition 5.3 For given x_1 and x_2 , $DV_n(x_1, x_2)$ is nonincreasing in period n .

Proof: We need to prove that $DV_n(x_1, x_2) \leq DV_{n-1}(x_1, x_2)$.

Based on (5-1) and (5-3), we have

$$\begin{aligned}
 &DV_n(x_1, x_2) - DV_{n-1}(x_1, x_2) \\
 &= V_n(x_1 + 1, x_2 + 1) - V_{n-1}(x_1 + 1, x_2 + 1) - \{V_n(x_1, x_2) - V_{n-1}(x_1, x_2)\} \\
 &= \sum_{i=1}^m P_{1i}^n \{\max\{r_i + D_1 V_{n-1}(x_1 + 1, x_2 + 1), 0\} \\
 &\quad - \max\{r_i + D_1 V_{n-1}(x_1, x_2), 0\}\} \\
 &+ \sum_{i=1}^m P_{2i}^n \{\max\{r_{2i} + D_2 V_{n-1}(x_1 + 1, x_2 + 1), 0\} \\
 &\quad - \max\{r_{2i} + D_2 V_{n-1}(x_1, x_2), 0\}\} \\
 &+ \sum_{i=1}^m P_{3i}^n \{\max\{r_{3i} + D V_{n-1}(x_1 + 1, x_2 + 1), 0\} - \max\{r_{3i} + D V_{n-1}(x_1, x_2), 0\}\}
 \end{aligned} \tag{5-4}$$

Based on properties c) and d), it is not difficult to show that each of the three terms on the RHS of

(5-4) is no more than 0. Hence, $DV_n(x_1, x_2)$ is nonincreasing in period n . ■

Proposition 5.3 tells that the optimal value function has modularity in $(x_1 + x_2, n)$. Therefore, there

are time thresholds for type 3 booking requests. We can define critical period

$$n^*(x_1, x_2, i) = \min\{n \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) - r_{3i} \geq 0\}$$

For $n \geq n^*$, at current state (x_1, x_2) , it is optimal to reject the type 3 booking request of fare class i

since $V_n(x_1, x_2) \geq r_{3i} + V_n(x_1 + 1, x_2 + 1)$; otherwise, when $n < n^*$, it is optimal to accept type 3

booking request. It tends to accept a booking request when the booking period approaches the end, i.e., n approaches 0.

Unfortunately, due to the loss of monotone properties of $D_1V_n(x_1, x_2)$ and $D_2V_n(x_1, x_2)$, we cannot define the critical period $n_1^*(x_1, x_2, i)$ and $n_2^*(x_1, x_2, i)$ for fare class i of type 1 and type 2 booking requests respectively. The reason is given as follow.

For the monotonicity of $D_1V_n(x_1, x_2)$, we need to prove that $D_1V_n(x_1, x_2) \leq D_1V_{n-1}(x_1, x_2)$. Based on (5-1), we have

$$\begin{aligned}
 & D_1V_n(x_1, x_2) - D_1V_{n-1}(x_1, x_2) \\
 &= V_n(x_1 + 1, x_2) - V_n(x_1, x_2) - \{V_{n-1}(x_1 + 1, x_2) - V_{n-1}(x_1, x_2)\} \\
 &= \sum_{i=1}^m P_{1i}^n \{ \max \{ r_{1i} + V_{n-1}(x_1 + 2, x_2) - V_{n-1}(x_1 + 1, x_2), 0 \} - \max \{ r_{1i} + V_{n-1}(x_1 + 1, x_2) - V_{n-1}(x_1, x_2), 0 \} \} \\
 &+ \sum_{i=1}^m P_{2i}^n \{ \max \{ r_{2i} + V_{n-1}(x_1 + 1, x_2 + 1) - V_{n-1}(x_1 + 1, x_2), 0 \} - \max \{ r_{2i} + V_{n-1}(x_1, x_2 + 1) - V_{n-1}(x_1, x_2), 0 \} \} \\
 &+ \sum_{i=1}^m P_{3i}^n \{ \max \{ r_{3i} + V_{n-1}(x_1 + 2, x_2 + 1) - V_{n-1}(x_1 + 1, x_2), 0 \} - \max \{ r_{3i} + V_{n-1}(x_1 + 1, x_2 + 1) - V_{n-1}(x_1, x_2), 0 \} \}
 \end{aligned}$$

Noted that the first term and the third term of the RHS of the last equality is less than or equal to zero. Therefore, we need to show that the second term of the RHS is also less than or equal to zero.

Unfortunately, we cannot ensure that it is less than or equal to zero since $V_{n-1}(x_1 + 1, x_2 + 1) - V_{n-1}(x_1 + 1, x_2) \geq V_{n-1}(x_1, x_2 + 1) - V_{n-1}(x_1, x_2)$, which is supermodularity of value function based on our structural property. Hence, we can not ensure that $D_1V_n(x_1, x_2)$ is nonincreasing in period n . Similar case applies to $D_2V_n(x_1, x_2)$.

Intuitively, when in period n , P_{2i}^n for type 2 are relatively higher than P_{1i}^n and P_{3i}^n , it tends to accept a booking request of type 2. After the acceptance a type 2, it tends to reject a type 3 at current period to reserve seat capacity for future type 3 booking request. Consequently, it might reject a fare class of type 1 that could be accepted at earlier period due to the lack of competitive advantage over future type 3 booking requests.

Therefore, for the round-trip/two-leg problem, due to the interactions of three types of booking requests, a fare class of type 1 and type 2 that can be accepted at an earlier period but might be rejected at a later period. This is different from the single-flight or single-leg problems which have the monotonicity of $DV(x)$ that ensures a fare class accepted at earlier period will be accepted at the later periods.

We use a numerical study to confirm the above conclusion. Please refer to the numerical example section for the results.

Proposition 5.4 The switching curves $S_{1n}(x_2, i)$, $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$ are monotonically nonincreasing in fare i .

Proof: First for the switching curve $S_{1n}(x_2, i)$, we need to prove that $S_{1n}(x_2, i) \geq S_{1n}(x_2, i+1)$.

Here, we use counter proof. From definition and assumption, we have

$$V_n(S_{1n}(x_2, i), x_2) - V_n(S_{1n}(x_2, i) + 1, x_2) - r_{1i} \geq 0$$

$$V_n(S_{1n}(x_2, i+1), x_2) - V_n(S_{1n}(x_2, i+1) + 1, x_2) - r_{1,i+1} \geq 0.$$

If $S_{1n}(x_2, i) < S_{1n}(x_2, i+1)$, then $V_n(S_{1n}(x_2, i), x_2) - V_n(S_{1n}(x_2, i) + 1, x_2) - r_{1,i+1} < 0$ and

$$V_n(S_{1n}(x_2, i), x_2) - V_n(S_{1n}(x_2, i) + 1, x_2) - r_{1i} \geq 0 > V_n(S_{1n}(x_2, i), x_2) - V_n(S_{1n}(x_2, i) + 1, x_2) - r_{1,i+1}$$

From the above inequality, we can infer that $r_{1i} < r_{1,i+1}$, which is conflict with $r_{1i} \geq r_{1,i+1}$. Hence,

$S_{1n}(x_2, i) \geq S_{1n}(x_2, i+1)$ holds.

The same results can be easily derived for $S_{2n}(x_1, i)$ and $S_{3n}(x_1, i)$. Then the three switching curves are all monotonically nonincreasing in fare i . ■

Proposition 5.4 is intuitive. It is more likely to reject a booking request of lower fares classes. The higher the fare class, the less of the chance of being rejected. Graphically, rejection areas of the high fare classes are monotonically smaller than those of lower fare classes.

Proposition 5.5 The switching curve $S_{3n}(x_1)$ is monotonic in period n .

Proof: For $S_{3n}(x_1)$ and $S_{3,n-1}(x_1)$, we have

$$S_{3n}(x_1) = \min\{x_2 \mid V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) - r_{3i} \geq 0, \text{ given } x_1\} \text{ and}$$

$$S_{3,n-1}(x_1) = \min\{x_2 \mid V_{n-1}(x_1, x_2) - V_{n-1}(x_1 + 1, x_2 + 1) - r_{3i} \geq 0, \text{ given } x_1\}$$

By proposition 4.3, we have $V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) \geq V_{n-1}(x_1, x_2) - V_{n-1}(x_1 + 1, x_2 + 1)$. Thus,

$$\begin{aligned} & V_n(x_1, S_{3,n-1}(x_1)) - V_n(x_1 + 1, S_{3,n-1}(x_1) + 1) - r_{3i} \\ & \geq V_{n-1}(x_1, S_{3,n-1}(x_1)) - V_{n-1}(x_1 + 1, S_{3,n-1}(x_1) + 1) - r_{3i} \geq 0 \end{aligned}$$

Since $S_{3n}(x_1)$ is the least value satisfying $V_n(x_1, x_2) - V_n(x_1 + 1, x_2 + 1) - r_{3i} \geq 0$, for any fixed x_1 , we have $S_{3n}(x_1) \leq S_{3,n-1}(x_1)$. Thus $S_{3n}(x_1)$ is monotonic in period n . ■

Proposition 5.5 tells us that it is less likely to reject type 3 booking requests as time approaching the end of the booking horizon.

5.3 The Two-Leg Problem

For the two-leg problem, if formulated as the discrete-time Markov decision process, the model is exactly the same as the round-trip problem that we have already studied. The structure of the optimal booking policy for two-leg problem is also characterized by monotone switching curves. Other monotone properties of the optimal value function are also preserved. The difference is the booking periods that we have mentioned in the introduction section.

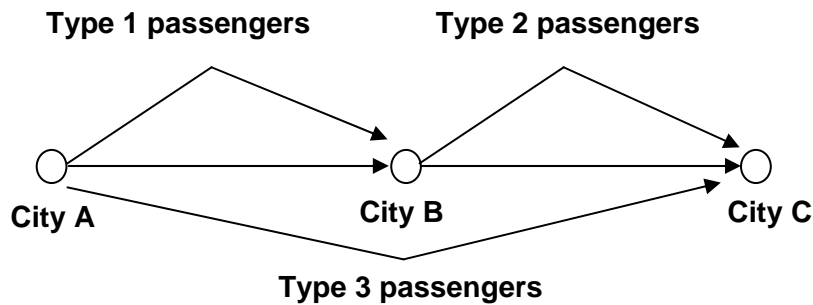


Figure 5.4: The Scenario of the Two-Leg Problem

5.4 Deterministic Linear Programming to Solve the Two-Leg Problem

For any realistic size of airline revenue management problem, it is hard to implement the computation because of the large dimensionality of the state space. Various approximation methods have been proposed to make the computation feasible. Among the widely used methods,

deterministic linear programming approach proved to be a simple and efficient one. For the round-trip/two-leg problem, let $\mu_{ij} = E[D_{ij}]$, $i = 1, 2, \dots, m$ $j = 1, 2, 3$ be the mean of the total demand of fare class i and type j passengers. The decision variables y_{ij} , $i = 1, 2, \dots, m$ $j = 1, 2, 3$ denote the partitioned allocation of capacity for each fare class and type of passengers. The deterministic linear programming (DLP) method is used to solve static seat capacity allocation by treating demands as deterministic ones with μ_{ij} , $i = 1, 2, \dots, m$ $j = 1, 2, 3$.

$$\max_y \sum_{j=1}^3 \sum_{i=1}^m r_{ji} y_{ij} \quad (5-5)$$

$$s.t. \quad \sum_{i=1}^m y_{i1} + \sum_{i=1}^m y_{i3} \leq C_1 \quad (5-6)$$

$$\sum_{i=1}^m y_{i2} + \sum_{i=1}^m y_{i3} \leq C_2 \quad (5-7)$$

$$0 \leq y_{ij} \leq \mu_{ij} \quad i = 1, 2, \dots, m \quad j = 1, 2, 3 \quad (5-8)$$

Constraint (5-6) denotes that the total capacities for type 1 and type 3 passengers are no more than C_1 which is the capacity of go-trip or the first leg. Constraint (5-7) has the same meaning as (5-6). Last constraints (5-8) are the total demand constraints indicating that the allocated capacity to each fare class and type of passenger is no more than the mean of its total demand. Here, we use the optimal prime solution y_{ij}^* for the DLP to construct a partitioned control. We regard $\sum_{i=1}^m y_{ij}$, $j = 1, 2, 3$ as the capacities for type 1, type 2 and type 3 passengers respectively. Thereafter, for each type of passengers, the one-dimensional dynamic model could be used for seat inventory control. The DLP method simplifies the analysis of the original complex problem and greatly reduces the computation workload for practical issues. Furthermore, it is convenient to formulate the multi-leg problem in DLP and solve it in the same way.

5.5 Extension to the Multiple-Seat Booking Problem

For single-leg multi-class multiple-seat booking problem, Brumelle and Walczak (2003) show the optimality of booking limit policy if partial acceptance of the multiple-seat booking request is

allowed. Such partial acceptance is classified into the cases “that is, given a demand for $k > 1$ units, we can sell any quantity u in the range $1 \leq u \leq k$, and the customer is willing to buy any amount in this range”— Talluri and van Ryzin (2004). Here, in this section, we investigate the extension to the two-leg multi-class multiple-seat booking problem. We allow that multiple demands can be partially satisfied. Thus, the structure of the optimal booking policy has been characterized as monotone switching curves. According to our knowledge, such issue has never been addressed before.

It is often the case that a passenger may demand multiple seats. It is natural to make an assumption that the number of seats of each booking request is a finite discrete random variable. The random variables of multiple seats for three types of booking requests are $q_1(d_1)$, $q_2(d_2)$ and $q_3(d_3)$ respectively, where

$$\sum_{d_1=1}^{Q_1} q_1(d_1) = 1, \quad \sum_{d_2=1}^{Q_2} q_2(d_2) = 1 \quad \text{and} \quad \sum_{d_3=1}^{Q_3} q_3(d_3) = 1$$

Here, for simplicity of notations, we assume that distributions for multiple seats are identical for all fare classes and identical throughout the whole booking horizon. To be noted, practically, we can make it different the probability distribution for each fare class and in each booking period. Such a refinement will not affect our following analysis.

Define operators $T_{1,i}$, $T_{2,i}$, $T_{3,i}$ and T^n as:

$$T_{1,i}V(x) = \max_{0 \leq j_1 \leq \{d_1, C\}} \{j_1 r_{1i} + V(x + j_1 e_1)\}$$

$$T_{2,i}V(x) = \max_{0 \leq j_2 \leq \{d_2, C\}} \{j_2 r_{2i} + V(x + j_2 e_2)\}$$

$$T_{3,i}V(x) = \max_{0 \leq j_3 \leq \{d_3, C\}} \{j_3 r_{3i} + V(x + j_3 e_1 + j_3 e_2)\}$$

$$T^n V(x) = P_0^n V(x) + \sum_{i=1}^m \sum_{d_1=1}^{Q_1} P_{1i}^n q_1(d_1) T_{1i} V(x) + \sum_{i=1}^m \sum_{d_2=1}^{Q_2} P_{2i}^n q_2(d_2) T_{2i} V(x) + \sum_{i=1}^m \sum_{d_3=1}^{Q_3} P_{3i}^n q_3(d_3) T_{3i} V(x)$$

Where, $n = N, N-1, \dots, 1$. $T_{1,i}V(x)$, $T_{2,i}V(x)$ and $T_{3,i}V(x)$ are the acceptance/rejection operators for type 1, type 2 and type 3 booking requests, respectively. Then the dynamic programming equation is

$$V_n(x) = T^n V_{n-1}(x)$$

Lemma 5.5 If $V \in \hat{V}$, then $T_{1,i}V, T_{2,i}V, T_{3,i}V$ and $T^n V \in \hat{V}$.

Proof: The proof is in appendix. ■

Based on a), b), c) and d), we continue our discussion on the structures of the optimal policy.

Define following functions for each fare class i and each number of demand d_j as:

$$S_{1n}(x_2, i, d_{1j}) = \min\{x_1 \mid V_n(x_1 + d_{1j} - 1, x_2) - V_n(x_1 + d_{1j}, x_2) - r_{1i} \geq 0, \text{ given } x_2\}$$

$$S_{2n}(x_1, i, d_{2j}) = \min\{x_2 \mid V_n(x_1, x_2 + d_{2j} - 1) - V_n(x_1, x_2 + d_{2j}) - r_{2i} \geq 0, \text{ given } x_1\}$$

$$S_{3n}(x_1, i, d_{3j}) = \min\{x_2 \mid V_n(x_1 + d_{3j} - 1, x_2 + d_{3j} - 1) - V_n(x_1 + d_{3j}, x_2 + d_{3j}) - r_{3i} \geq 0, \text{ given } x_1\}$$

$n = 1, 2, \dots, N, i = 1, 2, \dots, m, d_{1j} = 1, 2, \dots, Q_1, d_{2j} = 1, 2, \dots, Q_2$ and $d_{3j} = 1, 2, \dots, Q_3$.

Lemma 5.6 Switching curve $S_{1n}(x_2, i, d_{1j})$ is nondecreasing in x_2 for fixed d_{1j} ; is nonincreasing in d_{1j} for fixed x_2 . $S_{2n}(x_1, i, d_{2j})$ is nondecreasing in x_1 for fixed d_{2j} ; is nonincreasing in d_{2j} for fixed x_1 . $S_{3n}(x_1, i, d_{3j})$ is nonincreasing in x_1 for fixed d_{3j} ; is nonincreasing in d_{3j} for fixed x_1 .

Proof: The proof for the monotonicity of switching curves with respect to its two arguments can be derived in the same way as those of Lemma 5.3. It is also straightforward to demonstrate that switching curves are monotone nonincreasing with respect to the number of demand. For instance, for $S_{1n}(x_2, i, d_{1j})$, from its definition, we have

$$V_n(S_{1n}(x_2, i, d_{1j}) + d_{1j} - 1, x_2) - V_n(S_{1n}(x_2, i, d_{1j}) + d_{1j}, x_2) - r_{1i} \geq 0 \text{ and}$$

$$V_n(S_{1n}(x_2, i, d_{1j} + 1) + d_{1j} - 1, x_2) - V_n(S_{1n}(x_2, i, d_{1j} + 1) + d_{1j}, x_2) - r_{1i} < 0$$

Hence, our result $S_{1n}(x_2, i, d_{1j}) \geq S_{1n}(x_2, i, d_{1j} + 1)$ holds. ■

Theorem 5.2 The optimal booking policy is characterized by state-dependent monotone switching curves.

Proof: It is quite straightforward to obtain the result based on lemma 5.5 and lemma 5.6. ■

5.6 Extension to the Multi-Leg Problem

Suppose that there are $l + 1$ number of airport nodes and m number of fare classes. We formulate the multi-leg problem as a discrete-time Markov decision process. The dynamic programming equation is written as follow:

$$V_n(x) = P_0^n V_{n-1}(x) + \sum_{k=0}^{l-1} \sum_{j=k+1}^l \sum_{i=1}^m P_{i,j,k}^n \max \{ r_{i,k,j} + V_{n-1}(x + e_{k,j}), V_{n-1}(x) \} \quad (5-9)$$

The arrival probabilities should satisfy

$$P_0^n + \sum_{k=0}^{l-1} \sum_{j=k+1}^l \sum_{i=1}^m P_{i,j,k}^n = 1$$

where $e_{k,j}$ indicates the leg from node k to node j . It is a vector with 1 from $(k+1)^{\text{th}}$ argument to the j^{th} argument and the other arguments are 0. $r_{i,k,j}$ denotes the revenue from i fare class of the leg from node k to node j .

$$V_0(x) = - \sum_{i=1}^l R(x_i - C) I_{x_i > c} \quad (5-10)$$

where $R \geq \max_{i,k,j} \{ r_{i,k,j} \}$, $i = 1, \dots, m$, $k = 0, \dots, l-1$, $j = k+1, \dots, l$ and I is indicator function.

For ease of notation, we consider the structural properties and the optimal booking policy for the three-leg problem. Thus, the optimality equation (5-9) is reduced to

$$\begin{aligned} V_n(x) = & P_0^n V_{n-1}(x) + \sum_{j=1}^3 \sum_{i=1}^m P_{ji}^n \max \{ r_{ij} + V_{n-1}(x + e_j), V_{n-1}(x) \} \\ & + \sum_{i=1}^m P_{4i}^n \max \{ r_{4i} + V_{n-1}(x + e_1 + e_2), V_{n-1}(x) \} \\ & + \sum_{i=1}^m P_{5i}^n \max \{ r_{5i} + V_{n-1}(x + e_2 + e_3), V_{n-1}(x) \} \\ & + \sum_{i=1}^m P_{6i}^n \max \{ r_{6i} + V_{n-1}(x + e_1 + e_2 + e_3), V_{n-1}(x) \} \end{aligned} \quad (5-11)$$

$$V_0(x) = -\sum_{i=1}^3 R(x_i - C)I_{x_i > c} \quad (5-12)$$

where $R \geq \max_i \{r_{1i}, r_{2i}, r_{3i}, r_{4i}, r_{5i}, r_{6i}\}$ and

$$P_0^n + \sum_{j=1}^3 \sum_{i=1}^m P_{j,i}^n + \sum_{i=1}^m P_{4i}^n + \sum_{i=1}^m P_{5i}^n + \sum_{i=1}^m P_{6i}^n = 1$$

First, motivated by the structural result in Benjaafar and ELHafsi (2006), let us introduce a class of real-valued functions defined on the lattice Z_+^k , denoted by \mathcal{V} . We say $V \in \mathcal{V}$ if it possesses the following properties:

- 1) $V(x) - V(x + e_{i_1}) \uparrow e_{i_1} \downarrow e_{i_2}$ for any $i_2 \neq i_1$
- 2) $V(x) - V(x + e_{i_1} + e_{i_2}) \uparrow e_{i_1} \uparrow e_{i_2} \uparrow (e_{i_1} + e_{i_2}) \downarrow e_{i_3}$ for any $i_1 \neq i_2, i_3 \neq i_1$ and $i_3 \neq i_2$
- ⋮
- K) $V(x) - V(x + e_{i_1} + e_{i_2} + \dots + e_{i_k}) \uparrow e_{i_1} \uparrow (e_{i_2} + e_{i_3}) \uparrow (e_{i_4} + e_{i_5} + e_{i_6}) \uparrow \dots \uparrow (e_1 + e_2 + \dots + e_k)$

Here, $i_1, i_2, i_3, \dots, i_k$ are any different numbers selected from 1 to k respectively. Compared with the structural properties A1, A2 and A3 in the proof of theorem 1 in Benjaafar and ELHafsi (2006), ours are more general. For the three-leg problem, we say $V \in \mathcal{V}$ if it possesses the following properties:

- 1) $V(x) - V(x + e_1) \uparrow e_1 \downarrow e_2 \downarrow e_3$
- 2) $V(x) - V(x + e_2) \uparrow e_2 \downarrow e_1 \downarrow e_3$
- 3) $V(x) - V(x + e_3) \uparrow e_3 \downarrow e_1 \downarrow e_2$
- 4) $V(x) - V(x + e_1 + e_2) \uparrow e_1 \uparrow e_2 \uparrow (e_1 + e_2) \downarrow e_3$
- 5) $V(x) - V(x + e_1 + e_3) \uparrow e_1 \uparrow e_3 \uparrow (e_1 + e_3) \downarrow e_2$
- 6) $V(x) - V(x + e_2 + e_3) \uparrow e_2 \uparrow e_3 \uparrow (e_2 + e_3) \downarrow e_1$
- 7) $V(x) - V(x + e_1 + e_2 + e_3) \uparrow e_1 \uparrow e_2 \uparrow e_3 \uparrow (e_1 + e_2) \uparrow (e_1 + e_3) \uparrow (e_2 + e_3) \uparrow (e_1 + e_2 + e_3)$

For 1), $\uparrow e_1$ means that optimal value function is concave in x_1 . $\downarrow e_2$ and $\downarrow e_3$ denote the supermodular of value function in directions (e_1, e_2) and (e_1, e_3) , respectively. Same meaning applies to 2) and 3).

For 4), $\uparrow(e_1 + e_2)$ means that the optimal value function is concave in $(e_1 + e_2)$. $\uparrow e_1$ and $\uparrow e_2$ denote the submodular of value function in directions $(e_1 + e_2, e_1)$ and $(e_1 + e_2, e_2)$ respectively. Also, it can be explained in the same way for 5), 6) and 7). We do not elaborate on them any more.

By inductive approach, we intend to prove that these structural properties are preserved by the optimal value function in dynamic programming equation. Unfortunately, for instance, the maximization operation $\max\{r_{4i} + V_{n-1}(x + e_1 + e_2), V_{n-1}(x)\}$ does not propagate the structure property $V(x) - V(x + e_1) \downarrow e_3$. Hence, exact analysis of this problem seems extremely difficult. To derive the proposed structural properties for this three-leg problem, relaxation is needed for ease of treatment. We remove the control on two-leg itineraries, that is, admission control is only imposed on those types of passengers for one-leg and three-leg itineraries. Then, we have following dynamic programming equation.

$$\begin{aligned}
 V_n(x) = & P_0^n V_{n-1}(x) + \sum_{j=1}^3 \sum_{i=1}^m P_{ji}^n \max\{r_{ij} + V_{n-1}(x + e_j), V_{n-1}(x)\} \\
 & + \sum_{i=1}^m P_{4i}^n \{r_{4i} + V_{n-1}(x + e_1 + e_2)\} + \sum_{i=1}^m P_{5i}^n \{r_{5i} + V_{n-1}(x + e_2 + e_3)\} \\
 & + \sum_{i=1}^m P_{6i}^n \max\{r_{6i} + V_{n-1}(x + e_1 + e_2 + e_3), V_{n-1}(x)\}
 \end{aligned} \tag{5-13}$$

For this problem, we have following theorems.

Theorem 5.3 Optimal value function $V_n(x_1, x_2)$ has the structural properties from 1) to 7).

Proof: As we see in the proof of lemma 5.1, there are four states involved in each checking case.

For simplification, we can bypass checking the possibilities when actions are all the same in the four states. That is, if it is in advantage to have admission controls in all four states or rejection controls in four states, then, our result follows trivially from the structural properties 1) - 7) and the

induction hypothesis. Now, the only interest is when actions are different in four states. The remaining proof follows the same checking processes as those for lemma 5.1. It is quite straightforward but tedious. Hence, it is omitted. ■

We define four switching functions as those of the round-trip/two-leg problems. Then, we can prove the existence of switching functions and the following theorem.

Theorem 5.4 The optimal booking policy is characterized by four state-dependent monotone switching surfaces.

Proof: The proof is the same as that of lemma 5.3. Here, it is omitted. ■

Of the four monotone switching surfaces, three of them are used, respectively, for the admission/rejection decisions of three independent legs. The fourth switching surface is used for the booking request for whole three legs, i.e., from departure point to the ending point.

Finally, structural properties and optimal booking policies for the problems with more than three legs can be extended in the same way as above. We do not discuss it anymore.

5.7 Numerical Example

Practically, in implementing the booking reservation system, management needs to obtain the demand information in advance by various forecasting techniques. With demand information and predetermined fare class, airline can compute the recursive dynamic programming equation off-line to determine the state-dependent booking and routing control limits characterized as switching curves/surfaces. All the information of these booking limits will be stored in system for future use. Upon an arrival of the booking request in the booking horizon, seat reservation system will decide whether or not to accept a booking request based on its arrival time, fare class and current booking status on both flights/legs.

Since structures of the optimal policy and monotone properties of the optimal value function are same for both round-trip problem and two-leg problem, we do not differentiate the two cases in our numerical study. In our example, the booking horizon is divided into 15 periods with $N = 15$ denote the beginning period and $N = 0$ denote the departure time of the flight. The switching curves are all state-dependent. Hence, we have a switching curve for each fare class of the three

types of passengers in each booking period. Here, we only give the switching curves for the booking request of low-fare class in period 12. The numbers 0, 1 and 2 in our figures indicate rejection, acceptance to flight 1 and acceptance to flight 2, respectively.

Table 5.1: Probabilities for the Booking Requests

Probability	Period (from 15 to 1)				
	15--12	11--9	8---5	4--3	2--1
p_{11}	0.06	0.1	0.14	0.15	0.15
p_{12}	0.14	0.1	0.16	0.15	0
p_{21}	0.06	0.1	0.14	0.15	0
p_{22}	0.14	0.1	0.16	0	0
p_{31}	0.086	0.143	0.182	0.25	0
p_{32}	0.2	0.143	0.143	0.2	0
p_0	0.314	0.314	0.075	0.1	0.85

The capacities for the two flights or two legs are $C_1 = 10, C_2 = 10$. For simplicity, we consider two fares: low fare and high fare with

$$r_{11} = 150, \quad r_{12} = 100, \quad r_{21} = 110, \quad r_{22} = 90, \quad r_{31} = 220, \quad r_{32} = 150$$

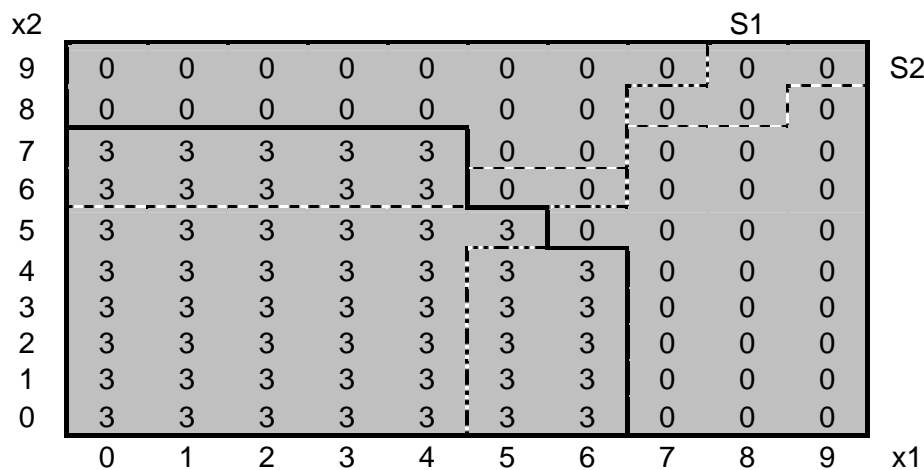


Figure 5.5: The Switching Curves for Type 3 Booking Requests

From the above figure, for low-fare booking request, we see that the dotted switching curves $S_1(x_2, 2)$ and $S_2(x_1, 2)$ are monotone nondecreasing and the solid curve $S_3(x_1, 2)$ are monotone nonincreasing. Upon an arrival of a booking request, airline seat reservation system will check current seat booking status on both flights/legs. If status lies above $S_3(x_1, 2)$, where decisions are indicated by 0, i.e., we reject the booking request. Otherwise, if current booking

status lies below $S_3(x_1, 2)$, where decisions are indicated by 3, accept the booking request. For instance, if currently 5 seats are booked on both flights/legs, the corresponding decision is indicated by 3 in the figure. Hence, it is better to accept the third type booking request. All the decision information in the figure is computed and stored before the start of the booking period. It will be used to make the decision upon the arrival of the booking request.

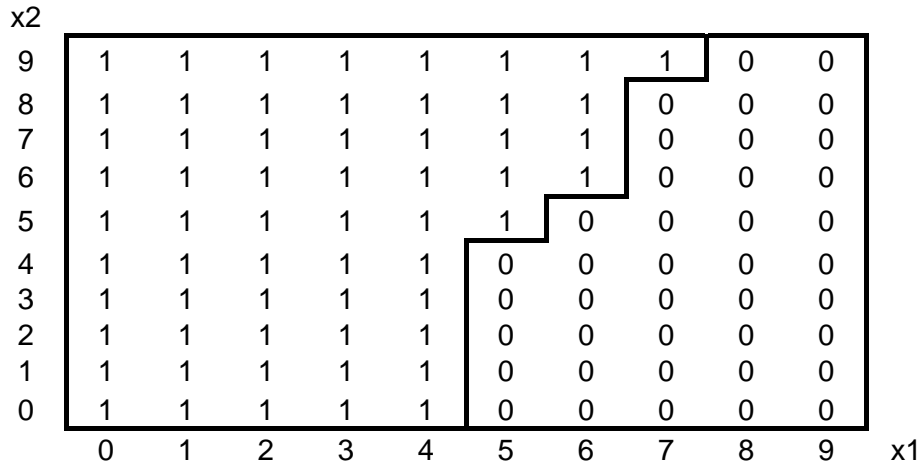


Figure 5.6: The Switching Curve for Type 1 Booking Requests

When a type 1 booking request for low fare arrives in period 12, airline seat reservation system will check the current seat booking status on both flight 1 and flight 2 or two legs. It tries to compare current booking status with the booking control information in Figure 5.6. If current booking status lies on the right side of the switching curve $S_1(x_2, 2)$, which is the area with decision 0, we reject the booking request for lower fare seat on flight 1 or leg 1. On the left side of the switching curve, which is the area with decision 1, accept the booking request. For instance, if the seat reservation system finds out that 4 seats and 5 seats are booked on flight 1 and flight 2 or leg 1 and leg 2 respectively, which corresponds to decision 1 in the figure, it will accept this booking request.

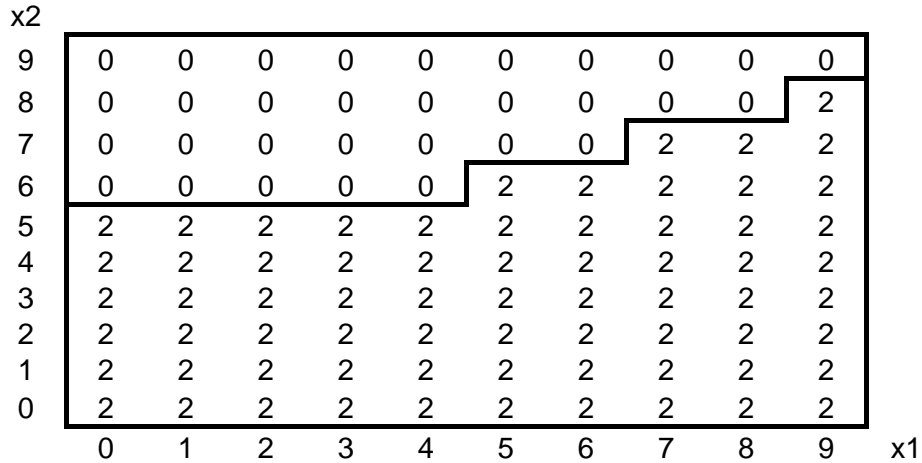


Figure 5.7: The Switching Curve for Type 2 Booking Requests

The above figure tells us that on the above of the switching curve of $S_2(x_1, 2)$, we reject the booking requests for low-fare class seats. Otherwise we accept the booking requests.

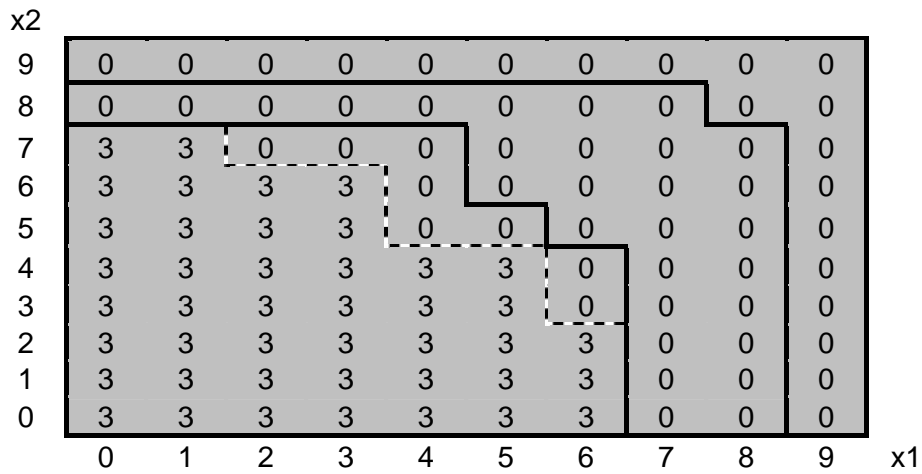


Figure 5.8: The Monotonicity of Switching Curve S_3

Here we give the switching curves for type 3 booking requests at periods 15, 12 and 6 in sequences as shown in the figure. It presents the monotone property of the switching curves in time periods for type 3 booking requests. As the departure approaching, the rejection region becomes smaller and smaller. It is more likely for airline to accept type 3 booking request.

Unfortunately, the monotonic in time periods has been lost in the switching curves of type 1 and type 2. To support our conclusion that $D_1V_n(x_1, x_2)$ and $D_2V_n(x_1, x_2)$ lose the monotonic in time period, we try an example to verify this. We adjust the arriving probabilities of booking requests in period 8 in table 5.1. The original probabilities $p_{11}=0.14$, $p_{12}=0.16$, $p_{21}=0.14$, $p_{22}=0.16$, $p_{31}=0.182$, $p_{32}=0.143$ are modified to $p_{11}=0.01$, $p_{12}=0.01$, $p_{21}=0.5$, $p_{22}=0.4$, $p_{31}=0.01$, $p_{32}=0.01$ with relatively

larger probabilities for type 2 booking requests. Our conclusion can be verified by the following figures.

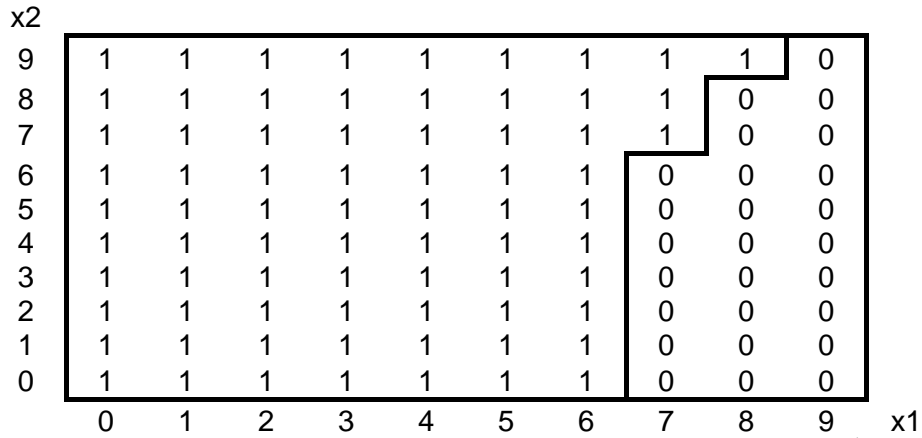


Figure 5.9: Switching Curve for the Lower Fare Class of Type 1 at the 7th period

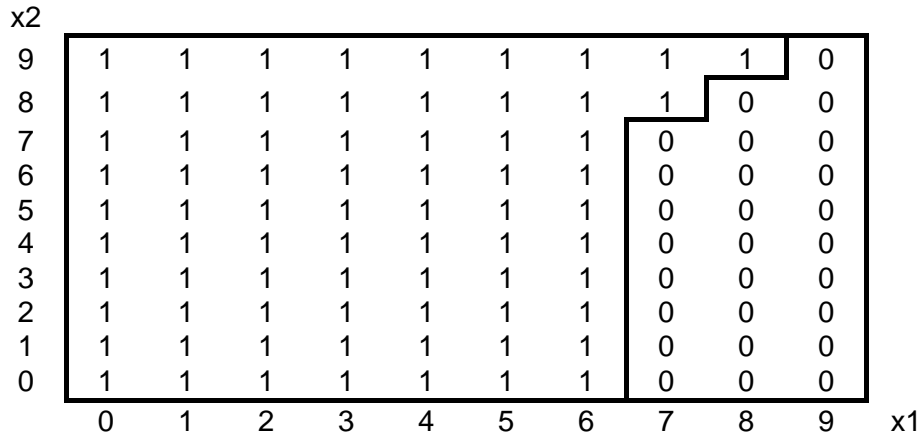


Figure 5.10: Switching Curve for the Lower Fare Class of Type 1 at the 8th period

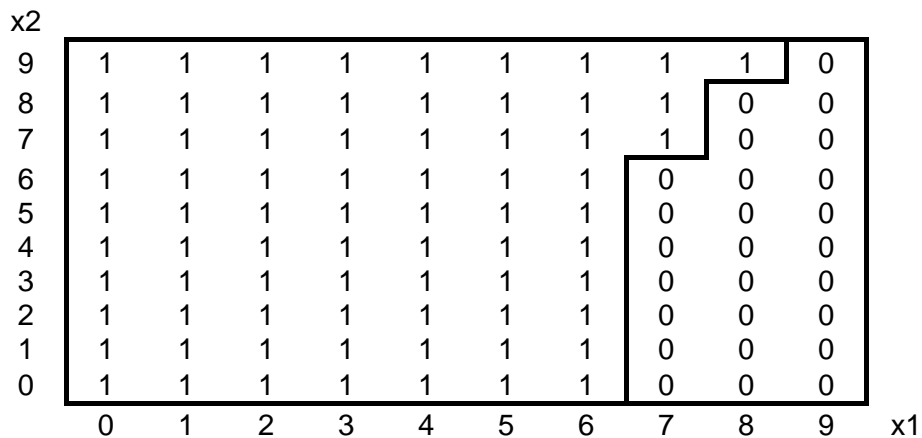


Figure 5.11: Switching Curve for the Lower Fare Class of Type 1 at the 9th period

We could see that the acceptance region in 8th period for type 1 lower fare class is no more than those of 7th period and 9th period, i.e. the switching curve is no longer monotonic in time period.

This implies that a booking request for the lower fare class of type 1 could be accepted in an earlier period but rejected in a later period.

Following figure shows the monotone property of $V(x_1, x_2) - V(x_1+1, x_2)$ at the 15th period, e.g. the starting period. It is increasing with respect to x_1 , e.g. concave in x_1 . With any fixed x_1 , it is decreasing with respect to x_2 , e.g. supermodular in (x_1, x_2) .

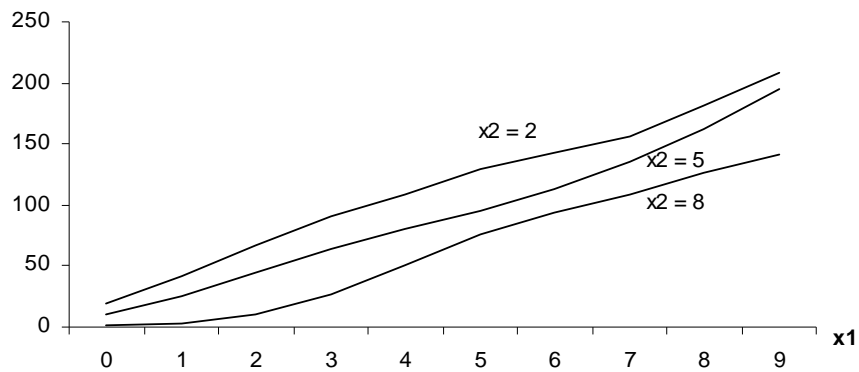


Figure 5.12: Monotone Properties of $V(x_1, x_2) - V(x_1+1, x_2)$.

Following figure illustrates, the monotone property of $V(x_1, x_2) - V(x_1, x_2+1)$ at the 15th period. It is increasing with respect to x_2 , e.g. concave in x_2 . With fixed x_2 , it is decreasing with respect to x_1 , e.g. supermodular in (x_1, x_2) .

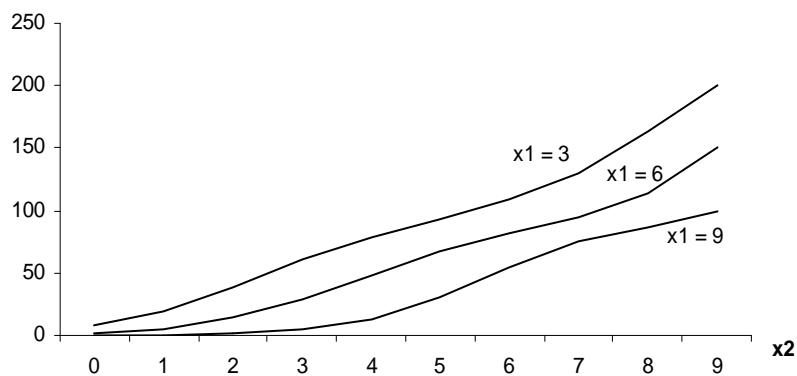


Figure 5.13: Monotone Properties of $V(x_1, x_2) - V(x_1, x_2+1)$.

Following figure illustrates, the monotone increasing property of $V(x_1, x_2) - V(x_1+1, x_2+1)$ at the 15th period. It is increasing with respect to x_2 for any fixed x_1 and increasing with respect to x_1 for any fixed x_2 .

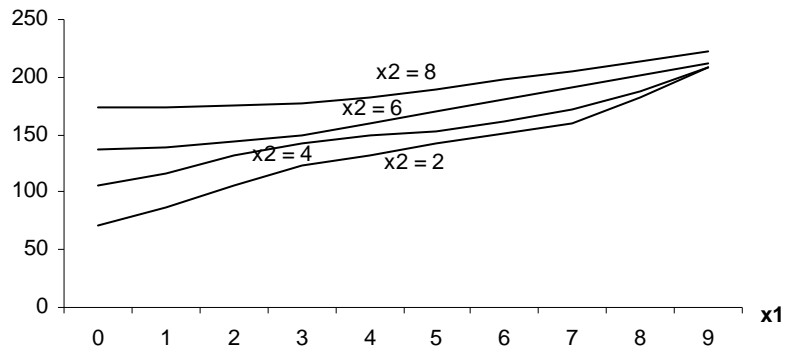


Figure 5.14: Monotone Properties of $V(x_1, x_2) - V(x_1+1, x_2+1)$.

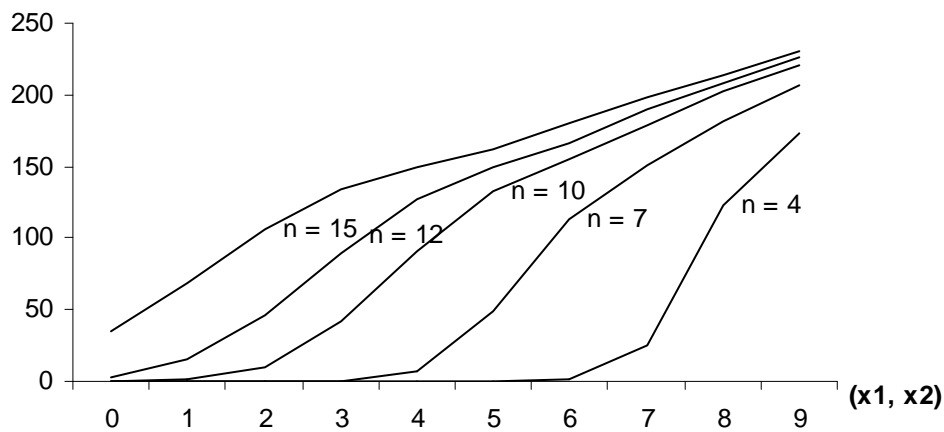


Figure 5.15: Monotonicity of $V(x_1, x_2) - V(x_1+1, x_2+1)$ with respect to n and $(x_1 + x_2)$.

In the above figure, a curve corresponds to the difference $V(x_1, x_2) - V(x_1+1, x_2+1)$ for any specified period n . It is monotone increasing in the direction of $(x_1 + x_2)$, e.g. the difference value is monotone increasing with respect to a simultaneous increase of x_1 and x_2 . Moreover, the

monotone increasing in period n for any fixed pair (x_1, x_2) implies the existence of time thresholds for the third type of passengers.

For the multiple-seat booking problem, we have following assumptions. First, $Q_1 = Q_2 = Q_3 = 3$, i.e. the largest demand quantity for each booking request is identically three seats with the discrete probability distribution for three types of booking requests as follow.

$$q_1(1) = 0.5, q_1(2) = 0.3, q_1(3) = 0.2$$

$$q_2(1) = 0.6, q_2(2) = 0.3, q_2(3) = 0.1$$

$$q_3(1) = 0.6, q_3(2) = 0.2, q_3(3) = 0.2$$

The other parameters are all the same as those of the single-seat booking example. The switching curves are all state-dependent. Hence, we only give the switching curves for the booking request of low-fare class in period 12. The numbers 0, 1, 2 and 3 in the figures for multiple-seat booking indicate decisions as rejection, acceptance with one seat, two seats and three seats, respectively.

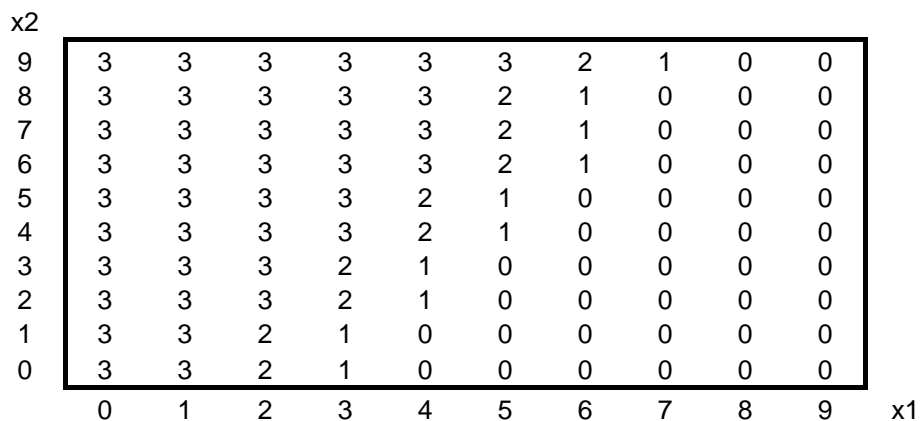


Figure 5.16: Decisions for Multiple-Seat Booking of Type 1 Booking Request

When a type 1 booking request for low fare arrives in period 12, airline seat reservation system will check the current seat booking status on both go-flight and return-flight or two legs. It tries to compare current booking status with the booking control information in Figure 5.16. If current booking status lies on the area with number 0, we reject the booking request. In the area with number 1, accept the booking request with one seat. In the area with number 2, accept it with two seats. Otherwise, it is optimal accept the booking request with three seats when current booking

status lies on the area with number 3. Boundaries of the decision zones of 0, 1, 2 and 3 are exactly the monotone switching curves defined in section 5.5.

x2										
9	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	1
6	0	0	0	0	0	0	0	1	1	2
5	0	0	0	0	1	1	1	2	2	3
4	0	0	0	1	2	2	2	3	3	3
3	1	1	1	2	3	3	3	3	3	3
2	2	2	2	3	3	3	3	3	3	3
1	3	3	3	3	3	3	3	3	3	3
0	3	3	3	3	3	3	3	3	3	3
	0	1	2	3	4	5	6	7	8	9
										x1

Figure 5.17: Decisions for Multiple-Seat Booking of Type 2 Booking Request

When a type 2 booking request for low fare arrives in period 12, airline seat reservation system will check the current seat booking status of x_1 and x_2 . If current booking status is in the area with number 0, reject the booking request. Correspondingly, in the area with number 1, accept the booking request with one seat. In the area with number 2, accept the booking request with two seats. Otherwise, it is optimal accept the booking request with three seats when current booking status lies on the area with number 3.

x2										
9	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
6	1	1	0	0	0	0	0	0	0	0
5	2	1	1	0	0	0	0	0	0	0
4	2	2	1	1	0	0	0	0	0	0
3	3	2	2	1	0	0	0	0	0	0
2	3	3	2	1	1	0	0	0	0	0
1	3	3	2	2	1	1	0	0	0	0
0	3	3	3	2	2	1	0	0	0	0
	0	1	2	3	4	5	6	7	8	9
										x1

Figure 5.18: Decisions for Multiple-Seat Booking of Type 3 Booking Request

The same interpretation for type 1 and type 2 booking requests can be applied for type 3 booking request as shown in Figure 5.18.

For a two-fare class multi-leg example, we randomly generate the arrival probabilities of different types of passengers. The fare prices for four types of passengers are shown as follow.

$$r_{11} = 100, r_{12} = 50, r_{21} = 90, r_{22} = 45, r_{31} = 80, r_{32} = 40, r_{41} = 240, r_{42} = 60$$

Numerical study yields similar results in round-trip/two-leg problems. We present the switching curves for lower fare class in the starting booking period. The admission/rejection switching surfaces in the three-leg problem are in three dimensions. Here, we give their projected switching curves on the planes X_1 - X_2 , X_1 - X_3 , and X_2 - X_3 , respectively.

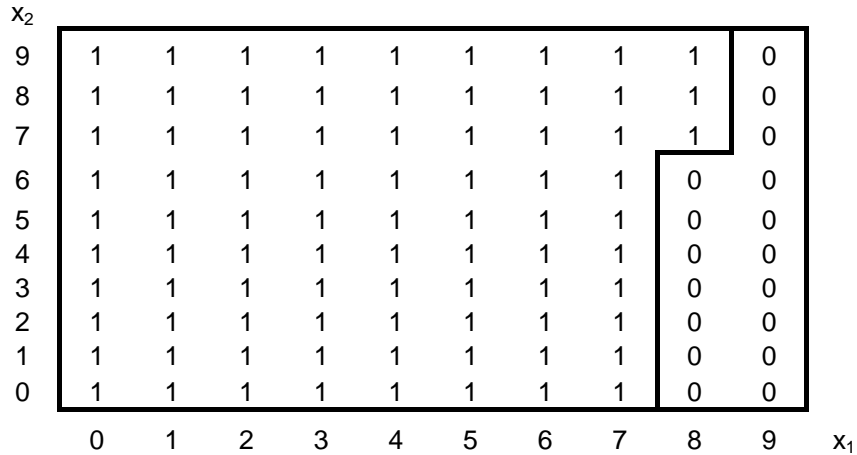


Figure 5.19: The Projected Switching Curve for Type 1

Figure 5.19 shows the projected switching curve on the plane X_1 - X_2 when $x_3 = 3$. It is used for the admission decisions of type 1 booking requests for leg 1. For instance, when a booking request arrives, seat reservation system will check the current booking status on three legs. It compares the booking status with the decision curve shown in the figure that has been computed prior to the booking horizon. If current booking status on three legs is $x_1 = 7, x_2 = 5, x_3 = 3$, which corresponds to decision 1 in the figure, it is better to accept the booking request. If current booking status on three legs is $x_1 = 8, x_2 = 6, x_3 = 3$, which corresponds to decision 0 in the figure, e.g., reject the booking request. In general, when booking status lies on the right side of the switching curve, accept the booking request, otherwise, reject it.

x_2										
9	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2
1	2	2	2	2	2	2	2	2	2	2
0	2	2	2	2	2	2	2	2	2	2
	0	1	2	3	4	5	6	7	8	9
										x_1

Figure 5.20: The Projected Switching Curve for Type 2

Figure 5.20 shows the projected switching curve on the plane X_1 - X_2 when $x_3 = 3$. It is intended for type 2 booking request. When a booking request arrives, seat reservation system will check current booking status on three legs. If current status is below the switching curve in the figure, e.g., decision area with 2, it is better to accept the booking request. On the above the switching curve with decision 0, e.g., reject the booking request.

x_3										
9	0	0	0	0	0	0	0	0	0	3
8	0	0	0	0	0	0	0	3	3	3
7	3	3	3	3	3	3	3	3	3	3
6	3	3	3	3	3	3	3	3	3	3
5	3	3	3	3	3	3	3	3	3	3
4	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3
2	3	3	3	3	3	3	3	3	3	3
1	3	3	3	3	3	3	3	3	3	3
0	3	3	3	3	3	3	3	3	3	3
	0	1	2	3	4	5	6	7	8	9
										x_1

Figure 5.21: The Projected Switching Curve for Type 3

Figure 5.21 is the projected switching curve on the plane X_1 - X_3 when $x_2 = 1$. It is used for the admission decisions for type 3 booking request which is for leg 3.

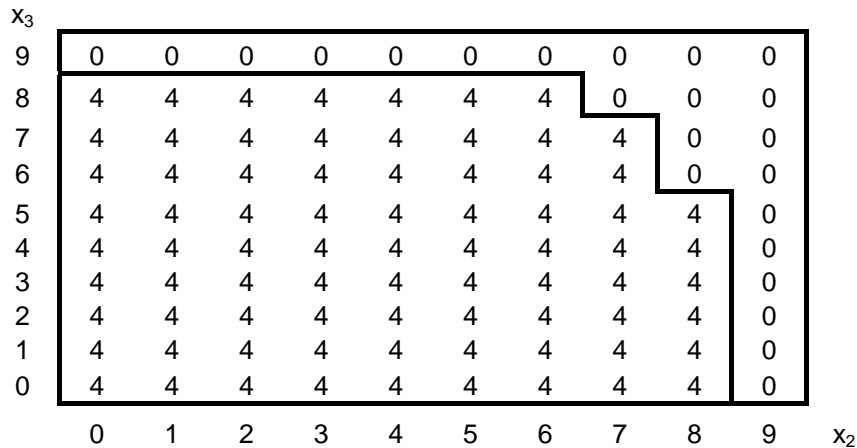


Figure 5.22: The Projected Switching Curve for Type 4

Figure 5.22 is the projected switching curve on the plane X_2 - X_3 with x_1 equal to 7. It is used for the decisions for type 4 booking request which is for the whole three legs. For instance, if current booking status for leg 1, leg 2 and leg 3 are 7, 7 and 6 seats, respectively, e.g., the corresponding decision is 4. Hence, it is better to accept the booking request. If the current booking status is $x_1 = 7$, $x_2 = 8$ and $x_3 = 6$ and the corresponding decision is indicated by 0, e.g., reject the booking request.

5.8 Conclusions

In this chapter, we have examined the structures of the optimal booking policy for round-trip/two-leg flight problems. Based on the structural property for the optimal value function, we have characterized the structure of the optimal booking policy as three monotone switching curves for each fare class. In an extension, optimal booking policy for multiple-seat booking problem has been completely characterized as switching curves if booking requests can be partially satisfied. Furthermore, for multi-leg problem, we explore the structure properties of the optimal value function and partially characterize the structure of the optimal booking policy. Practically, many airline seat inventory control problems are multi-dimensional when formulated as Markov decision process. However, as we know, such problems are hard to tackle owing to the curse of dimension for dynamic programming. Therefore, it is crucial to find some basic structures for the optimal policy. These structures can either improve the efficiency of computation or help us design effective heuristic booking policies.

For future research, we might consider round-trip/two-leg problems with cancellations and overbooking. Moreover, for the round-trip with multi-flight, problem would be more complicated. Therefore, efficient heuristic algorithms are definitely needed to solve these problems.

5.9 Appendix:

Proof of Lemma 5.1

In our proof, we directly use increasing and decreasing to indicate nondecreasing and nonincreasing respectively.

First, we will show that if $V \in \hat{V}$, then $T_{1,i}V \in \hat{V}$. Define

$$f(x_1, x_2) = T_{1,i}V(x_1, x_2) = \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\}$$

Proof of Property (a): $f(x_1, x_2) - f(x_1 + 1, x_2) \uparrow x_1 \downarrow x_2$

Note that

$$\begin{aligned} & f(x_1, x_2) - f(x_1 + 1, x_2) \\ &= \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \end{aligned} \quad (\text{A1})$$

and

$$\begin{aligned} & f(x_1 + 1, x_2) - f(x_1 + 2, x_2) \\ &= \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A2})$$

In order to show that $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$, i.e., $f(x_1, x_2)$ is concave in x_1 , we have to check all cases.

Since $V \in \hat{V}$, from (A1), (A2) and structural properties of $V(x_1, x_2)$, yields

$$\begin{aligned} & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 2, x_2) \\ & \leq r_{1i} + V(x_1 + 2, x_2) - r_{1i} - V(x_1 + 3, x_2) \\ & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - r_{1i} - V(x_1 + 3, x_2) \\ & \\ & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 2, x_2) \\ & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2) \end{aligned}$$

Combining the above two inequalities, yields

$$\begin{aligned} & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A3})$$

Similarly, we can check that

$$\begin{aligned} & V(x_1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned} \quad (\text{A4})$$

Combining (A3) and (A4), we have

$$\begin{aligned} & \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} \\ & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2), V(x_1 + 2, x_2)\} \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$, that is, $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_1 .

We now proceed to show that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is decreasing in x_2 , i.e.,

$$f(x_1, x_2) - f(x_1 + 1, x_2) \geq f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1)$$

or equivalently, $f(x_1, x_2) - f(x_1, x_2 + 1) \geq f(x_1 + 1, x_2) - f(x_1 + 1, x_2 + 1)$.

This is exactly the supermodularity of the value function.

Note that

$$\begin{aligned} & f(x_1, x_2) - f(x_1, x_2 + 1) \\ & = \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{1i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \end{aligned} \quad (\text{A5})$$

and

$$\begin{aligned} & f(x_1 + 1, x_2) - f(x_1 + 1, x_2 + 1) \\ & = \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A6})$$

By the structural properties of $V(x_1, x_2)$, we have

$$\begin{aligned} & \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - r_{1i} - V(x_1 + 1, x_2 + 1) \\ & \geq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 1, x_2 + 1) \\ & \geq r_{1i} + V(x_1 + 2, x_2) - r_{1i} - V(x_1 + 2, x_2 + 1) \\ & \geq r_{1i} + V(x_1 + 2, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} & \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - r_{1i} - V(x_1 + 1, x_2 + 1) \\ & \geq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 1, x_2 + 1) = V(x_1 + 1, x_2) - V(x_1 + 1, x_2 + 1) \\ & \geq V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A8})$$

Combining (A7) and (A8), yields

$$\begin{aligned} & \max \{r_{i_i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - r_{i_i} - V(x_1 + 1, x_2 + 1) \\ & \geq \max \{r_{i_i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{i_i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A9})$$

Similarly, we can check that

$$\begin{aligned} & \max \{r_{i_i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - V(x_1, x_2 + 1) \\ & \geq \max \{r_{i_i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{i_i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned} \quad (\text{A10})$$

Combining (A9) and (A10), we have

$$\begin{aligned} & \max \{r_{i_i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{i_i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2 + 1)\} \\ & \geq \max \{r_{i_i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{i_i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2)$ is decreasing in x_2 .

Proof of Property (b): $f(x_1, x_2) - f(x_1, x_2 + 1) \downarrow x_1 \uparrow x_2$

It is clear that Property (b) is symmetrical to Property (a) with respect to x_1 and x_2 , respectively, and hence our result applies.

Proof of Property (c): $f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \uparrow x_1 \uparrow x_2$

By symmetry, we need to prove the case of increasing in x_1 only. This is to show that $f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2 + 1)$.

Note that

$$\begin{aligned} & f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \\ & = \max \{r_{i_i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{i_i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\ & f(x_1 + 1, x_2) - f(x_1 + 2, x_2 + 1) \\ & = \max \{r_{i_i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{i_i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2 + 1)\} \end{aligned}$$

We check different cases. First we have

$$\begin{aligned} & r_{i_i} + V(x_1 + 1, x_2) - \max \{r_{i_i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\ & \leq r_{i_i} + V(x_1 + 1, x_2) - r_{i_i} - V(x_1 + 2, x_2 + 1) \\ & \leq r_{i_i} + V(x_1 + 2, x_2) - r_{i_i} - V(x_1 + 3, x_2 + 1) \\ & \leq \max \{r_{i_i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - r_{i_i} - V(x_1 + 3, x_2 + 1) \end{aligned} \quad (\text{A11})$$

$$\begin{aligned}
 & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 2, x_2 + 1) \\
 & = V(x_1 + 1, x_2) - V(x_1 + 2, x_2 + 1) \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2 + 1)
 \end{aligned} \tag{A12}$$

Thus, we have

$$\begin{aligned}
 & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2 + 1)\}
 \end{aligned} \tag{A13}$$

By the same checking process as the above, we can obtain

$$\begin{aligned}
 & V(x_1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2 + 1)\}
 \end{aligned} \tag{A14}$$

Combining (A13) and (A14) gives that

$$\begin{aligned}
 & \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2), V(x_1 + 1, x_2)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2 + 1)\}
 \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2 + 1)$ is increasing in x_1 . The rest half of Property (c) is obvious due to the symmetry between x_1 and x_2 .

Proof of Property (d): $f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \uparrow (x_1 + x_2)$

We check different cases. First we have

$$\begin{aligned}
 & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 2, x_2 + 1) \\
 & \leq r_{1i} + V(x_1 + 2, x_2 + 1) - r_{1i} - V(x_1 + 3, x_2 + 2) \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - r_{1i} - V(x_1 + 3, x_2 + 2)
 \end{aligned} \tag{A15}$$

$$\begin{aligned}
 & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq r_{1i} + V(x_1 + 1, x_2) - r_{1i} - V(x_1 + 2, x_2 + 1) \\
 & \leq V(x_1 + 1, x_2 + 1) - V(x_1 + 2, x_2 + 2) \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - V(x_1 + 2, x_2 + 2)
 \end{aligned} \tag{A16}$$

Thus, the above two inequalities give that

$$\begin{aligned}
 & r_{1i} + V(x_1 + 1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 2)\}
 \end{aligned} \tag{A17}$$

By the same approach, we can obtain

$$\begin{aligned}
 & V(x_1, x_2) - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 2)\}
 \end{aligned} \tag{A18}$$

Combining (A17) and (A18) gives that

$$\begin{aligned}
 & \max \{r_{1i} + V(x_1 + 1, x_2), V(x_1, x_2)\} - \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{1i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2 + 1)\} - \max \{r_{1i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 2)\}
 \end{aligned}$$

Therefore, $f(x_1, x_2) - f(x_1 + 1, x_2 + 1)$ is increasing in $(x_1 + x_2)$.

Thus, $f(x_1, x_2) = T_{1,i}V \in \hat{V}$, if $V \in \hat{V}$.

The verification for $T_{2,i}V \in \hat{V}$ if $V \in \hat{V}$ is the same as the above proof, and is thus omitted.

We now proceed to prove that $T_{3,i}V \in \hat{V}$ if $V \in \hat{V}$. The proof technique is the same as the proof of $T_{1,i}V \in \hat{V}$ and $T_{2,i}V \in \hat{V}$. Let

$$f(x_1, x_2) = T_{3,i}V = \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\}$$

We need to prove that $f(x_1, x_2)$ satisfies the properties (a), (b), (c) and (d).

For Property (a), we first show that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_1 or

$f(x_1, x_2)$ is concave in x_1 , i.e., $f(x_1, x_2) - f(x_1 + 1, x_2) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2)$.

$$\begin{aligned}
 & f(x_1, x_2) - f(x_1 + 1, x_2) \\
 & = \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}
 \end{aligned} \tag{A19}$$

$$\begin{aligned}
 & f(x_1 + 1, x_2) - f(x_1 + 2, x_2) \\
 & = \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2)\}
 \end{aligned} \tag{A20}$$

Checking through different cases, we have

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - r_{3i} - V(x_1 + 2, x_2 + 1) \\
 & \leq r_{3i} + V(x_1 + 2, x_2 + 1) - r_{3i} - V(x_1 + 3, x_2 + 1) \\
 & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - r_{3i} - V(x_1 + 3, x_2 + 1)
 \end{aligned} \tag{A21}$$

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - r_{3i} - V(x_1 + 2, x_2 + 1) \\
 & \leq V(x_1 + 1, x_2) - V(x_1 + 2, x_2) \\
 & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2)
 \end{aligned} \tag{A22}$$

Combining (A21) and (A22) we obtain that

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2)\}
 \end{aligned} \tag{A23}$$

Similarly we can obtain

$$\begin{aligned}
 & V(x_1, x_2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2)\}
 \end{aligned} \tag{A24}$$

and combine (A23) and (A24) we have

$$\begin{aligned}
 & \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 1), V(x_1 + 2, x_2)\}
 \end{aligned}$$

Thus $f(x_1, x_2) - f(x_1 + 1, x_2)$ is increasing in x_1 . By the same method we will prove that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is decreasing in x_2 . That is

$$f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1) \leq f(x_1, x_2) - f(x_1 + 1, x_2). \text{ Note that}$$

$$f(x_1, x_2 + 1) = \max \{r_{3i} + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 1)\}$$

$$f(x_1 + 1, x_2 + 1) = \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\}$$

$$f(x_1, x_2) = \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\}$$

$$f(x_1 + 1, x_2) = \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}$$

Checking for different cases, we have

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\
 & \leq r_{3i} + V(x_1 + 1, x_2 + 2) - r_{3i} - V(x_1 + 2, x_2 + 2) \leq r_{3i} + V(x_1 + 1, x_2 + 1) - r_{3i} - V(x_1 + 2, x_2 + 1) \\
 & \leq \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - r_{3i} - V(x_1 + 2, x_2 + 1)
 \end{aligned}$$

and

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\
 & \leq r_{3i} + V(x_1 + 1, x_2 + 2) - V(x_1 + 1, x_2 + 1) \leq r_{3i} + V(x_1 + 1, x_2 + 1) - V(x_1 + 1, x_2) \\
 & \leq \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - V(x_1 + 1, x_2)
 \end{aligned}$$

Combining the above two inequalities we have

$$\begin{aligned}
 & r_{3i} + V(x_1 + 1, x_2 + 2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}
 \end{aligned}$$

Similarly, we can obtain

$$\begin{aligned}
 & V(x_1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}
 \end{aligned}$$

Finally combining the above two inequality leads to

$$\begin{aligned}
 & \max \{r_{3i} + V(x_1 + 1, x_2 + 2), V(x_1, x_2 + 1)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\
 & \leq \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}
 \end{aligned}$$

This is exactly the inequality $f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1) \leq f(x_1, x_2) - f(x_1 + 1, x_2)$ which indicates that $f(x_1, x_2) - f(x_1 + 1, x_2)$ is decreasing in x_2 .

Hence, $f(x_1, x_2)$ holds the property a). Similarly, we can prove that $f(x_1, x_2)$ holds **property b)**.

For Property (c), we first show that $f(x_1, x_2) - f(x_1 + 1, x_2 + 1)$ is increasing in x_1 , that is,

$$f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \leq f(x_1 + 1, x_2) - f(x_1 + 2, x_2 + 1). \text{ Note that}$$

$$f(x_1, x_2) = \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\}$$

$$f(x_1 + 1, x_2 + 1) = \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\}$$

$$f(x_1 + 1, x_2) = \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\}$$

$$f(x_1 + 2, x_2 + 1) = \max \{r_{3i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 1)\}$$

Checking for different cases, we have

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - r_{3i} - V(x_1 + 2, x_2 + 2) \leq r_{3i} + V(x_1 + 2, x_2 + 1) - r_{3i} - V(x_1 + 3, x_2 + 2) \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - r_{3i} - V(x_1 + 3, x_2 + 2) \end{aligned}$$

and

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - V(x_1 + 1, x_2 + 1) = r_{3i} + V(x_1 + 2, x_2 + 1) - V(x_1 + 2, x_2 + 1) \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - V(x_1 + 2, x_2 + 1) \end{aligned}$$

Combining the above two inequalities we have

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 1)\} \end{aligned}$$

Similarly, we can obtain

$$\begin{aligned} & V(x_1, x_2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 1)\} \end{aligned}$$

Finally combining the above two inequality leads to

$$\begin{aligned} & \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 1), V(x_1 + 1, x_2)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 2), V(x_1 + 2, x_2 + 1)\} \end{aligned}$$

Thus, $f(x_1, x_2) - f(x_1 + 1, x_2 + 1)$ is increasing in x_1 . Symmetrically it is also increasing in x_2 ,

and hence Property (c) applies.

For Property (d), we check different cases. First we have

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - r_{3i} - V(x_1 + 2, x_2 + 2) \\ & \leq r_{3i} + V(x_1 + 2, x_2 + 2) - r_{3i} - V(x_1 + 3, x_2 + 3) \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} - r_{3i} - V(x_1 + 3, x_2 + 3) \end{aligned} \tag{A25}$$

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq r_{3i} + V(x_1 + 1, x_2 + 1) - V(x_1 + 1, x_2 + 1) \\ & = r_{3i} + V(x_1 + 2, x_2 + 2) - V(x_1 + 2, x_2 + 2) \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} - V(x_1 + 2, x_2 + 2) \end{aligned} \tag{A26}$$

Thus, we have

$$\begin{aligned} & r_{3i} + V(x_1 + 1, x_2 + 1) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 3), V(x_1 + 2, x_2 + 2)\} \end{aligned} \quad (\text{A27})$$

By the same approach as above, we can obtain

$$\begin{aligned} & V(x_1, x_2) - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 3), V(x_1 + 2, x_2 + 2)\} \end{aligned} \quad (\text{A28})$$

Combining (A27) and (A28) gives that

$$\begin{aligned} & \max \{r_{3i} + V(x_1 + 1, x_2 + 1), V(x_1, x_2)\} - \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} \\ & \leq \max \{r_{3i} + V(x_1 + 2, x_2 + 2), V(x_1 + 1, x_2 + 1)\} - \max \{r_{3i} + V(x_1 + 3, x_2 + 3), V(x_1 + 2, x_2 + 2)\} \end{aligned}$$

This is the inequality $f(x_1, x_2) - f(x_1 + 1, x_2 + 1) \leq f(x_1 + 1, x_2 + 1) - f(x_1 + 2, x_2 + 2)$

Therefore, $T_{3,i}V \in \hat{V}$ if $V \in \hat{V}$.

We have shown that properties (a), (b), (c) and (d) of the optimal value function are preserved under the operators $T_{1,i}$, $T_{2,i}$ and $T_{3,i}$. Since operator T^n is a convex combination of $T_{1,i}$,

$T_{2,i}$, $T_{3,i}$ and V , and thus $T^nV \in \hat{V}$ if $V \in \hat{V}$. This completes the proof. ■

Proof of Lemma 5.5

First, we will show that if $V \in \hat{V}$, then $T_{1,i}V \in \hat{V}$. Since demand is finite, we will prove the results by induction on demand. For type 1 passengers, we have the operator

$$T_{1,i}V(x) = \max_{0 \leq j_1 \leq \{d_1, C\}} \{j_1 r_{1i} + V(x + j_1 e_1)\}$$

Since demand d_1 is finite, we only need to show that

$$f(x) = \max \{V(x), r_{1i} + V(x + e_1), \dots, k r_{1i} + V(x + k e_1)\}$$

preserves the structural properties (a), (b), (c) and (d) for any positive integer k .

Note that for the case of single-unit booking request, our structural properties are preserved under the control operators. Suppose that it is also true for case of $k - 1$ unit of the booking request.

For Property (a): $f(x) - f(x + e_1) \uparrow x_1 \downarrow x_2$, note that

$$\begin{aligned}
 & f(x) - f(x + e_1) \\
 &= \max \left\{ \max \left\{ V(x), \dots, (k-1)r_i + V(x + (k-1)e_1) \right\}, kr_i + V(x + ke_1) \right\} \\
 &\quad - \max \left\{ \max \left\{ V(x + e_1), \dots, (k-1)r_i + V(x + e_1) \right\}, kr_i + V(x + (k+1)e_1) \right\} \\
 &= LHS1 \{ LHS1(1), LHS1(2) \} - LHS2 \{ LHS2(1), LHS2(2) \}
 \end{aligned} \tag{A29}$$

and

$$\begin{aligned}
 & f(x + e_1) - f(x + 2e_1) \\
 &= \max \left\{ \max \left\{ V(x + e_1), \dots, (k-1)r_i + V(x + ke_1) \right\}, kr_i + V(x + (k+1)e_1) \right\} \\
 &\quad - \max \left\{ \max \left\{ V(x + 2e_1), \dots, (k-1)r_i + V(x + (k+1)e_1) \right\}, kr_i + V(x + (k+2)e_1) \right\} \\
 &= RHS1 \{ RHS1(1), RHS1(2) \} - RHS2 \{ RHS2(1), RHS2(2) \}
 \end{aligned} \tag{A30}$$

In order to show that $f(x) - f(x + e_1) \leq f(x + e_1) - f(x + 2e_1)$, i.e., $f(x)$ is concave in x_1 , we have to check all cases. To be noted that if all operation actions take the first term under the outer maximization operation, e.g. we need to prove

$$LHS1(1) - LHS2(1) \leq RHS1(1) - RHS2(1) \tag{A31}$$

This is true based on our assumption for $k - 1$.

If all operation actions take the second term under the outer maximization operation, e.g. we need to prove

$$LHS1(2) - LHS2(2) \leq RHS1(2) - RHS2(2), \tag{A32}$$

i.e.

$$kr_i + V(x + ke_1) - \{kr_i + V(x + (k+1)e_1)\} \leq kr_i + V(x + (k+1)e_1) - \{kr_i + V(x + (k+2)e_1)\}$$

This is an evident result. Next, we only need to consider following two cases.

$$LHS1(1) - LHS2 \leq RHS1 - RHS2(2) \tag{A33}$$

$$\text{and } LHS1(2) - LHS2 \leq RHS1 - RHS2(1). \tag{A34}$$

It is readily to check each case based on the structural properties (a), (b), (c) and (d) and (P.a), (P.b) and (P.c). The checking process is tedious but quite straightforward. For instance, for (A33), it is true from the fact that

$$\begin{aligned}
 V(x) - V(x + e_1) &\leq kr_i + V(x + (k + 1)e_1) - \{kr_i + V(x + (k + 2)e_1)\} \\
 r_i + V(x + e_1) - \{r_i + V(x + 2e_1)\} &\leq kr_i + V(x + (k + 1)e_1) - \{kr_i + V(x + (k + 2)e_1)\} \\
 &\vdots \\
 &\vdots \\
 (k - 1)r_i + V(x + (k - 1)e_1) - \{(k - 1)r_i + V(x + e_1)\} \\
 &\leq kr_i + V(x + (k + 1)e_1) - \{kr_i + V(x + (k + 2)e_1)\}
 \end{aligned}$$

that is $LHS1(1) - LHS2(1) \leq RHS1(2) - RHS2(2)$. Further, in the same checking process, we can demonstrate the truth of (A34). Consequently, based on the results (A31), (A32), (A33) and (A34), for any positive number k , $f(x)$ preserves Property (a). Hence, $T_{1,i}V$ preserves Property (a). Similarly, we can prove that $T_{2,i}V(x)$, $T_{3,i}V(x)$ and $T^nV(x)$ all preserve Property (a).

Moreover, it can be argued in the same way that $T_{1,i}V(x)$, $T_{2,i}V(x)$, $T_{3,i}V(x)$ and $T^nV(x)$ preserve Properties (b), (c) and (d). That completes our proof. ■

6. Stock Rationing of a Make-To-Stock Production/Inventory System

6.1 Introduction

In the research of production/inventory systems, many of the previous studies primarily focus on the single-class demand. Currently, it is a trend with the increasing product proliferation and demand variety. To achieve a better service for customers, it needs to differentiate customer demand for market segment. This arouses interests both in academia and industries to consider multiple classes of demand. Since resource is limited and customer demands are of different economic value to suppliers, it is reasonable to reserve stock for some more important class of demands. Hence, to minimize the total penalty cost by denying high priority demand, management tries to ration stock to satisfy different priority class demands. When a class of customer demand arrives, it needs to decide whether satisfy the demands or leave the demand unsatisfied and hold the stock for later higher priority demands.

In this chapter, following the make-to-stock queue setting, we examine a two-item, two production facilities production/inventory system with lost sales. Customer demands are classified into three types with several classes. Customer orders in each class arrive according to Poisson process and each customer requests a single unit of item. There are three types of customers: the first and second type order, respectively, for the first and the second item only; whereas the third type is flexible and willing to take either of the two items. Upon the arrival of demand, management first identifies the type of customers, and then makes the decision on whether accept it and, when a third type is accepted, which item to satisfy it. These two items are of the same generic type but may differ in other aspects like quality, price, and etc. The flexible customers can be satisfied by their main functionality regardless of others. The objective is to minimize the total expected cost through controlling of inventories of two items and rationing of demands. For this two-item stock rationing systems, we completely characterize the structures of the optimal policy as switching curves.

Similarly, Gallego and Phillips (2004) also considered three types of customers with decision horizon of two periods with demand postponement strategy. Flexible demands accepted in the first period will be satisfied with a product in the second period or denied with a penalty payment. Differently, our model follows the stock rationing setting and the acceptance/rejection decisions are made immediately upon the arrival of customer demand.

6.2 Model Formulation

Suppose that there are two stocks of items for customers. The management tries to minimize the total expected cost of these two items through stock rationing and production control. Here, we assume that the production processes for the two items are separate. It means that the two items are independently produced on two distinct facilities. The customer demands are classified into three types. The demands of the first two types are specifically for one of the two items. The third type is flexible, which can be satisfied by either of the two items. Further, it is assumed that the arrival processes for the three types of customer demands are independent.

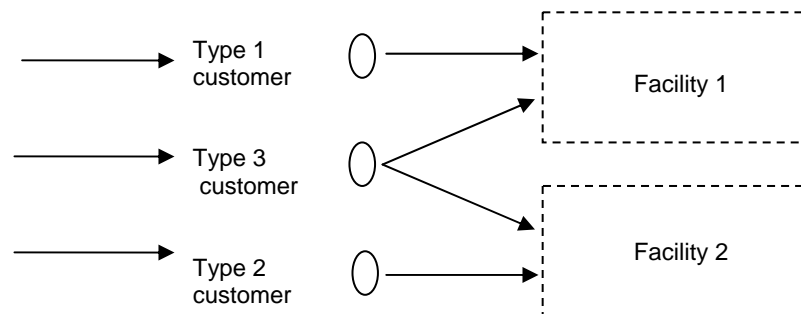


Figure 6.1: Three Types of Demands for Two Items

6.2.1 Dynamic Programming Equation

Assume that there are m fare classes, and for $i = 1, \dots, m$, π_{1i} , π_{2i} and π_{3i} are the penalty costs incurred by rejecting type 1, type 2 and type 3 customers of class i , respectively. Without loss of generality, $\pi_{j1} > \pi_{j2} > \dots > \pi_{jm}$, $j = 1, 2, 3$. The priorities of customers are in the descending order with class 1 being the highest one. Here, x_1 and x_2 are state variables, denoting the inventory levels of two items, respectively. $h(x_1, x_2) = h_1(x_1) + h_2(x_2)$ is the holding cost per unit of time for the two items where $h_1(x_1) = 0$ when $x_1 \leq 0$ and $h_2(x_2) = 0$ when $x_2 \leq 0$. Here,

different from previous research with objective of reducing inventory holding costs and penalty costs, we consider different production costs with various production rates. Assume that we have $i = 1, \dots, n$ classes of production rates for two production facilities with corresponding production cost of c_{1i} and c_{2i} , respectively with $i = 1, \dots, n$. Just after one item is finished, a unit production cost with corresponding production rate is incurred. Without loss of generality, we assume $c_{j1} < c_{j2} < \dots < c_{jn}$, $j = 1, 2$. Hence, the total costs considered in this paper include inventory holding costs, penalty costs incurred by rejecting customer demand and the production cost with the selected production rate.

It is assumed that the arrival rates for each class of three types of customer demands are $\lambda_{1i}, \lambda_{2i}$, and λ_{3i} , respectively, with $i = 1, \dots, m$. Let μ_{1i} and μ_{2i} be the production rates of two items respectively with $i = 1, \dots, n$. Following the uniformization technique in Lippman (1975), let

$$\Lambda = \sum_{i=1}^m (\lambda_{1i} + \lambda_{2i} + \lambda_{3i}) + \sum_{i=1}^n (\mu_{1i} + \mu_{2i})$$

After uniformization, the continuous-time system is only observed at discrete, randomly spaced points in time, and each point of time corresponds to an arrival of a customer or a finishing production of one unit of item. $V(x)$ is the expected minimum α -discounted cost function when current state is $x = (x_1, x_2)$. Policy u is a rule for choosing actions at the state transition points. Among the admissible policies, we will find the optimal policy that minimizes the total expected discount cost over the infinite horizon. Given initial inventory $x = (x_1, x_2)$, we have

$$J_u(x) = E_{x_0}^u \left[\int_0^\infty e^{-\alpha t} h(x_1, x_2) dt + \sum_{j=1}^3 \sum_{i=1}^m \int_0^\infty e^{-\alpha t} \pi_{ji} dM_{ji}(t) + \sum_{j=1}^2 \sum_{i=1}^n \int_0^\infty e^{-\alpha t} c_{ji} dN_{ji}(t) \right]$$

Where $M_{ji}(t)$ and $N_{ji}(t)$ represent, respectively, up to time t , the number rejected customers from the i -th class and j -th type, and the number of items finished with the i -th production rate, j -th production facility. And the value function $V(x)$ is the infimum of $J_u(x)$ over all admissible policies $u \in U$

$$V(x) = \inf_{u \in U} J_u(x)$$

where U is total set of admissible control policies.

When customer demands arrive, rationing decisions $u_d \in u$ can take values 0, 1, 2 which indicate the decisions of rejecting, satisfying the demand with item 1 and satisfying the demand with item 2, respectively. At any point in time, production decisions $u_p \in u$ can take values 0, 1, 2 denoting stop production, produce item 1 and produce item 2 respectively. Let us assume, given current state (x_1, x_2) and for each control value of u_d and u_p , the next state after transition is y . Define the transition probability function

$$p(y | x, u) = \sum_{i=1}^m \frac{\lambda_{1i} + \lambda_{3i}}{\Lambda} I\{y = (x_1 - 1, x_2), u_d = 1\} + \sum_{i=1}^m \frac{\lambda_{2i} + \lambda_{3i}}{\Lambda} I\{y = (x_1, x_2 - 1), u_d = 2\} \\ + \sum_{i=1}^n \frac{\mu_{1i}}{\Lambda} I\{y = (x_1 + 1, x_2), u_p = 1\} + \sum_{i=1}^n \frac{\mu_{2i}}{\Lambda} I\{y = (x_1, x_2 + 1), u_p = 2\}$$

This illustrates that with probability $\sum_{i=1}^m \frac{\lambda_{1i} + \lambda_{3i}}{\Lambda} I\{u_d = 1\}$, one unit of item 1 is taken away by the customer, hence, the next state after transition is $(x_1 - 1, x_2)$. Similarly, with probability $\sum_{i=1}^m \frac{\lambda_{2i} + \lambda_{3i}}{\Lambda} I\{u_d = 2\}$, one unit of item 2 is taken away by the customer, hence, the next state after transitions is $(x_1, x_2 - 1)$. For the control of production processes, with probabilities

$$\sum_{i=1}^n \frac{\mu_{1i}}{\Lambda} I\{u_p = 1\} \text{ and } \sum_{i=1}^n \frac{\mu_{2i}}{\Lambda} I\{u_p = 2\}, \text{ the next state after transitions are } (x_1 + 1, x_2) \text{ and } (x_1, x_2 + 1), \text{ respectively.}$$

It is evident that Markovian policies are optimal for this Markov decision model with unified transition rate Λ and the transition probability p defined above. Without loss of generality, by scaling the time, we have $\alpha + \Lambda = 1$. Here, $0 < \alpha < 1$ is the discount rate.

Now we define following operators $T_{1,i}$, $T_{2,i}$, $T_{3,i}$, $T_{4,i}$, $T_{5,i}$, and T on the value function $V(x)$ with state variables $x = (x_1, x_2) \in Z_+^2$.

$$T_{1,i}V = \begin{cases} \min \{V(x - e_1), V(x) + \pi_{1i}\} & \text{if } x_1 > 0 \\ V(x) + \pi_{1i} & \text{Otherwise} \end{cases}$$

$$T_{2,i}V = \begin{cases} \min \{V(x - e_2), V(x) + \pi_{2i}\} & \text{if } x_2 > 0 \\ V(x) + \pi_{2i} & \text{Otherwise} \end{cases}$$

$$T_{3,i}V = \begin{cases} \min \{V(x - e_1), V(x - e_2), V(x) + \pi_{3i}\} & \text{if } x_1 > 0 \text{ and } x_2 > 0 \\ \min \{V(x - e_1), V(x) + \pi_{3i}\} & \text{if } x_1 > 0 \text{ and } x_2 \leq 0 \\ \min \{V(x - e_2), V(x) + \pi_{3i}\} & \text{if } x_2 > 0 \text{ and } x_1 \leq 0 \\ V(x) + \pi_{3i} & \text{Otherwise} \end{cases}$$

$$T_{4,i}V = \min \{c_{1i} + V(x + e_1), V(x)\}$$

$$T_{5,i}V = \min \{c_{2i} + V(x + e_2), V(x)\}$$

$$TV = h(x) + \sum_{i=1}^m \lambda_{1i} T_{1,i}V + \sum_{i=1}^m \lambda_{2i} T_{2,i}V + \sum_{i=1}^m \lambda_{3i} T_{3,i}V + \sum_{i=1}^n \mu_{1i} T_{4,i}V + \sum_{i=1}^n \mu_{2i} T_{5,i}V$$

Then the dynamic programming equation of our problem is as follow.

$$V = TV \tag{6-1}$$

Here, for simplicity, we only consider the case of “horizontal routing” which means the accepted flexible demand should be routed within the same demand class of the two items. Complex shifting among different classes could be formulated if more details for the underlying probability structures are known.

6.2.2 Structures of the Optimal Policy

In this subsection, we will explore the structure of the optimal rationing policy and production/idle policy. Before we proceed on, it is convenience for notations to define two differences of the value functions as follow.

$$\Delta_1 V(x) = V(x + e_1) - V(x)$$

$$\Delta_2 V(x) = V(x + e_2) - V(x)$$

First, define \hat{V} as the set of functions on nonnegative integer set Z_+^2 such that if $V \in \hat{V}$, then:

$$(a) \Delta_1 V(x) \uparrow x_1 \uparrow x_2$$

$$(b) \Delta_2 V(x) \quad \uparrow x_1 \uparrow x_2$$

$$(c) \Delta_1 V(x) - \Delta_2 V(x) \quad \uparrow x_1 \downarrow x_2$$

where \uparrow and \downarrow denote, respectively, nondecreasing and nonincreasing.

Respectively, these three structure properties are the discrete convex, supermodular and monotone difference of the optimal value function. The term monotone difference is from structural property of value function in Chen (2004). In the following, for simplicity, we use increasing and decreasing to indicate nondecreasing and nonincreasing respectively.

Lemma 6.1 If $h(x) \in \hat{V}$ and $V \in \hat{V}$, then $T_{1,i}V, T_{2,i}V, T_{3,i}V, T_{4,i}V, T_{5,i}V$ and $TV \in \hat{V}$.

Proof. See the Appendix. ■

The proof of lemma 6.1 is a case-by-case checking process. Since the proof for various operators defined above are all in the same way, we only choose to verify that $T_{1,i}V \in \hat{V}$. Lemma 6.1 says that the dynamic programming operator T preserves the structure properties (a), (b) and (c) of value function.

To further investigate the structures of the optimal policy, we define the switching functions as follows:

$$S_1(x_2, i) = \min \{x_1 \mid \pi_{1i} + V(x) - V(x - e_1) \geq 0, \text{ given } x_2\}$$

$$S_2(x_1, i) = \min \{x_2 \mid \pi_{2i} + V(x) - V(x - e_2) \geq 0, \text{ given } x_1\}$$

$$S_3(x_1, i) = \min \{x_2 \mid \pi_{3i} + V(x) - V(x - e_1) \geq 0, \text{ or} \\ \pi_{3i} + V(x) - V(x - e_2) \geq 0, \text{ given } x_1\}$$

$$S_4(x_1, i) = \min \{x_2 \mid V(x - e_1) - V(x - e_2) \geq 0, \text{ given } x_1\}$$

$$S_5(x_2, i) = \min \{x_1 \mid c_{1i} + V(x + e_1) - V(x) \geq 0, \text{ given } x_2\}$$

$$S_6(x_1, i) = \min \{x_2 \mid c_{2i} + V(x + e_2) - V(x) \geq 0, \text{ given } x_1\}$$

For S_1 and S_5 , alternatively, it is equivalent to define them as follow.

$$S_1(x_1, i) = \min \{x_2 \mid \pi_{1i} + V(x) - V(x - e_1) \geq 0, \text{ given } x_1\}$$

$$S_5(x_1, i) = \min \{x_2 \mid c_{1i} + V(x + e_1) - V(x) \geq 0, \text{ given } x_1\}$$

Therefore, we use $S_1(x_1, i)$ instead of $S_1(x_2, i)$ in the definition of S_3 . From the above definition, S_1, S_2, S_3 and S_4 are four switching functions associated with rationing/routing controls. The switching functions S_5 and S_6 are used to control the production processes of item 1 and item 2 respectively. Since constants r_i, R_i of class i do not affect the structure properties of the switching functions, to simplify notations, we suppress i in functions $S(\cdot, \cdot)$.

Lemma 6.2: Functions $S_1(x_2)$ and $S_5(x_2)$ are decreasing in x_2 ; $S_2(x_1), S_3(x_1)$ and $S_6(x_1)$ are decreasing in x_1 ; $S_4(x_1)$ is increasing in x_1 .

Proof. See the Appendix. ■

Theorem 6.1. The optimal policy is characterized by six monotone switching curves $S_1(x_2), S_2(x_1), S_3(x_1), S_4(x_1), S_5(x_2)$ and $S_6(x_1)$ according to the following decision rules upon an arrival demand, given the current inventory on two items (x_1, x_2) :

- (i). Type 1 demand of class i will be accepted if and only if $x_1 \geq S_1(x_2)$;
- (ii). Type 2 demand of class i will be accepted if and only if $x_2 \geq S_2(x_1)$;
- (iii). For Type 3 demand of class i , there are the following four cases:
 - (a) satisfy the demand with item 2 if $x_2 \geq S_3(x_1)$ and $x_2 > S_4(x_1)$;
 - (b) satisfy the demand with item 1 if $x_2 \geq S_3(x_1)$ and $x_2 < S_4(x_1)$;
 - (c) satisfy the demand with either item 1 or item 2 if $x_2 \geq S_3(x_1)$ and $x_2 = S_4(x_1)$;
 - (d) reject the demand if $x_2 < S_3(x_1)$.
- (iv). For the production of item 1, stop using production rate μ_{1i} when $x_1 \geq S_5(x_2, i)$, otherwise produce with it.
- (v). For the production of item 2, stop using production rate μ_{2i} when $x_2 \geq S_6(x_1, i)$, otherwise produce with it.

Proof. From lemma 6.2, we know that the monotone curve $S_1(x_2)$ is decreasing in x_2 . It is a switching curve used to decide whether type 1 demand should be satisfied by item 1 or not. On the

right side of $S_1(x_2)$, for any given x_2 , we have $x_1 \geq S_1(x_2)$. The definition of S_1 and the convex of value function implies that $\pi_{1i} + V(x) - V(x - e_1) \geq 0$ whenever $x_1 \geq S_1(x_2)$. Thus, it is optimal to satisfy type 1 demand. Otherwise reject the demand when x_1 is on the left side of the curve $S_1(x_2)$, that is $x_1 < S_1(x_2)$.

Similarly, for fixed x_1 , on the above of $S_2(x_1)$, we always have $\pi_{2i} + V(x) - V(x - e_2) \geq 0$. Thus, it is optimal to accept type 2 demand for item 2. Otherwise, reject type 2 demand.

From figure 6.2, we see that $S_3(x_1)$ is the solid curve decreasing in x_1 . Above $S_3(x_1)$, for fixed x_1 , we have either $\pi_{3i} + V(x) - V(x - e_1) \geq 0$ or $\pi_{3i} + V(x) - V(x - e_2) \geq 0$. Hence, it is optimal to accept the flexible demand. Below $S_3(x_1)$, it is optimal to reject it.

$S_4(x_1)$ is used for routing control of the flexible demand when it is accepted. On the above of $S_4(x_1)$, for any fixed x_1 , we have $V(x - e_1) - V(x - e_2) \geq 0$. This is true from structure property c). Therefore, it is optimal to fulfill the flexible demand with item 2. Below $S_4(x_1)$, it is optimal to satisfy the demand with item 1.

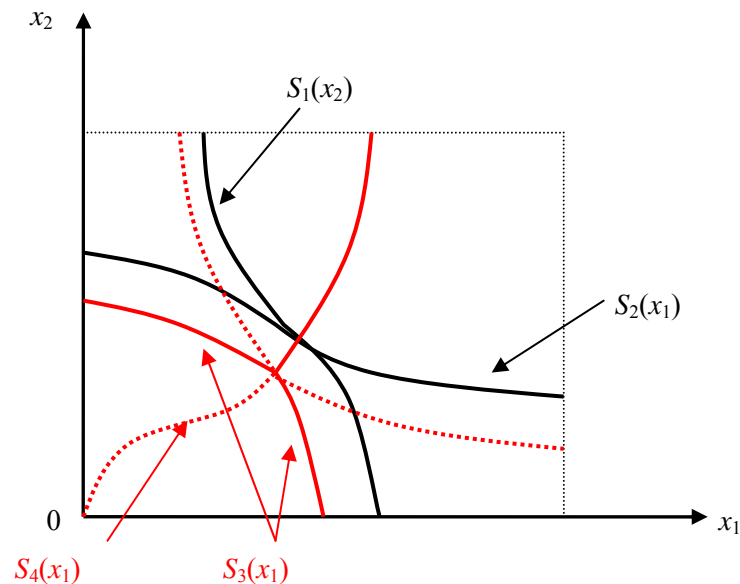


Figure 6.2: Switching curves for admission and routing controls

From lemma 6.2, $S_5(x_2)$ is decreasing in x_2 and it is used to make the production/idle decision of item 1. On the right side of $S_5(x_2, i)$, for fixed x_2 , $x_1 > S_5(x_2, i)$. Then, by property a), we have $c_{1i} + V(x + e_1) - V(x) \geq 0$ on the right of the curve. Hence, it is optimal to stop the production rate μ_{1i} . Otherwise produce item 1 with rate μ_{1i} when $x_1 \leq S_5(x_2, i)$.

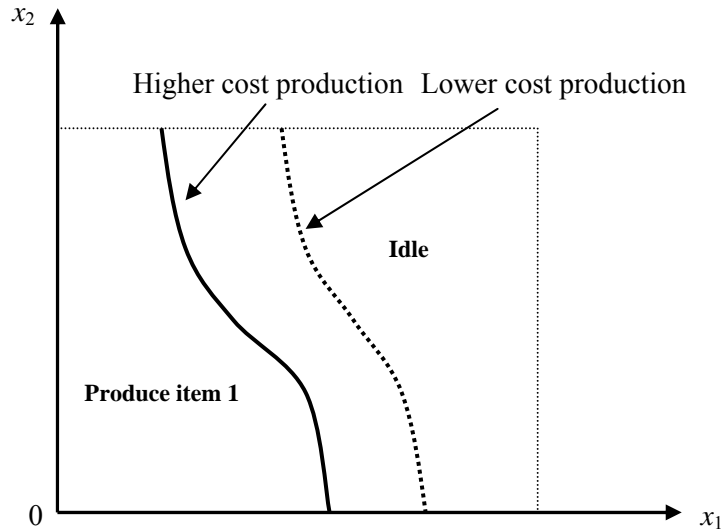


Figure 6.3: The decision curves for production of item 1

For $S_6(x_1, i)$ and fixed x_1 , on the above of the curve, we have $x_2 > S_6(x_1, i)$ and $c_{2i} + V(x + e_2) - V(x) \geq 0$. Thus, it is optimal to stop the production rate μ_{2i} . Otherwise, it is optimal to produce item 2 with the rate μ_{2i} when x_2 is below the curve. ■

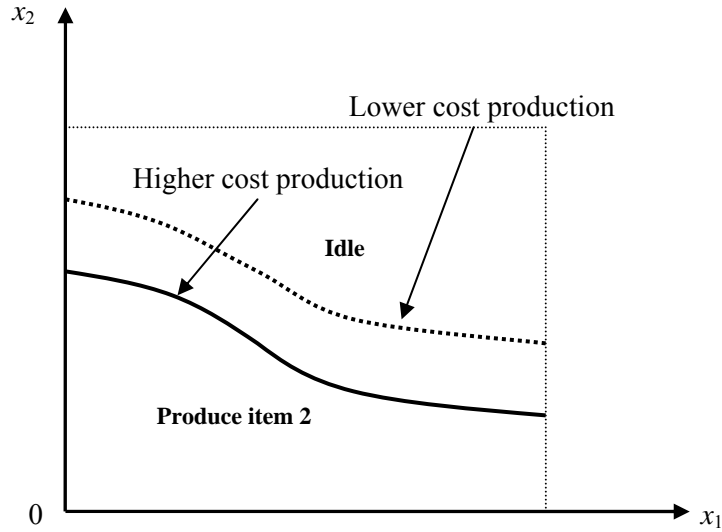


Figure 6.4: The decision curves for production of item 2

Proposition 6.1. Switching curves $S_1(x_2, i)$, $S_2(x_1, i)$ and $S_3(x_1, i)$ are monotone increasing in demand classes; Switching curves $S_5(x_2, i)$ and $S_6(x_1, i)$ are monotone decreasing in the classes of discrete production rate.

Proof: See the Appendix. ■

This proposition tells us that switching curves for m classes of demand are nested. On the other hand, production/idle switching curves are nested in the classes of discrete production rates. Proposition 6.1 is intuitive. Since $V(x+1) - V(x)$ is increasing in one of the argument x_2 , by fixing the other argument of x_1 , it will cross $-\pi_i$ one by one regarding the demand classes as x_2 increases. It also demonstrates that demands in lower classes are more likely to be rejected. The higher the demand class, the more likely it is being accepted. Graphically, rejection regions of high demand classes are smaller than those of the lower demand classes. For production processes, the production rate with lowest cost is firstly used. More production rates with higher cost could be employed when inventory level goes down across certain threshold since the switching curves for production/idle are monotone decreasing in the classes of production rate. Readers may refer to figures 6.3 and figure 6.4.

6.3 The Case of Average Cost Criterion

In the previous sections, we concentrate on the discount-cost criterion. Here, we argue that all structure results derived in the discount-cost criterion can be extended to the average-cost criterion case. Given a policy u , the average cost for our problem is given by:

$$J_u(X_0) = \inf_u \lim_{T \rightarrow \infty} \frac{E_{X_0}^u \left[\int_0^T h(x_1, x_2) dt + \sum_{j=1}^3 \sum_{i=1}^m \int_0^T \pi_{ji} dM_{ji}(t) + \sum_{j=1}^2 \sum_{i=1}^n \int_0^T c_{ji} dN_{ji}(t) \right]}{T}$$

And the objective is to find a optimal control policy u^* yielding value function $V(X_0)$ which is the infimum of $J_u(X_0)$ over all admissible policies $u \in U$

$$V(X_0) = \inf_{u \in U} J_u(X_0)$$

where U is total set of admissible control policies.

Let $V(x)$ be the relative value function of being in state (x_1, x_2) and g denote the optimal average cost per unit time after the rescaling of time. Then, we can write the optimality equation as follow.

$$V + g = TV$$

Where T is operator defined in the previous section. Without rescaling of time, it can be written as

$$V + \frac{g}{\Lambda} = \frac{1}{\Lambda} TV$$

Theorem 6.2. There exists a stationary policy that is optimal under the average cost criterion. The policy retains the structural properties of the optimal policy for the discount-cost criterion.

Proof. Our proof is straightforward. To argue the existence of an optimal policy for the average cost criterion, we follow the proof of the conditions in Weber and Stidham (1987), under which the average cost problem can be regarded as the limit of the discount cost problem by taking $\alpha \rightarrow 0$. First, we need to show that our model is unichain with finite state space and finite action space. According to decision rules in Theorem 6.1, we know that production/idle of an item is characterized by a switching curve or threshold. It is optimal to produce only when inventory levels are below the thresholds. Thus, the state space of the original problem can be confined to a finite state space and action space. As a result, we could have a finite average cost J_u . Secondly, it

can be readily demonstrated that the number of states for which $h(x) \leq J_u$ are finite based on the fact that holding cost $h(x)$ is increasing and component-wise convex. Hence, applying the same argument in Weber and Stidham (1987), the existence of an optimal policy follows naturally. Moreover, it can be shown that there exists g^* and f , such that

$$f + g^* \geq Tf$$

Since the stationary policy that minimizes Tf for each state x is an optimal policy for the average cost criterion and yields a constant g^* , properties of the optimal policy for average-cost criterion through f can be obtained in the same way as those of discount-cost criterion. Hence, optimal policy under average-cost criterion retains all the structures as those of discount-cost criterion. ■

6.4 Numerical Example

Table 6.1: Arrival Rates for Three Types of Demands

	demand for item 1	demand for item 2	flexible demand
High Priority	0.11	0.1	0.248
Low Priority	0.1	0.15	0.21

In the numerical example, for simplicity, we assume that there are only two classes: high priority and low priority classes with $\pi_{11} = 50$, $\pi_{12} = 30$, $\pi_{21} = 50$, $\pi_{22} = 30$, $\pi_{31} = 50$ and $\pi_{32} = 35$. The holding costs in our example are assumed to be separately linear with $h_1 = 3$ and $h_2 = 2$. Let discount rate be $\alpha = 0.002$. The rescaled production rates for the two items are $\mu_{1i} = 0.02$ and $\mu_{2i} = 0.02$ for $i = 1, 2$ indicating the classes. The production costs are $c_{11} = 1$, $c_{12} = 2$, $c_{21} = 1$ and $c_{22} = 2$ respectively. The rescaled demand arrival rates are listed in table 6.1. Since our problem deals with the admission control of the arrival process, the stability condition

$$\sum_{i=1}^n (\mu_{1i} + \mu_{2i}) > \sum_{i=1}^m (\lambda_{1i} + \lambda_{2i} + \lambda_{3i})$$

is not strictly required. In the following figures 6.5, 6.6 and

6.7, we use 0, 1, 2 to denote the rejection, satisfying the demand with item 1 and satisfying the demand with item 2 respectively. In figures 6.8 and 6.9, decisions 0, 1, and 2 represent idle, produce item 1 and produce item 2 respectively.

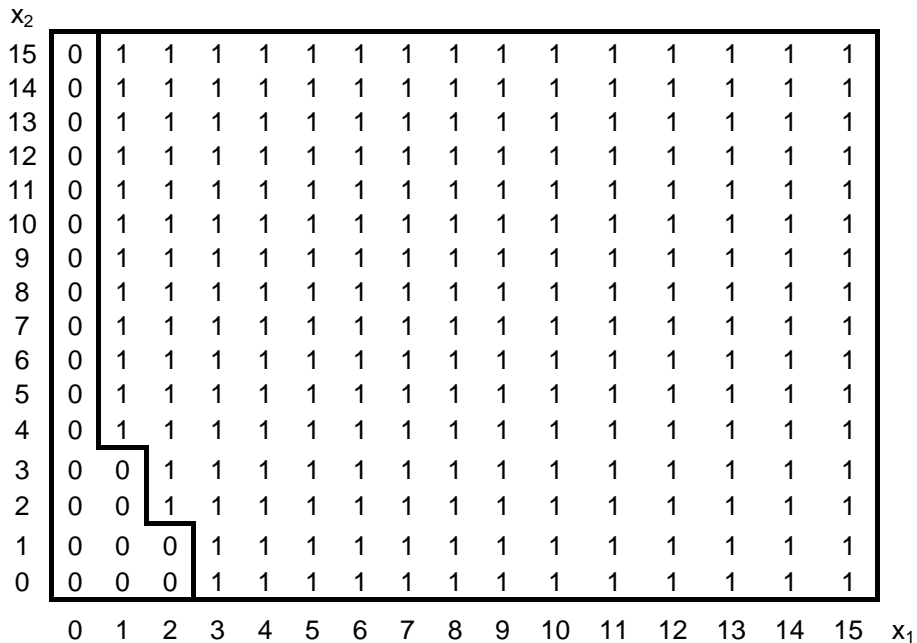


Figure 6.5: Switching curves for Lower-Priority Demand on Item 1.

The above figure illustrates the monotone decreasing curve $S_1(x_2, 2)$ which is used for lower-priority demand on item 1. When type 1 demand arrives, it needs to check the inventory states of both items. If current inventory state lies on the right side of $S_1(x_2, 2)$, which is the decision area with number 1, it is optimal to satisfy the demand with item 1. Otherwise, reject the demand.

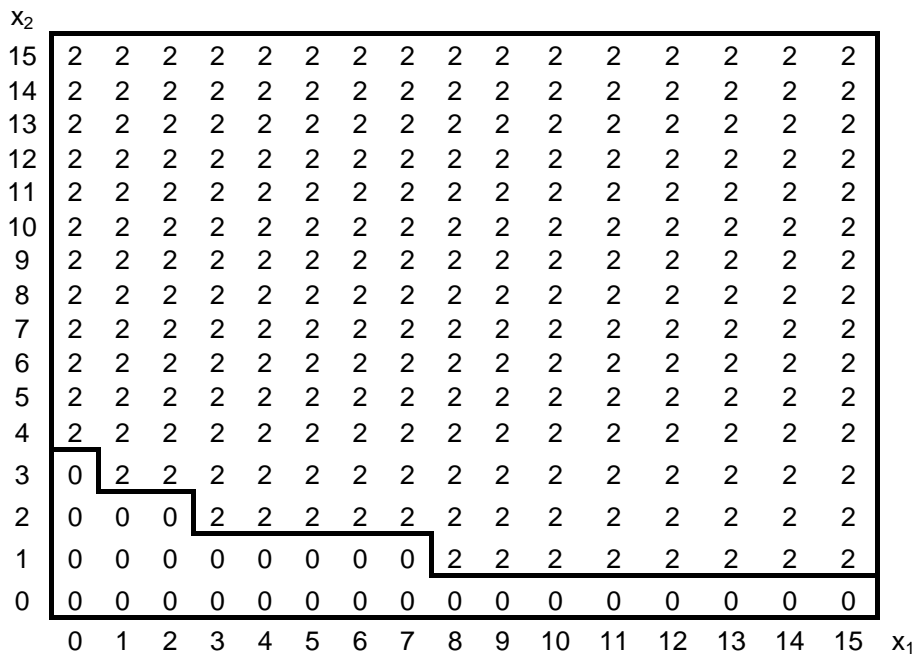


Figure 6.6: Switching curves for Lower-Priority Demand on Item 2.

The above figure illustrates the monotone decreasing curve $S_2(x_1, 2)$ which is used for lower-priority demand on item 2. If current inventory state is above $S_2(x_1, 2)$, which is the area with number 2, it is optimal to satisfy an arrival demand with item 2. Otherwise, reject the demand.

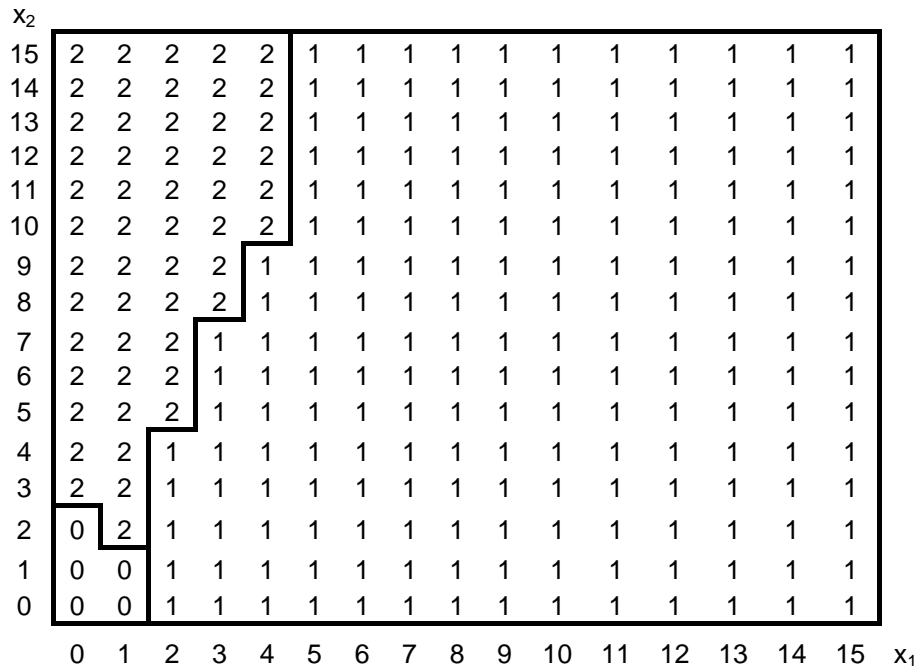


Figure 6.7: Switching curves for flexible Lower-Priority Demand.

The above figure shows the monotone decreasing curve $S_3(x_1, 2)$ and monotone increasing curve $S_4(x_1, 2)$ which are used for the flexible lower-priority demand. If current inventory state is below $S_3(x_1, 2)$, where is the area with decision 0 in the figure, it is optimal to reject the flexible demand. Otherwise, accept the demand. Thereafter, we use $S_4(x_1, 2)$ to dispose the accepted flexible demand. If inventory state lies in the area with number 2 which is just above $S_4(x_1, 2)$, it is optimal to satisfy the demand with item 2 for the reason that inventory of item 2 is relatively higher than that of item 1. Accordingly, when current inventory state is in the area with decision 1, it is optimal to satisfy the demand with item 1 which has relatively higher inventory than that of item 2.

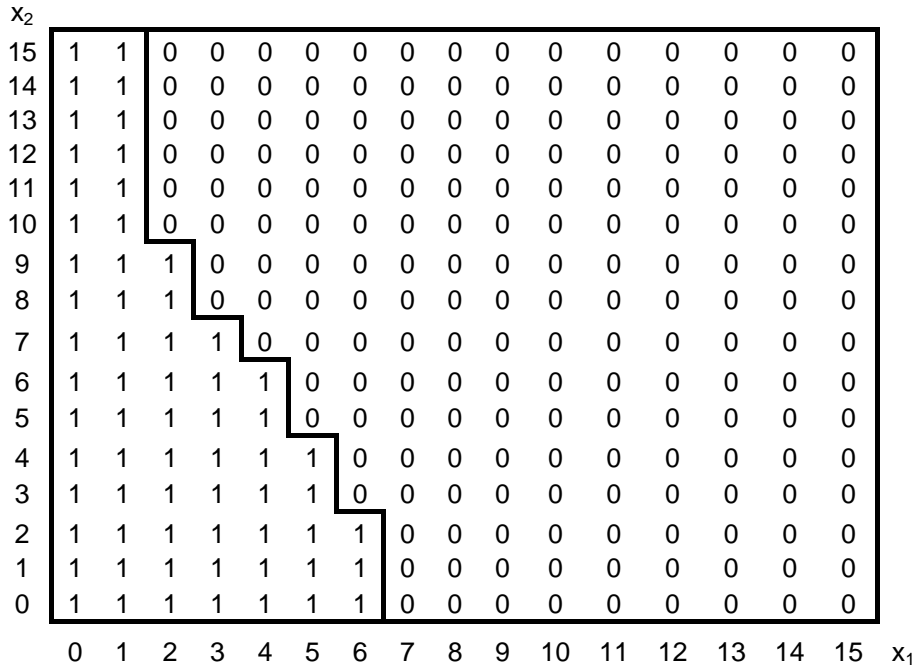


Figure 6.8: Switching Curve for Production of Item 1 with Lower Cost.

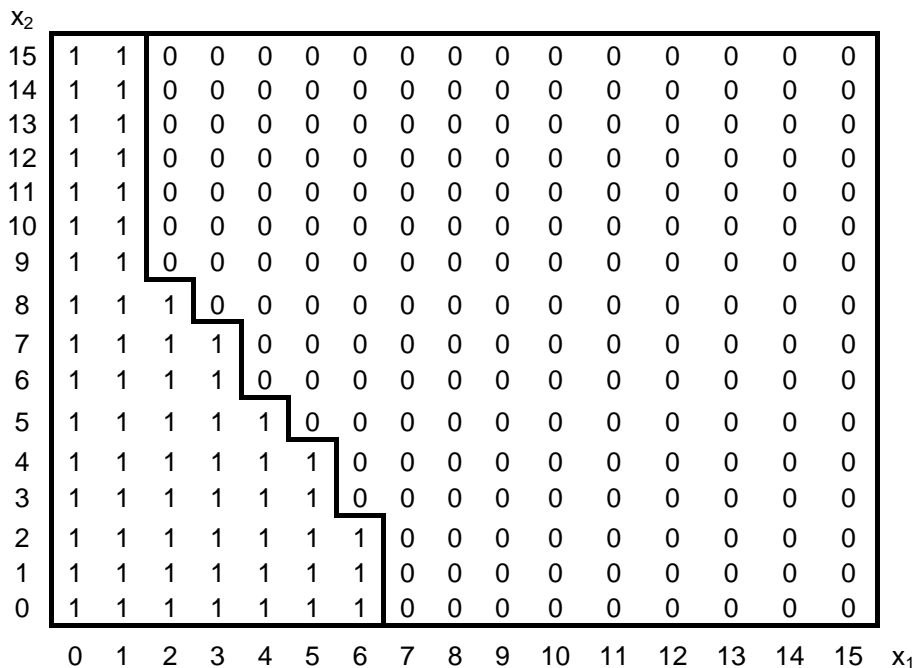


Figure 6.9: Switching Curve for Production of Item 1 with Higher Cost.

Figure 6.8 and Figure 6.9 show the production switching curves for item 1 with lower cost and higher cost, respectively. We can see that the production switching curve is monotone decreasing in production cost. Intuitively, management will resort to more production capacity and working staff to increase the production rate when inventory is relatively lower. On right side of the

switching curve, which is the area with decision 0, e.g., it is optimal to stop the production rate μ_{i_1} of item 1 which has a relatively higher inventory. Otherwise, it is optimal to produce item 1 with the production rate when the inventory state is on the left side of the switching curve.

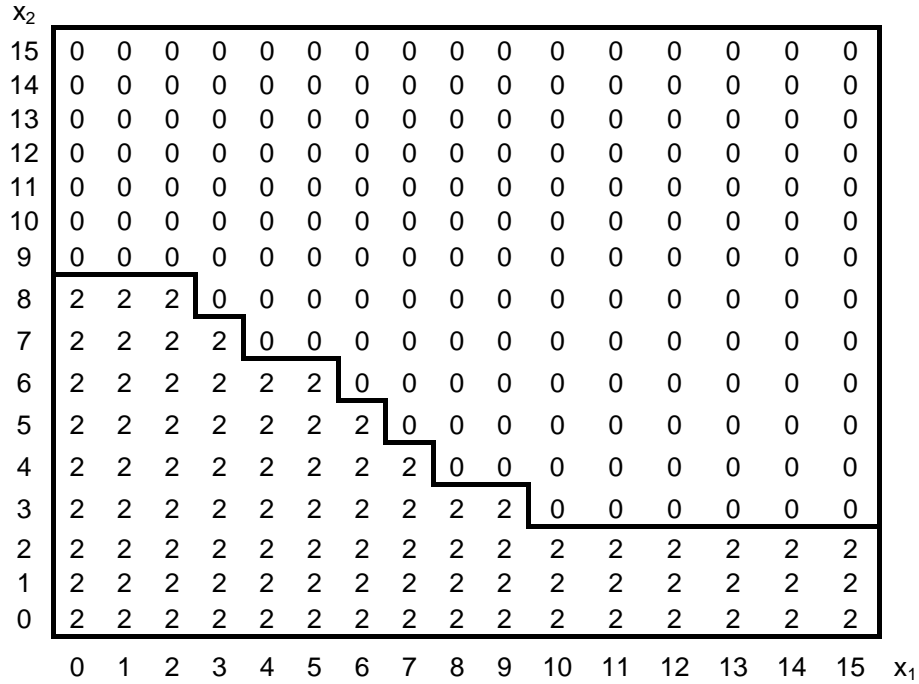


Figure 6.10: Switching Curve for Production of Item 2 with Lower Cost.

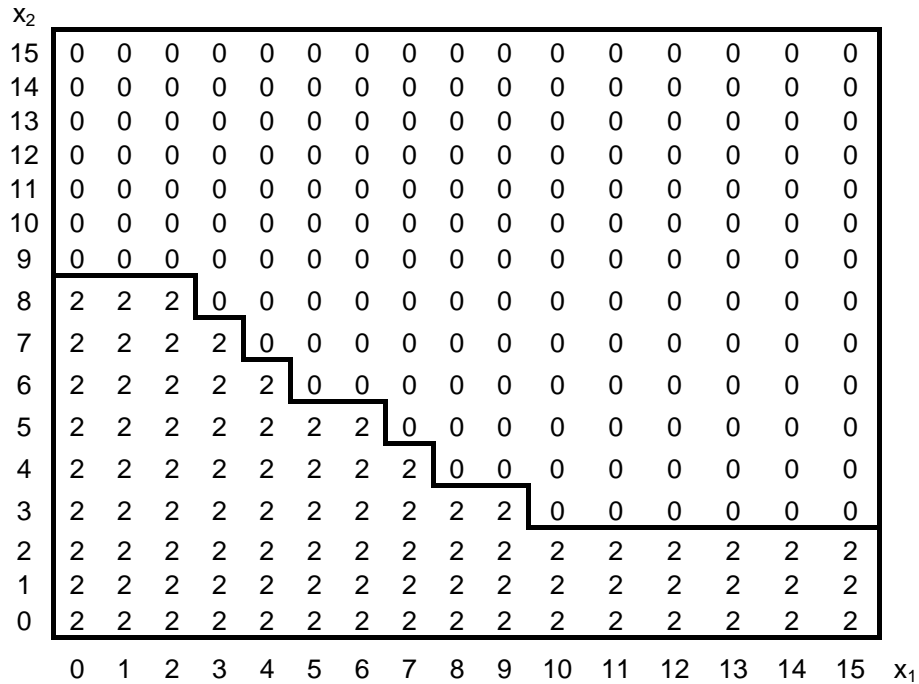


Figure 6.11: Switching Curve for Production of Item 2 with Higher Cost.

Figure 6.10 and Figure 6.11 show the production switching curves for item 2 with lower cost and higher cost respectively. The production switching curve is also monotone decreasing in production cost. It is optimal to produce with rate μ_{2i} when inventory level of both items lies below the switching curve. Otherwise, stop the production rate μ_{2i} of item 2.

Following figure shows the monotone property of the difference $V(x_1+1, x_2) - V(x_1, x_2)$. For any fixed x_2 , it is an increasing curve with respect to x_1 , e.g. convex in x_1 . For any fixed x_1 , the difference is increasing with respect to x_2 , e.g. supermodular in (x_1, x_2) .

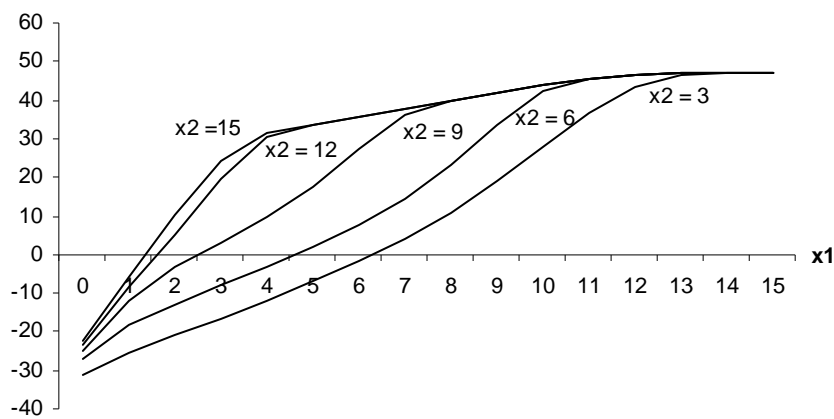


Figure 6.12: Monotone Property of $V(x_1+1, x_2) - V(x_1, x_2)$.

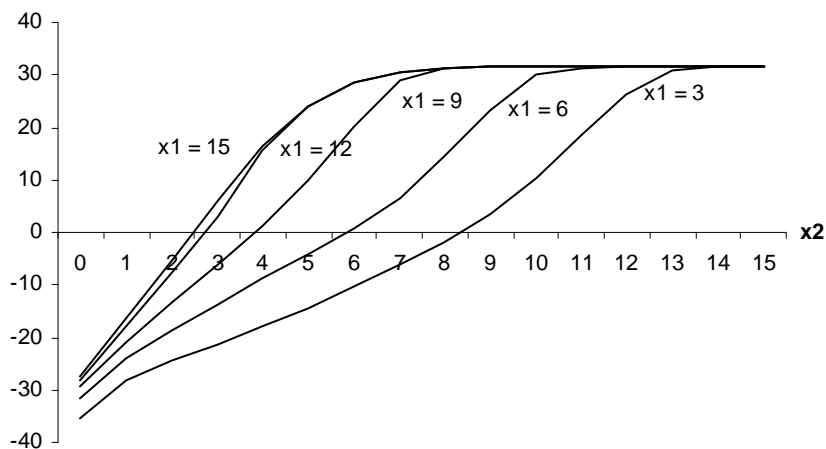


Figure 6.13: Monotone Property of $V(x_1, x_2+1) - V(x_1, x_2)$.

The above figure illustrates the monotone property of $V(x_1, x_2+1) - V(x_1, x_2)$ which is an increasing curve with respect to x_2 , e.g. convex in x_2 for any fixed x_1 . For any fixed x_2 , the difference is increasing with respect to x_1 , e.g. supermodular in (x_1, x_2) .

6.5 Conclusions

In this chapter, an interesting two-item, two production-facility stock rationing problem has been studied. Model formulations follow the make-to-stock queue setting. To derive some intuitive structure results, we concentrate on the analysis of the optimal policies for the lost-sales case. By inductive approach of the dynamic programming, we obtain the structure properties of optimal value function. Furthermore, it is shown that the structure of the optimal policy can be characterized by monotone switching curves, which separates the decision space into several regions. Our theoretical results can be regarded as a two-dimensional extension to those on the single-item, single-production facility problems.

Two-item problem has special attribute that it involve two items substitutable with each other in use for some of the customers. Thus, in recognizing this substitutability, we differentiate the type of customer demand by market segment. Customers who are indifferent between the two items can be singled out and classified as flexible ones. From previous discussion, it can be argued that this flexibility adds value to management. On the other hand, flexible demand may have more opportunities to be satisfied by the suppliers. Many medicines in a pharmaceutical store, for instance, have substitutable properties. More examples in Gallego and Phillips (2004) discuss flexible products which have similar traits of flexible demand. Commonly, the essential idea of rationing is to differentiate the demands that have more or less value to the suppliers. Consideration of flexibility of demand is a further extension of this idea of rationing demand. In addition to the rationing and production decisions, we have included the routing of flexible demand. Both suppliers and customers benefit from this identification of flexible demand.

Finally, our research can be extended in many directions. In the original model, we make the assumption that the arrival rates of the demand are time-homogeneous. In practice, many of the demand arrival processes are non-homogeneous. We believe, after reformulate our continuous-

time model into discrete-time model approaching the non-homogeneous case, the structures of the optimal control policies are anticipated to be preserved. Although this chapter presents the analysis of the lost-sales case, we believe that continuing research on the backorder case in the same scenario as ours may yield more intuitive results. Another direction of research is to remove the assumption of exponential processing time in our model. Current $M/M/2$ setting could be extended to $M/E_k/2$ and $M/G/2$. It can be expected that these extensions are more difficult to analyze. Here, we have only taken into account the continuous-review stock rationing system. There are many papers give intensive discussions on periodic-review systems. The periodic-review production/inventory system with two items is completely characterized by Chen (2004) where optimal policy is also proved to have the switching curve structures. Therefore, we conjecture that periodic-review model of our peers can be analyzed and similar theoretical results might be retained. Furthermore, future research may investigate the multi-item (more than two items) problems. Above all, our work on two-item, two-production facility stock rationing problem is only a prelude to many more general cases.

6.6 Appendix:

Proof of Lemma 6.1

First, we will show that if $V \in \hat{V}$, then $T_{1,i}V \in \hat{V}$. From

$$T_{1,i}V = \begin{cases} \min\{V(x-e_1), V(x)+\pi_{1i}\} & \text{if } x_1 > 0 \\ V(x)+\pi_{1i} & \text{Otherwise} \end{cases}$$

Our result is evident for $x_1 \leq 0$. If $x_1 > 0$, define

$$f(x) = \min\{V(x-e_1), V(x)+\pi_{1i}\}$$

Proof of Property (a): $\Delta_1 f(x) \uparrow x_1 \uparrow x_2$

Note that

$$\begin{aligned} \Delta_1 f(x) &= f(x+e_1) - f(x) \\ &= \min\{V(x), V(x+e_1)+\pi_{1i}\} - \min\{V(x-e_1), V(x)+\pi_{1i}\} \end{aligned} \quad (\text{A1})$$

and

$$\begin{aligned} \Delta_1 f(x+e_1) &= f(x+2e_1) - f(x+e_1) \\ &= \min\{V(x+e_1), V(x+2e_1)+\pi_{1i}\} - \min\{V(x), V(x+e_1)+\pi_{1i}\} \end{aligned} \quad (\text{A2})$$

In order to show that $\Delta_1 f(x) \leq \Delta_1 f(x+e_1)$, i.e., $f(x)$ is convex in x_1 , a simple way is to check all cases.

$$\begin{aligned} &\min\{V(x), V(x+e_1)+\pi_{1i}\} - V(x-e_1) \\ &\leq V(x) - V(x-e_1) \leq V(x+e_1) - V(x) \\ &\leq V(x+e_1) - \min\{V(x), V(x+e_1)+\pi_{1i}\} \end{aligned}$$

and

$$\begin{aligned} &\min\{V(x), V(x+e_1)+\pi_{1i}\} - V(x-e_1) \\ &\leq V(x) - V(x-e_1) \leq V(x+2e_1)+\pi_{1i} - V(x+e_1) - \pi_{1i} \\ &\leq V(x+2e_1)+\pi_{1i} - \min\{V(x), V(x+e_1)+\pi_{1i}\} \end{aligned}$$

Combining the above two inequalities, we derive

$$\begin{aligned} &\min\{V(x), V(x+e_1)+\pi_{1i}\} - V(x-e_1) \\ &\leq \min\{V(x+e_1), V(x+2e_1)+\pi_{1i}\} - \min\{V(x), V(x+e_1)+\pi_{1i}\} \end{aligned} \quad (\text{A3})$$

Similarly, we can obtain

$$\begin{aligned} & \min \{V(x), V(x + e_1) + \pi_{1i}\} - V(x) - \pi_{1i} \\ & \leq \min \{V(x + e_1), V(x + 2e_1) + \pi_{1i}\} - \min \{V(x), V(x + e_1) + \pi_{1i}\} \end{aligned} \quad (\text{A4})$$

Combining (A3) and (A4), we have

$$\begin{aligned} & \min \{V(x), V(x + e_1) + \pi_{1i}\} - \min \{V(x - e_1), V(x) + \pi_{1i}\} \\ & \leq \min \{V(x + e_1), V(x + 2e_1) + \pi_{1i}\} - \min \{V(x), V(x + e_1) + \pi_{1i}\} \end{aligned}$$

We are now proceeding to show that $\Delta_1 f(x) \uparrow x_2$, the supermodular of the value function, i.e.,

$$f(x + e_1) - f(x) \leq f(x + e_1 + e_2) - f(x + e_2)$$

Note that

$$\begin{aligned} & f(x + e_1) - f(x) \\ & = \min \{V(x), V(x + e_1) + \pi_{1i}\} - \min \{V(x - e_1), V(x) + \pi_{1i}\} \end{aligned} \quad (\text{A5})$$

and

$$\begin{aligned} & f(x + e_1 + e_2) - f(x + e_2) \\ & = \min \{V(x + e_2), V(x + e_1 + e_2) + \pi_{1i}\} - \min \{V(x - e_1 + e_2), V(x + e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A6})$$

Since $V \in \hat{V}$, we have

$$\begin{aligned} & \min \{V(x), V(x + e_1) + \pi_{1i}\} - V(x - e_1) \\ & \leq V(x) - V(x - e_1) \leq V(x + e_2) - V(x - e_1 + e_2) \\ & \leq V(x + e_2) - \min \{V(x - e_1 + e_2), V(x + e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A7})$$

and

$$\begin{aligned} & \min \{V(x), V(x + e_1) + \pi_{1i}\} - V(x - e_1) \\ & \leq V(x) - V(x - e_1) \leq V(x + e_1 + e_2) - V(x + e_2) \\ & \leq V(x + e_1 + e_2) + \pi_{1i} - \min \{V(x - e_1 + e_2), V(x + e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A8})$$

Combining (A7) and (A8), we derive

$$\begin{aligned} & \min \{V(x), V(x + e_1) + \pi_{1i}\} - V(x - e_1) \\ & \leq \min \{V(x + e_2), V(x + e_1 + e_2) + \pi_{1i}\} - \min \{V(x - e_1 + e_2), V(x + e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A9})$$

Similarly, we can obtain

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - V(x) - \pi_{1i} \\ & \leq \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} - \min \{V(x-e_1+e_2), V(x+e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A10})$$

Combining (A9) and (A10), we have

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - \min \{V(x-e_1), V(x) + \pi_{1i}\} \\ & \leq \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} - \min \{V(x-e_1+e_2), V(x+e_2) + \pi_{1i}\} \end{aligned}$$

Proof of Property (b): $\Delta_2 f(x) \uparrow x_1 \uparrow x_2$

It is clear that Property (b) is symmetrical to Property (a) with respect to x_1 and x_2 , and hence the result applies.

Proof of Property (c): $\Delta_1 f(x) - \Delta_2 f(x) \uparrow x_1 \downarrow x_2$

First, we prove that $\Delta_1 f(x) - \Delta_2 f(x) \uparrow x_1$, which is equivalent to requiring $f(x+e_1) - f(x+e_2) \leq f(x+2e_1) - f(x+e_1+e_2)$. Note that

$$\begin{aligned} & f(x+e_1) - f(x+e_2) \\ & = \min \{V(x), V(x+e_1) + \pi_{1i}\} - \min \{V(x-e_1+e_2), V(x+e_2) + \pi_{1i}\} \\ \\ & f(x+2e_1) - f(x+e_1+e_2) \\ & = \min \{V(x+e_1), V(x+2e_1) + \pi_{1i}\} - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned}$$

We check different cases. First we have

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - V(x-e_1+e_2) \\ & \leq V(x) - V(x-e_1+e_2) \leq V(x+e_1) - V(x+e_2) \\ & \leq V(x+e_1) - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A11})$$

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - V(x-e_1+e_2) \\ & \leq V(x+e_1) + \pi_{1i} - V(x-e_1+e_2) \\ & \leq V(x+e_1) + \pi_{1i} - V(x+e_2) + V(x+e_2) - V(x-e_1+e_2) \\ & \leq V(x+2e_1) + \pi_{1i} - V(x+e_1+e_2) + V(x+e_1+e_2) - V(x+e_2) \\ & \leq V(x+2e_1) + \pi_{1i} - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A12})$$

Thus, we have

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - V(x-e_1+e_2) \\ & \leq \min \{V(x+e_1), V(x+2e_1) + \pi_{1i}\} - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A13})$$

Similarly, we can obtain

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - V(x+e_2) - \pi_{1i} \\ & \leq \min \{V(x+e_1), V(x+2e_1) + \pi_{1i}\} - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned} \quad (\text{A14})$$

Combining (A13) and (A14), we obtain that

$$\begin{aligned} & \min \{V(x), V(x+e_1) + \pi_{1i}\} - \min \{V(x-e_1+e_2), V(x+e_2) + \pi_{1i}\} \\ & \leq \min \{V(x+e_1), V(x+2e_1) + \pi_{1i}\} - \min \{V(x+e_2), V(x+e_1+e_2) + \pi_{1i}\} \end{aligned}$$

Therefore, $\Delta_1 f(x) - \Delta_2 f(x) \uparrow x_1$. The rest half of Property (c) can be proved in the same way.

Thus, $f(x_1, x_2) = T_{1,i}V \in \hat{V}$, if $V \in \hat{V}$.

The verification for $T_{2,i}V$, $T_{3,i}V$, $T_{4,i}V$ and $T_{5,i}V \in \hat{V}$ if $V \in \hat{V}$ is the same as the above checking process, and hence is thus omitted. Since operator T is a linear and nonnegative combination of $T_{1,i}V$, $T_{2,i}V$, $T_{3,i}V$, $T_{4,i}V$ and $T_{5,i}V$, and thus $TV \in \hat{V}$ if $V \in \hat{V}$. ■

Proof of Lemma 6.2

From definition, given x_2 , $S_1(x_2)$ is the least value of x_1 satisfying

$$\pi_{1i} + V(x) - V(x-e_1) \geq 0$$

The supermodularity of $V(x)$ implies

$$\pi_{1i} + V(S_1(x_2), x_2+1) - V(S_1(x_2)-1, x_2+1) \geq 0 \quad (\text{A-15})$$

By the definition of $S_1(x_2+1)$, we have

$$\pi_{1i} + V(S_1(x_2+1), x_2+1) - V(S_1(x_2+1)-1, x_2+1) \geq 0 \quad (\text{A-16})$$

Comparing (A-15) with (A-16), we can infer that $S_1(x_2) \geq S_1(x_2+1)$ based on the fact that

$S_1(x_2+1)$ is the least value of x_1 satisfying (A-16). Therefore, $S_1(x_2)$ is decreasing in x_2 .

Similarly, we can prove that $S_5(x_2)$ is decreasing in x_2 and $S_2(x_1)$, $S_3(x_1)$ and $S_6(x_1)$ are decreasing in x_1 .

For $S_4(x_1)$, by definition, $V(x_1-1, S_4(x_1)) - V(x_1, S_4(x_1)-1) \geq 0$

From property c)

$$V(x_1 - 2, S_4(x_1)) - V(x_1 - 1, S_4(x_1) - 1) \geq 0. \quad (\text{A-17})$$

By the definition of $S_4(x_1 - 1)$, we have

$$V(x_1 - 2, S_4(x_1 - 1)) - V(x_1 - 1, S_4(x_1 - 1) - 1) \geq 0 \quad (\text{A-18})$$

Comparing (A-17) with (A-18), from the fact that $S_4(x_1 - 1)$ is the least value satisfying (A-18), we can infer that $S_4(x_1) \geq S_4(x_1 - 1)$. Therefore, $S_4(x_1)$ is increasing in x_1 . ■

Proof of Proposition 6.1

First, for the switching curve $S_1(x_2, i) = \min\{x_1 \mid \pi_{1i} + V(x) - V(x - e_1) \geq 0, \text{ given } x_2\}$, we need to prove that $S_1(x_2, i) \leq S_1(x_2, i + 1)$ given fixed x_1 . From the definition, we have

$$\pi_{1i} + V(S_1(x_2, i), x_2) - V(S_1(x_2, i) - 1, x_2) \geq 0 \quad (\text{A-19})$$

$$\pi_{1,i+1} + V(S_1(x_2, i + 1), x_2) - V(S_1(x_2, i + 1) - 1, x_2) \geq 0 \quad (\text{A-20})$$

and $\pi_{1i} > \pi_{1,i+1}$. Here, we use counter proof. If $S_1(x_2, i) > S_1(x_2, i + 1)$, then from (A-19)

$$\pi_{1i} + V(S_1(x_2, i + 1), x_2) - V(S_1(x_2, i + 1) - 1, x_2) < 0 \quad (\text{A-21})$$

then from (A-20) and (A-21), we have

$$\begin{aligned} & \pi_{1i} + V(S_1(x_2, i + 1), x_2) - V(S_1(x_2, i + 1) - 1, x_2) \\ & < \pi_{1,i+1} + V(S_1(x_2, i + 1), x_2) - V(S_1(x_2, i + 1) - 1, x_2) \end{aligned}$$

This leads to $\pi_{1i} < \pi_{1,i+1}$, which contradicts with $\pi_{1i} > \pi_{1,i+1}$. Hence, $S_1(x_2, i) \leq S_1(x_2, i + 1)$.

Results can be derived for $S_2(x_1, i)$, $S_3(x_1, i)$, $S_5(x_2, i)$ and $S_6(x_1, i)$ in the same logic. Then switching curves $S_2(x_1, i)$ and $S_3(x_1, i)$ are monotonically increasing in demand classes.

$S_5(x_2, i)$ and $S_6(x_1, i)$ are monotone decreasing in the classes of discrete production rate. ■

7. Concluding Remarks

In this chapter, we finish our studies on airline seat inventory control and rationing of the make-to-stock production/inventory system with some concluding remarks.

The mathematical formulations for our dynamic control problems are Markov decision processes (MDP) or stochastic dynamic programming, which is a popular analytical model among the researchers of Operations Research and Management Science. Our research concentrates on characterizing the structures of the optimal policies for multi-flight problem, round-trip/multi-leg problem and two-item, two production-facility production/inventory system, respectively.

Firstly, we thoroughly analyze the two-flight problem with a flexible demand segment, and interesting structural results are derived. The main contribution is the proof that structure of the optimal booking policy can be characterized as monotone switching curves. To our knowledge, the only paper in the revenue management literature with a same spirit is Bitran and Mondschein (1995); their Figure 1 (p. 432) portrays a switching curve for a simple two-type room, one-class allocation problem in a hotel context. Their model follows a classic dynamic programming approach that needs to develop good heuristic solution methods due to dimensionality difficulty, which is in the fact the main focus on their study. This problem can also be classified into the categories of flexible product in Gallego and Phillips (2004).

For the round-trip/two-leg problem, the structure of the optimal booking policy is characterized by three switching curves. In addition, structure of the optimal booking policy for multiple-seat booking problem has been completely characterized as switching curves based on the assumption that booking requests can be partially satisfied. Moreover, this thesis investigates the structure properties of the optimal value function for multi-leg problem. Based on the relaxation of booking controls on some legs, we partially characterize the optimal booking policy as switching surfaces. Approximation method has been used to solve our problems.

For rationing of the make-to-stock production/inventory system, we consider two items, two production facilities. Difference between this study and previous research lies in (i) it considers a

flexible demand segment which brings in significant value to the management; (ii) it considers multiple discrete product rates for two production facilities. Management may choose different production rates according to the inventory status. Structures of the optimal control policy for admission/rejection/routing decisions are characterized by three switching curves for each of demand classes. Additionally, we examine relevant monotone properties associated with the optimal value functions.

Today, major airlines may have as many as hundreds of flights available at any time. Real-time control of the seat reservation of so many flights poses great challenge both to academia and carriers. This motivates us to study the multi-flight and multi-leg problems in this dissertation. Although optimal booking policies has been completely characterized for two-flight and two-leg/round-trip problem, it seems to be difficult to extend our results to multiple flights /legs (more than two legs) problems which require cumbersome computation work even though optimal booking policies are completely characterized. In the future, there are still many research works to do for the manipulation of huge volume of demand data. From the side of industrial application, it needs effective heuristic methods to greatly reduce the computation work and more efficient seat reservation systems to handle large scale problems.

In the research of rationing of make-to-stock production/inventory system, single-item problems have been fairly-well studied. However, for two-item problem, there are not many works on such a topic. Here, we only consider the lost-sales case. It is expected that the present results could be extended to the backlog case. For problems with more than two items, complete characterization of the structure of the optimal policy seems to be difficult. Instead, heuristics and approximation methods might be developed to solve the problems.

One of the special characteristics in our research is to differentiate the customer demands and classify them as dedicated customers and flexible customers. Dedicated customers can be only satisfied with their favorite carriers/flights/items while the flexible customers are indifferent among the carriers/flights/items. In this dissertation, it is demonstrated that identification and management of the flexible demand segment can bring in relative higher value than those without

consideration of the flexibility of customers. Such a topic is quite interesting and has not been thoroughly addressed in other research works. This is the initiative motivation for this dissertation to give an in-depth analysis of customer demand.

References

- BELOBABA, P.P. 1989. Application of a Probabilistic Decision Model to Airline Seat Inventory control. *Operations Research*. **37**, 183-197.
- BENJAAFAR S. AND M. ELHAFSI. 2006. Production and Inventory Control of a Single Product Assemble-To-Order System with Multiple Customer Classes. *Management Science*. **52**, 1896-1912.
- BERTSIMAS D. AND I. POPESCU. 2003. Revenue Management in a Dynamic Network Environment *Transportation Science*. **37**, 257-277.
- BERTSIMAS D. AND R. SHIODA. 2003. Restaurant Revenue Management. *Operations Research*. **51**, 472-486.
- BREMAUD, P. 1980. Point Processes and Queues, Martingale Dynamics. Springer-Verlag, New York.
- BITRAN, G.R. AND S.V. MONDSCHHEIN. 1995. Application of Yield Management to Hotel Industry Considering Multiple Day Stays. *Operations Research*. **43**, 427-443.
- BITRAN, G.R. AND S.M. GILBERT. 1996. Managing Hotel Reservations with Uncertain Arrivals. *Operations Research*. **44**, 35-49.
- BRUMELLE, S.L. AND J.I. MCGILL. 1993. Airline Seat Allocation with Multiple Nested Fare Classes. *Operations Research*. **41**, 127-137.
- BRUMELLE, S.L. AND D. WALCZAK. 2003. Dynamic Airline Revenue Management with Multiple Semi-Markov Demand. *Operations Research*. **51**, 137-148
- CALDENTEY, R. AND L. WEIN. 2002. Revenue Management of a Make-To-Stock Queue. *Manufacturing and Service Operations Management* Vol.4, No.1, 4-6.
- CALDENTEY, R. AND L. WEIN. 2005. Revenue Management of a Make-To-Stock Queue. Working Paper, Stern Business School, New York University, New York.
- CHATWIN, R.E. 1999. Continuous-Time Airline Overbooking with Time-Dependent Fares and Refunds. *Transportation Science*. **33**, 192-206.
- CHEN, S. 2004. The Optimality of Hedging Point Policies for Stochastic Two-Product Flexible Manufacturing Systems. *Operations Research*. **52**, 312-322.
- COHEN, M.A., P.R. KLEINDORFER AND H.L. LEE. 1988. Service Constrained (s, S) Inventory Systems with Priority Demand Classes and Lost Sales. *Management Science*. **34**, 482-499.
- CURRY, R.E. 1990. Optimal Airline Seat Allocation with Fare Classes Nested By Origins and Destinations. *Transportation Science*. **24**, 193-204.
- DANA, J.D. 1999. Using Yield Management to Shift Demand When the Peak Time is Unknown. *RAND Journal of Economics*. **30**, 456-474.

- DE VERICOURT, F. F. KARAESMEN AND Y. DALLERY 2002. Optimal Stock Allocation for Capacitated Supply System. *Management Science*. **48**, 1486-1500.
- DESHPANDE, V., COHEN, M.A. AND K. DONOHUE 2003. A Threshold Inventory Rationing Policy for Service-Differentiated Demand Classes. *Management Science*. **49**, 683-703.
- FENG, Y. AND G. GALLEGRO. 1995. Optimal Stopping Times for End of Season Sales and Optimal Stopping Times for Promotional Fares. *Management Science*. **41**, 1371-1391.
- FENG, Y. AND P. LIN. 2001. Monotone Properties of Optimal Booking Control on a Multi-Leg Flight. Working Paper. National University of Singapore, Singapore.
- FENG, Y. AND B. XIAO. 2000. Optimal Policies of Yield Management with Multiple Predetermined Prices. *Operations Research*. **48**, 332-343.
- FENG, Y. AND B. XIAO. 2001. A Dynamic Airline Seat Inventory Control Model and Its Optimal Policy. *Operations Research*. **49**, 938-949.
- FRANK, K.C. , R.C. ZHANG AND I. DUENYAS. 2003. Optimal Inventory Policy for Inventory System with Priority Demand Classes. *Operations Research*. **51**, 993-1002.
- GALLEGRO, G. AND R. PHILLIPS. 2004. Revenue Management of Flexible Products. *Manufacturing and Service Operations Management* **6**, 321-337.
- GALLEGRO, G. AND G.J. VAN RYZIN. 1994. Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons. *Management Science*. **40**, 999-1020.
- GALLEGRO, G. AND G.J. VAN RYZIN. 1997. A Multi-Product Dynamic Pricing Problem and Its Applications to Network Yield Management. *Operations Research*. **45**, 24-41.
- GERCHAK, Y. M. PARLAR AND T. YEE. 1985. Optimal Rationing Policies and Production Quantities for Products with Several Demand Classes. *Canadian Journal of Administrative Sciences*. **1**, 161-176.
- GLASSERMAN, P. AND D.D. YAO. 1994. Monotone Structure in Discrete-Event Systems. New York: Wiley.
- HA, A.Y. 1997a. Inventory Rationing in a Make-To-Stock Production System with Several Demand Classes and Lost Sales. *Management Science*. **43**, 1093-1103.
- HA, A.Y. 1997b. Stock Rationing Policy for a Make-To-Stock Production System with Two Priority Classes and backordering. *Naval Research Logistics* **44**, 457-472.
- HA, A.Y. 1997c. Optimal Dynamic Scheduling Policy for a Make-To-Stock Production. *Operations Research*. **45**, 42-53.
- HA, A.Y. 2000. Stock Rationing in an $M/E_k/1$ Make-To-Stock Queue. *Management Science*. **46**, 77-87.
- IYEEY, A. V., V. DESHPANDE AND Z. WU. 2003. A Postponement Model for Demand Management. *Management Science*. **49**, 983-1002.

- HAJEK, B. 1984. Optimal Control of Two Interacting Service Stations. *IEEE Transactions on Automatic Control*. **29**, 491-498.
- HERSH, M. AND S.P. LADANY 1978. Optimal Seat Allocation for Flights with Intermediate Stop. *Computers and Operations Research*. **5**, 31-37.
- KLEYWEGT, A.J. AND J.D. PAPASTAVROU. 1998. The Dynamic and Stochastic Knapsack Problem. *Operations Research*. **46**, 17-35.
- KLEYWEGT, A.J. AND J.D. PAPASTAVROU. 2001. The Dynamic and Stochastic Knapsack Problem with Random Sized Items. *Operations Research*. **46**, 17-35.
- KOOLE. G.M. Structure Results for the Control of Queueing Systems Using Event-Based Dynamic Programming, *Queueing Systems*. **30**(1998), 323-339.
- LADANY, S.P. AND D.N. BEDI 1977. Dynamic Booking Rules For Flights with Intermediate Stops. *OMEGA* **5**, 721-730.
- LAUTENBACHER, C.J. AND S. STIDHAM. 1999. The Underlying Markov Decision Process in the Single-Leg Airline Yield Management Problem. *Transportation Science*. **33**, 136-146.
- LEE T. C. AND M. HERSH. 1993. A Model of Dynamic Airline Seat Inventory Control with Multiple Seat Bookings. *Transportation Science*. **27**, 252-265.
- LI, M.Z.F. AND T.H. OUM. 2002. A Note on the Single Leg, Multifare Seat Allocation Problem. *Transportation Science*. **36**, 349-353.
- LIANG, Y. 1999. Solution to the Continuous Time Dynamic Yield Management Model. *Transportation Science*. **33**, 117-123.
- LIN, K.Y. 2004. A Sequential Dynamic Pricing Model and Its Applications. *Naval Research Logistics*. **51**, 501-621
- LIPPMAN, S. 1975. Applying a New Device in the Optimization of Exponential Queueing Systems. *Operations Research*. **23**, 687-710.
- LITTLEWOOD, K. 1972. Forecasting and Control of Passenger Bookings. *AGIFORS Symp. Proc.* **12**, 103-105.
- MAGLARAS, C. 2003. Revenue Management for a Multi-Class Single-Server Queue. Working Paper, Dept. of Decision, Risk & Operations, Columbia University, NY
- MCGILL, J.I. AND G.J. VAN RYZIN. 1999. Revenue Management: Research Overview and Prospects. *Transportation Science*. **33**, 233-256.
- NAHMIA, S. AND W.S. DEMMY. 1981. Operating Characteristics of an Inventory System with Rationing. *Management Science*. **27**, 1236-1245.
- ROTHSTEIN, M. 1971. An Airline Overbooking Model. *Transportation Science*. **5**, 180-192
- ROBINSON, L.W. 1995. Optimal and Approximate Control Policies for Airline Booking with Sequential Non-Monotonic Fare Classes. *Operations Research*. **43**, 252-263.

- SONG D. AND Y. SUN. 1999. Optimal Control Structure of an Unreliable Manufacturing System with Random Demands, *IEEE Transactions on Automatic Control*. **44**, 619-622.
- STIDHAM S. AND R. WEBER. 1993. A Survey of Markov Decision Models for Control of Networks of Queues. *Queueing Systems*. **13**, 291-314.
- TALLURI, K. 2001. Airline Revenue Management with Passenger Routing Control: a New Model with Solution Approaches. *International Journal of Services Technology and Management*. **2**. 102-115.
- TALLURI, K. AND G.J. VAN RYZIN. 1999. An Analysis of Bid-Price Controls for Network Revenue Management. *Management Science*. **44**, 1577-1593.
- TALLURI, K. AND G.J. VAN RYZIN. 2004. Revenue Management under a General Discrete Choice Model of Customer Behavior. *Management Science*. **50**, 15-33.
- TALLURI, K. AND G.J. VAN RYZIN. 2004. The Theory and Practice of Revenue Management. Kluwer, New York.
- TOPKIS, D.M. 1968. Optimal Ordering and Rating Policies in a Nonstationary Dynamic Inventory Model with n Demand Classes. *Management Science*. **15**, 160-176.
- TOPKIS, D.M. 1978. Minimizing a submodular function on a lattice. *Operations Research*. **26**, 305-321.
- TOPKIS, D.M. 1998. Submodularity and Complementarity. Princeton University Press.
- VAN SLYKE, R. AND Y. YUNG 2000. Finite Horizon Stochastic Knapsacks with Applications to Yield Management. *Operations Research*. **48**, 155-172.
- VAN RYZIN, G. AND Q. LIU 2004. On the Choice-Based Linear Programming Model for Network Revenue Management. Working Paper. Columbia Business School, Columbia University, New York.
- VEINOTT, A.F. 1965. Optimal Policy in Dynamic, Single Product Nonstationary Inventory Model with Several Demand Classes. *Operations Research*. **13**, 761-778.
- WEBER, R.R. AND S. STIDHAM 1987. Optimal Control of Service Rates in Networks of Queues. *Advances in Applied Probability*. **19**, 202-218.
- WOLLMER, R.D. 1986. A Hub-Spoke Seat Management Model. McDonnell Douglas Corporation, Long Beach, CA
- WOLLMER, R.D. 1992. An Airline Seat Management Model for Single Leg Route When Low Fare Classed Book First. *Operations Research*. **40**, 26-37.
- WONG, J.T. F. KOPPELMAN AND M.S. DASKIN. 1993. Flexible Assignment Approach to Itinerary Seat Allocation. *Transportation Research. B*. **27B**, 33-48.
- YOU, P.S. 1999. Dynamic Pricing in Airline Seat Management for Flights with Multiple Flight Legs. *Transportation Science*. **33**, 192-206.

- ZHANG, D. AND W.L., COOPER. 2005. Revenue Management for Parallel Flights with Customer Choice Behavior. *Operations Research*. **53**, 415-431.
- ZHAO, H. V. DESHPANDE AND J.K. RYAN 2005. Inventory Sharing and Rationing in Decentralized Dealer Networks. *Management Science*. **51**, 531-547.
- ZHAO W. AND Y.S. ZHENG. 2000. Optimal Dynamic Pricing for Perishable Assets with Non-Homogeneous Demand. *Management Science*. **46**, 375-388.
- ZHAO W. AND Y.S. ZHENG. 2001. A Dynamic Model for Airline Seat Allocation with Passenger Diversion and No-Shows. *Transportation Science*. **35**, 80-98.