597 LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS

Christiane Barz

Risk-Averse Capacity Control in Revenue Management

Springer

Lecture Notes in Economics
and Mathematical Systems 597

Christiane Barz

Risk-Averse Capacity Control in Revenue Management

With 32 Figures and 10 Tables

Springer

Christiane Barz
Universität Karlsruhe (TH)
Fakultät für Wirtschaftswissenschaften
Kaiserstraße 12
76128 Karlsruhe
Germany
barz@wior.uni-karlsruhe.de

Dissertation, genehmigt von der Fakultät für Wirtschaftswissenschaften der Universität Fridericiana zu Karlsruhe, gefördert durch die DFG. Referent: Prof. Dr. Karl-Heinz Waldmann; Korreferentin: Prof. Dr. Marliese Uhrig-Homburg; Tag der mündlichen Prüfung: 19.12.2006.

Library of Congress Control Number: 2007930764

ISSN 0075-8442

ISBN 978-3-540-73013-2 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover-design: WMX Design GmbH, Heidelberg

SPIN 12075102   88/3180YL - 5 4 3 2 1 0   Printed on acid-free paper

Acknowledgements

"A fool ... is a man who never tried an experiment in his life."
(Erasmus Darwin)

I would like to thank all of those who supported me during my biggest experiment so far, the completion of this thesis.

In particular, I would like to thank my supervisor, Prof. Dr. Karl-Heinz Waldmann. Without his help, this work would not have been possible. I also thank my co-advisor, Prof. Dr. Marliese Uhrig-Homburg, for her valuable and constructive comments on my work and beyond.

I am deeply indebted to Dr. Alfred Müller. His stimulating suggestions and unflinching encouragement were invaluable. I thank him and all the other colleagues at the chair for their support.

This work was funded by the DFG Graduate School for Information Management and Market Engineering. Many thanks go to all members of this program.

Words fail me to express my appreciation to Thorsten Friedrich for his patience, love, and persistent confidence in me.

Finally, I thank my parents and my brother for their faith and constant support.

Christiane Barz Karlsruhe, June 2007

Contents

1 Introduction 1
1.1 The Basic Capacity Control Problem 4
1.1.1 Assumptions 4
1.1.2 The Attitude Towards Risk in Capacity Control 8
1.2 Risk-Aversion in Revenue Management 10
1.3 Chapter Organization 11

Part I Basic Principles

2 Markov Decision Processes and the Total Reward Criterion 19
2.1 Finite Horizon Markov Decision Processes 19
2.2 Infinite Horizon Markov Decision Processes 22
2.2.1 Markov Decision Processes with an Absorbing Set 23
2.2.2 Markov Decision Processes in a Random Environment . 25

3 Expected Utility Theory for Sequential Decision Making . . 29
3.1 Static Decision Problems 29
3.1.1 The Concept of Risk-Aversion 31
3.1.2 Special Utility Functions 32
3.2 Sequential Decision Problems 35
3.2.1 Additive Time-Separable Utility Functions 36
3.2.2 General Atemporal Utility Functions 38
3.2.3 Atemporal Exponential Utility Functions 40

Part II Expected Revenue Maximizing Capacity Control

4 Capacity Control in a Random Environment 47
4.1 The Decision Model 48
4.1.1 The Environmental Process and the Arrival of Requests 48

4.1.2 Accepting Requests ... 49
4.1.3 Cancelations, No-Shows, and Overbooking ... 49
4.1.4 The Underlying Markov Decision Process ... 51
4.2 Structural Results ... 52
4.3 Examples for the Random Environment ... 55
4.3.1 Exogenous Effects Based on the Evolution of a Markov Chain ... 55
4.3.2 Capacity Control Under Uncertainty ... 55
4.3.3 General Demand Patterns ... 56
4.4 Numerical Examples ... 57
4.4.1 Example 1: Capacity Control with Exogenous Effects ... 57
4.4.2 Example 2: Capacity Control with Dependent Demand ... 59

5 Basic Single Resource Capacity Control Models in Revenue Management ... 63
5.1 The Dynamic Model ... 63
5.1.1 The Decision Model ... 64
5.1.2 Structural Results ... 65
5.1.3 A Numerical Example ... 69
5.2 The Static Model ... 70
5.2.1 The Decision Model ... 71
5.2.2 Structural Results ... 72
5.2.3 A Numerical Example ... 74
5.2.4 The EMSR Heuristics ... 75

Part III Expected Utility Maximizing Capacity Control

6 Capacity Control Maximizing Additive Time-Separable Utility ... 83
6.1 The Dynamic Model ... 83
6.1.1 Structural Results ... 84
6.1.2 Numerical Examples ... 86
6.2 The Static Model ... 88
6.2.1 Structural Results ... 89
6.2.2 Numerical Examples ... 93

7 Capacity Control Maximizing Atemporal Utility ... 95
7.1 The Dynamic Model ... 96
7.1.1 Structural Results ... 96
7.1.2 Numerical Examples ... 97
7.1.3 Structural Results in the Case of an Exponential Utility Function ... 100
7.1.4 Numerical Example in the Case of an Exponential Utility Function ... 105

7.2 The Static Model 106
7.2.1 Numerical Examples 107
7.2.2 Structural Results in the Case of an Exponential Utility Function 109
7.2.3 Numerical Example in the Case of an Exponential Utility Function 114
7.2.4 Possible Extensions of the EMSR Heuristics 118

8 An Extension: Capacity Control Under a General Discrete Choice Model of Consumer Behavior 125
8.1 The Capacity Control Model 125
8.2 Maximizing Expected Revenue 126
8.2.1 Efficient Sets 127
8.2.2 Structure of an Optimal Policy 128
8.2.3 A Numerical Example 129
8.3 Maximizing Expected Utility 131
8.3.1 Efficient Sets 133
8.3.2 Characterization of the Optimal Policy 133
8.3.3 A Numerical Example 136

9 Conclusion 139
9.1 Summary 139
9.2 Directions for Future Work 141

List of Symbols 143

An Extension of Stidham's Lemma 147

Stochastic Concavity and the Binomial Distribution 149

References 151

List of Figures

3.1 Exponential utility function for different values of γ. 33

4.1 Transition graph of the external Markov chain from Example 1. 58
4.2 Transition graph of the external Markov chain from Example 2. 60

5.1 Protection levels of an optimal policy in the case of a risk-neutral decision-maker. 69

6.1 Protection levels $y_i^{\text{add}*}(n)$ of an add-optimal policy given exponential one-stage utility functions with $\gamma = 0.002$. 87
6.2 Protection levels $y_i^{\text{add}*}(n)$ of an optimal policy given exponential one-stage utility functions with $\gamma = 0.01$. 88
6.3 Capacity dependent controls for classes 2 and 3 of an add-optimal policy given exponential one-stage utility functions with $\gamma = 0.05$. 94
6.4 Controls of an add-optimal policy for classes 1, 2 and 3 given 85 remaining seats and one-stage exponential utility functions with $\gamma = 0.0001, 0.0002, 0.0005$ and 0.001. 94

7.1 Protection levels of an atmp-optimal policy given a decision-maker with logarithmic utility function and current wealth $w = 0$. 99
7.2 Protection levels of an atmp-optimal policy given a decision-maker with logarithmic utility function and current wealth $w = 500$. 99
7.3 Protection levels of a γ-optimal policy with $\gamma = 0.001$. 106
7.4 Protection levels of a γ-optimal policy with $\gamma = 0.005$. 106
7.5 Atmp-optimal controls given a decision-maker with logarithmic utility function at current wealth $w = 0$. 109
7.6 Atmp-optimal controls given a decision-maker with logarithmic utility function at current wealth $w = 50000$. 109

7.7 Protection levels of a γ-optimal policy. 114
7.8 Simulated average load factor given a γ-optimal policy. 115
7.9 Expected total revenue when applying a γ-optimal policy. 115
7.10 Simulated standard deviation when applying a γ-optimal policy. 116
7.11 Expected total revenue vs. standard deviation of total revenue when applying different γ-optimal policies. 116
7.12 Relative frequencies of total revenue for 500000 simulation runs when applying an optimal (risk-neutral) policy and applying a γ-optimal policy with $\gamma = 0.0001$, 0.0004, and 0.01. . 117
7.13 Expected total revenue and standard deviation of total revenue given policies obtained by MSCE-a and MSCE-b for different values of γ. 121
7.14 Expected total revenue and standard deviation of total revenue given policies obtained by EMSU and γ-optimal policies. 123
7.15 Expected total revenue and standard deviation of total revenue given a γ-optimal policy and policies obtained by MSCE-a, MSCE-b, MSCE-b', and EMSU. 124

8.1 Scatter plot of λ and ω as well as the efficient sets. 130
8.2 Optimal policy in the case of a risk-neutral decision-maker given the example data for $N = 40$ and $C = 20$. 130
8.3 Scatter plot of λ, $\omega^{0.001}$, and $\omega^{0.01}$ as well as the efficient frontiers. 137
8.4 γ-optimal policy with $\gamma = 0.001$. 138
8.5 γ-optimal policy with $\gamma = 0.01$. 138

List of Tables

4.1 Protection levels $y^*(d, i, (w, t))$ of an optimal policy without (and including) cancelations. 59
4.2 Protection levels $y^*(d, i, 1)$ of an optimal policy. 61
4.3 Protection levels $y^*(d, i, 1)$ of an optimal policy assuming an emphasized effect of low-value demand. 62

5.1 Request probabilities $\hat{p}_n(i)$. 69
5.2 Expectations and standard deviations of the number of requests. 75
5.3 Protection levels and expected revenue from solving the exact model, EMSR-a, and EMSR-b. 78

7.1 Protection levels $y_1^{\gamma a}$, $y_2^{\gamma a}$ and $y_3^{\gamma a}$ determined by MSCE-a. 120
7.2 Protection levels $y_1^{\gamma b}$, $y_2^{\gamma b}$ and $y_3^{\gamma b}$ determined by MSCE-b. 121

8.1 Choice probabilities $\hat{p}_i(\mathfrak{M})$, total probability of purchase $\lambda(\mathfrak{M})$, and expected one-stage revenues $\omega(\mathfrak{M})$. 130
8.2 Choice probabilities $\hat{p}_i(\mathfrak{M})$, total probability of purchase $\lambda(\mathfrak{M})$, and expected one-stage utilities $\omega^\gamma(\mathfrak{M})$ with $\gamma = 0.001$ and $\gamma = 0.01$. 137

1

Introduction

"If necessity is the mother of invention, then deregulation is the father, and revenue management (also known as yield management) is the couple's golden child – at least as far as operations research is concerned." (Horner, 2000, p. 47)

Deregulation had a significant impact on the U.S. airline industry in the late 1970s. Charter and low-cost airlines such as People Express and Southwest were able to offer seats at a fraction of the price charged by established carriers like Pan Am and American Airlines. Due to their different cost structure, it seemed to be impossible for the big carriers to offer tickets at the same low price. Yet they had to find a way to compete.

Robert L. Crandall from American Airlines is widely credited with the solution to the problem: *yield management* – today called *revenue management*, since it maximizes revenue earned on a flight rather than yield (revenue per passenger mile).

The idea was simple: American Airlines flights were only half full on average. Offering the empty seats at a discount price would not only enable the carriers to compete with the low-cost airlines but even create additional revenue, if (1) it were possible to prevent cannibalization, i.e. the sale of discount tickets to consumers who would otherwise be willing to pay full fare, and if (2) it could be assured that only the seats that would otherwise fly empty were sold at the low price.

Implementing this strategy, American Airlines matched the low-cost airlines' prices with a limited number of seats that had to be booked several weeks or months in advance. Due to this purchase restriction – the lack of flexibility – the offer was not attractive for the late-arriving demand (typically business travelers) willing to pay the full fare.

Note, however, that if too many seats were sold at low prices, the airline would run the risk of filling the plane too early and losing full-fare customers (risk of revenue dilution). On the other hand, they were taking the chance that discount demand would be rejected, with full-fare demand not sufficing to fill the airplane. The plane would then depart with more empty seats than necessary (demand spoilage).

The ability of American Airlines to control the availability of discount seats had a dramatic effect on its low-cost competitors. People Express was hit especially hard. For details on the "battle" of American Airlines versus People

Express, see Cross (1998, Chap. 4). Revenue management contributed not only to the bankruptcy of People Express but to the demise of several other carriers as well. At the same time, it generated significant additional revenue for the airlines that applied it. According to Smith et al. (1992), American Airlines estimates that over a three-year period at the end of the 1980s, quantifiable benefits of over 1.4 billion dollars were attributable to the control of discount price capacity and overbooking (i.e. selling more reservations than there are seats on the plane). Today, revenue management is both prevalent and mature in the airline industry. In fact, Talluri and van Ryzin (2004b, p. 10) state that the additional revenue generated by revenue management practices accounts for 4 to 5 percent of overall revenue, a value roughly comparable to many airlines' profits in a good year.

One major factor that enabled American Airlines to effectively apply revenue management practices was the use of information technology, namely central reservation systems, to manage the sale of seats. In addition to recording the number of seats sold and the number left to sell, central reservation systems also enabled better price and inventory management.

Capacity control mechanisms allow airlines to open and close the offer of discount fares depending on the number of seats still available, the time remaining until departure, and demand forecasts. Usually, these mechanisms are deeply embedded in the software logic and are expensive and difficult to change (Talluri and van Ryzin, 2004b, p. 28). According to Zhang and Cooper (2005) nested protection levels dominate airline practice due to the fact that many distribution channels allow only these types of controls.

A protection level y specifies the number of seats to reserve (protect) for a particular class or set of classes. If the plane's capacity was 100 and the protection level for full-fare demand was 70, a maximum of 30 seats could be sold at a discount price. Beyond this limit, the discount fare class is closed. In this example, "nested" means that full-fare demand has access to all the capacity reserved for lower fare demand. So if e.g. only 10 seats were sold at the discount fare but there is a high demand for seats at the full price, the airline could sell up to 90 seats at the full fare even if some of them had originally been assigned to the discount fare class. In the case of partitioned (non-nested) classes, only 70 seats would be offered at the full fare. As nested protection levels are so common in practice, one could argue that the formulation of a capacity control problem should require the selection of a control by protection levels (Zhang and Cooper, 2005). Yet the question remains whether (and if, when) this is restrictive.

These central reservation systems constitute one of the earliest examples of e-commerce. Based on the airline industry's success story, it is expected that the use of revenue management will be enhanced by the emerging role of Internet-based e-commerce, see Copeland and McKenney (1988), Smith et al. (2001), Baker et al. (2001), Boyd and Bilegan (2003), and Klein and Loebbecke (2003). But also apart from Internet-based e-commerce, today's information technology is enabling more and more industries to adopt revenue management practices. The hotel and hospitality sector, cruise lines, event promotion firms, and car rental companies are all examples of traditional applications that could be modeled in a similar way to the airline industry. Today, however, entities as diverse as the broadcasting, hospital, casino, and utility industries are starting to use revenue management practices. For surveys on the variety of applications, see McGill and van Ryzin (1999), Talluri and van Ryzin (2004b, Chap. 10), Yeoman and McMahon-Beattie (2004), Kimms and Klein (2005), and Sfodera (2006).

In the recent literature, the term "revenue management" encompasses more general demand management decisions, covering not only the determination of the number of tickets to offer at low prices (which is then called *capacity control*, *seat inventory control*, or *discount allocation*) but also *overbooking*. In addition to these quantity decisions, some authors even extend the meaning of revenue management to include the problem of product creation to ensure that full-fare demand is not sacrificed by offering tickets at discount prices (*market segmentation*) as well as the problem of finding the right prices and adjusting them over time (*dynamic pricing*).

For a survey of the different subtopics, see Kimes (1989), Weatherford and Bodily (1992), Harris and Peacock (1995), Weatherford (1998), McGill and van Ryzin (1999), Boyd and Bilegan (2003), Talluri and van Ryzin (2004b) and Phillips (2005). Different approaches to quantity decisions are summarized in Zehle (1991), Daudel and Vialle (1992), Klein (2001), and Tscheulin and Lindenmeier (2003). Dynamic pricing surveys can be found in Chan et al. (2004), Elmaghraby and Keskinocak (2003), and Bitran and Caldentey (2003); for legal aspects, see Weiss and Mehrotra (2001). Hybrid approaches that deal with both capacity allocation and optimal pricing of the fare classes can be found in Weatherford (1997) and Feng and Xiao (2006b). Badinelli (2000), Walczak (2001), Chatwin (2002), Chatwin (2003), and Maglaras and Meissner (2006) discuss the differences and similarities between capacity control and dynamic pricing. Note that dynamic pricing problems in revenue management do not consider replenishment; surveys on dynamic pricing and inventory decisions are provided by Elmaghraby and Keskinocak (2003) and Chan et al. (2004). Issues of price discrimination in the context of revenue management are discussed e.g. in Faßnacht and Homburg (1998); Talluri and van Ryzin (2004b, Chap. 8) furnish an overview.

To prevent misunderstanding, we will call the above-mentioned practice of determining the number of seats to protect for full-fare demand the "capacity control problem in revenue management", or "capacity control" for short. Overviews that exclusively deal with mathematical models for capacity control can be found in Ben-Yosef (2005, Chap. 7), Pak and Piersma (2002), and Kimms and Müller-Bungart (2004).

A variety of problems is summarized under the term revenue management, and its techniques are applied in many different industries. This work aims neither at an industry-specific nor at an all-embracing approach to revenue management. Instead, the goal is to provide a deeper understanding of the generic single-resource capacity control problem that forms the basis of many revenue management systems. The general terminology of capacity control will be introduced, but for simplicity, we will stick to the terminology of the airline industry thereafter.

1.1 The Basic Capacity Control Problem

The basic capacity control model is concerned with making efficient use of a certain, fixed capacity C of a single resource with homogeneous units that becomes worthless after a given time T.

The company sells its capacity as $i_{\max}$ distinct products. Each product $i = 1, \ldots, i_{\max}$ consists of one unit of this resource and is offered at a price (or fare) of ϱ_i. Without loss of generality, we assume that the products are indexed such that $0 < \varrho_{i_{\max}} \leq \cdots \leq \varrho_i \leq \varrho_1$.

In the airline industry, the capacity could be the number of seats in the economy compartment of a single-leg flight, i.e. a non-stop flight from one origin to one destination departing at some future point in time. All products, also called booking or fare classes, represent one (reservation for a) seat on that flight. They might only differ in price and/or in purchase restrictions such as a Saturday night restriction or early booking conditions. These artificial differences, known as "fencing conditions", ensure that the same resource can be sold at different prices. In the above-mentioned simplified case of American Airlines, there were $i_{\max} = 2$ products or booking classes: tickets at a discount fare that were available only several weeks before departure and the full-fare tickets. We assume the fencing conditions as well as the product prices to be fixed and exogenously given.

The question of capacity control is which products to offer for sale at a given point in time. Frequently, it is advantageous (and feasible) to reword this question and decide whether one should accept or reject an incoming request for product i given a certain amount of remaining capacity and time until departure.

1.1.1 Assumptions

We speak of a basic capacity control problem, if the following assumptions are made:

i) After a certain time T, the whole amount of capacity C is worthless. No additional units of capacity can be ordered.
ii) It is assumed that the major part of the costs is already sunk and that variable costs are negligible, so that the aim of profit maximization can be approximated by maximizing the revenue gained from the selling process.
iii) The resource has to be allocated dynamically as demand materializes. Rejected demand is lost and cannot be stored for the future. Once accepted, a customer cannot be rejected later without significant cost.

iv) There is considerable uncertainty about the quantity and the type of future demand. Future demand for the products offered can be described in terms of a random variable with a known probability distribution.
v) Products consist of a single (homogeneous) resource. Products that are composed of multiple resources, called network problems, are not considered.
vi) The company has monopolistic market power and customers are myopic.
vii) Demand (i.e. the number of requests) and time are discrete.
viii) Group bookings that have to be completely accepted or rejected are not considered. If there is demand for more than one ticket at a time, this demand may be partially accepted.
ix) Customers do not cancel (strictly) prior to the time of service. No-shows, i.e. customers that do not show up at the time of service, are not considered.
x) Demand for the products is independent of the availability of other products.
xi) The decision-maker's preferences can be approximated by a maximization of expected revenue; he is assumed to be risk-neutral.

This basic single-resource capacity control model does not reflect the state of the art in the revenue management literature, but it forms the basis for a lot of more advanced models and for most models used in practice.

Assumptions i), ii), and iii) form the heart of the capacity control problem in revenue management: a fixed amount of perishable capacity, high fixed costs and non-storable demand. (If it were possible to store demand, one would store all of it and sell seats right before capacity perished in decreasing price order.) If variable costs cannot be neglected, the product's contribution margin can usually be considered instead of the price (Zehle, 1991).

In some recent articles, however, the development of new products is propagated to allow the seller to reject some of the demand that has already been accepted against a certain compensation or to reassign it to a different type of capacity; see Biyalogorsky et al. (1999), Biyalogorsky and Gerstner (2004), Gallego and Phillips (2004), and Gallego et al. (2004). The latter products are frequently referred to as "flexible products" and are actually used in the hotel and cruise line industries. Generally, however, concerns about public image preclude their application (Biyalogorsky et al., 1999).

Although forecasting is considered an important ingredient for successful implementation, it is usually omitted from capacity control models. Yet assumption iv) is critical if new routes are offered or schedules are changed. The problem is discussed in van Ryzin and McGill (2000) and solved by an adaptive algorithm. To the author's knowledge, Bayesian demand learning is only considered in the dynamic pricing context; see e.g. Farias and Van Roy (2006) and the references given there. For a different perspective on capacity control that can do completely without demand forecasts, see Ball and Queyranne (2006).

Assumption v) is even more crucial in practice. Many large carriers operate on a flight network with hubs. Due to the curse of dimensionality, network capacity control is often tackled with approximate dynamic programming, heuristics, or simulations. For the basics on capacity control for flight networks, see Phillips (2005, Chap. 8) or Talluri and van Ryzin (2004b, Chap. 3) and the references given there. When the existence of more compartments (such as the business and economy compartments) is considered explicitly, issues of upgrading are discussed e.g. in Biyalogorsky et al. (2005) and Lukaschewitsch (2005). Although revenues are reported to improve by an additional 2.5 percent when network carriers optimize on the network level rather than on single-legs, single-leg models are still widely used (Talluri and van Ryzin, 2004b, p. 82). In addition, they form important building blocks in many heuristics for the network case (Talluri and van Ryzin, 2004b, p. 27).

Assumption vi) is standard in revenue management models. The first model to consider a basic capacity control model in a competitive framework is Netessine and Shumsky (2005). Strategic customers are considered in Anderson and Wilson (2003) and Liu and van Ryzin (2005).

To facilitate mathematics, some models assume demand to be continuous in contrast to the first part of assumption vii), see e.g. Curry (1990), Belobaba (1987a), or Bodily and Weatherford (1995). However, discrete demand seems more natural in real applications. The assumption of discrete time is not a hard to implement, since time can be discretized e.g. by counting arbitrarily small time intervals, by uniformization (Lippman, 1975), or by looking only at the times when demand materializes (as Lin, 2004, did in a dynamic pricing model). The latter approach turns the planning horizon into a random variable representing the number of points in time when demand arrives. Although most capacity control models use a discrete time approach, Liang (1999), Zhao and Zheng (2001), and Feng and Xiao (2006a) use a continuous time approach; a semi-Markov decision process is modeled in Walczak (2001) and Brumelle and Walczak (2003).

Papastavrou et al. (1996), Kleywegt and Papastavrou (1998, 2001) and van Slyke and Young (2000) demonstrate that the capacity control problem can also be formulated as a (stochastic) knapsack problem. They use this approach to handle group bookings that must be accepted or rejected as a whole. Lee and Hersh (1993) and Brumelle and Walczak (2003) consider group bookings within the framework of Markov decision processes. Among other things, they show that capacity control mechanisms that are suitable under assumption viii), such as the control by protection levels, are not optimal for total accept/deny decisions. According to Farley (2003, p. 155), small group bookings are usually either treated as individual bookings; airlines rely on manual processes to price and book larger groups. Eguchi and Belobaba (2004) give a recent overview of the literature on implemented group booking processes in airline revenue management.

In contrast to assumption ix), a significant proportion of tickets are canceled (strictly) prior to departure in airline applications. In addition, some