REVENUE MANAGEMENT BEYOND
"ESTIMATE, THEN OPTIMIZE"

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Vivek Francis Farias
June 2007

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Benjamin Van Roy) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Balaji Prabhakar)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Ramesh Johari (Management Science and Engineering))

Approved for the University Committee on Graduate Studies.

Abstract

Modern revenue management systems enable firms to make sophisticated pricing decisions over the course of a sales season. Many of these systems operate within what one might refer to as the "estimate, then optimize" paradigm where estimation and optimization are two distinct, interleaved activities. This thesis will discuss two research efforts that explore moving away from such a paradigm.

We begin with the study of a simple model of one-resource revenue management that incorporates uncertainty in demand statistics. The opportunity to learn more about demand over the course of the sales season introduces a tension between "exploratory" pricing that attempts to learn quickly and "exploitative" pricing that attempts to exploit existing demand knowledge so as to maximize revenues. We present a simple heuristic that addresses this trade-off in a transparent, operationally intuitive, manner. We establish that pricing decisions that account for this trade-off offer significant increases in revenue over repeated cycles of "estimate, then optimize."

We next turn our attention to the dynamic capacity allocation problem that airlines face. We will present new approximate dynamic programming based algorithms for this large-scale problem that allow for a seamless integration of complex demand forecast models within the optimization framework. The algorithms we develop are scalable to large problems and we present computational results that suggest a significant improvement over methods that rely on frequent re-optimization using a popular linear programming based heuristic.

To my parents – Francis and Juliana Farias

Acknowledgements

This thesis is the product of, what for me has been, a very fruitful collaboration with my advisor and friend Professor Benjamin Van Roy. Ben has taught me the importance of seeking problems that are of broad interest, formulating them correctly, and not quitting when the going gets tough. And he has taught me all of this by example, while treating me no different from a colleague. Life in graduate school can be very trying at times - for social, academic and other reasons. Ben has, in many ways, displayed a genuine concern for my well being and this has been as important to me as his guidance in academic matters. I would be hard-pressed to distinguish my respect for Ben from awe.

Much of the work in this thesis has benefited from the advice of Professor Michael Harrison, who has always been amazingly accessible. Chapter 2 in this thesis was motivated by a problem Mike told me about over the course of a Revenue Management class he taught in 2005 and his emphasis on the importance of moving beyond Poisson models for Network RM led to Chapter 4 in its current form. I've collaborated on other research problems with Professors Balaji Prabhakar and Tsachy Weissman. I admire Balaji's style and his keen knack for formulating elegant problems. The little information theory I know, I learned from Tsachy who is in addition to being a gracious teacher, a formidable problem solver. In recent years I have come to learn some game theory thanks to Professor Ramesh Johari. We continue to have very stimulating discussions and although we haven't written a paper together yet, I hope that changes soon. Finally, I owe a debt of gratitude to Professor Dunbar Birnie and the late Professor Michael Weinberg at the University of Arizona for starting me out with a great deal of care and support on the path I have chosen.

Graduate school would probably have been very different without my office mate and good friend, Ciamac Moallemi. Ciamac and I have written two papers together, and I am certain that our collaboration and friendship will continue over the years to come. Paat Rusmevichientong and Gabriel Weintraub graciously read parts of this work. They were also generous with advice for my job search for which I am grateful. I was fortunate to have as fellow group members David Choi, Jared Han, Kahn Mason, Michael Padilla, Waraporn Tongprasit and Robbie Yan. The Information Systems Laboratory at Stanford (ISL) has been a nice place to call home over the last five years – and it wouldn't be nearly quite as nice without our administrator Denise Murphy and some of the frighteningly clever and down-to-earth people I have befriended here.

There are a number of close friends without whose help I wouldn't have made it this far in one piece. My good friend Pieter Abbeel has been a sounding board for many ideas and a constant source of encouragement through times of self doubt. My roommate and partner-in-crime Jian Wong and I have done a number of things over the last five years that will hopefully never make it to print – I will really miss those fun times. Among my other close friends are David Wakayama, Scott Coughlin and Melissa Kar (from my undergraduate days at Arizona), and Chinmay Belthangady, Ankit Bhagatwala, James Mammen and Jimmy Zhang here at Stanford. I am grateful to all of them for their untiring friendship over these years.

My parents Francis and Juliana have over the years sacrificed a great deal to get me to where I stand today. They are honest and hardworking and I have tried – and will continue to try – to make myself in their mould. My accomplishments to this point are primarily the outcome of their parenting. Together with my baby brother Vinay, they are the people I love the most, and all of this would be hopeless were it not for them.

Contents

List of Tables

List of Figures

Chapter 1

Revenue Management Beyond "Estimate, Then Optimize"

"Revenue Management" (or RM for short) is today a ubiquitous area of operations research that is concerned with developing into a science, the art of selling the right item, to the right person, at the right price. This thesis is concerned with addressing the flaws of an operational paradigm we refer to as "estimate, then optimize", that has become commonplace in the modern practice of revenue management. Our objective is to propose new schemes that, by appropriately addressing these flaws, are capable of adding to the revenues generated by existing RM mechanisms.

This chapter begins with a brief introduction to the field of revenue management. We then introduce the "estimate, then optimize" paradigm in the context of two fairly commonplace RM models. After a discussion of the flaws inherent to this paradigm, we provide a brief overview of the contributions made by this thesis in addressing those flaws. We restrict ourselves in this chapter to a high level, qualitative discussion; subsequent chapters will make more detailed arguments supported either by mathematical results or computational evidence.

1.1 What is Revenue Management?

Revenue management has over the years come to be associated with a broad variety of sales practices across an equally broad array of industries. It is therefore natural that any attempt at a concise description of the term is likely to fail to capture the finer nuances of the practice. Nonetheless, we attempt to provide such a description: Revenue management refers to the optimal or near-optimal use of a sales mechanism so as to maximize over some period, expected revenues from the sale of one or more products, each requiring quantities of one or more potentially scarce resources. While certainly not comprehensive, this definition is rather broad. In particular, the products in question could be essentially anything ranging from airline tickets, hotel rooms and fashion goods on the one hand, to power and natural gas on the other. The sales mechanisms employed could also vary widely. For example, a product may be sold at a posted price in which case the seller must decide on what prices to post over time. As another example, items of a product may be auctioned in which case the seller must design a suitable auction mechanism. Finally, the phrases "optimal use" and "maximize expected revenues" suggest a well specified mathematical optimization problem and a number of reasonable formulations are likely to be viable for a given revenue management problem. The very act of selling a product entails making the types of decisions we have alluded to and it is this last feature – the mathematical systematization of the typical decisions that must be made by a seller – that distinguishes RM.

To get a sense for the diversity inherent to the practice of RM, consider the following instances: Airlines offer an array of "fare-products" (which are essentially itinerary-price-restriction combinations) whose availability are carefully modulated over time using RM tools. Potential customers arrive through varied sources such as internet travel sites or through a travel agent which in turn, through an interface with the airline's reservation management system, allow the customer to choose from fare-products that are available at that point in time. Fashion goods are typically sold through retail outlets where prices posted for various items are adjusted multiple times over the course of a sales season based on a multitude of factors including the perceived popularity of the item and its availability. Again such price adjustments often require the support of an

RM system. As yet another example, Natural gas pipeline operators offer a variety of gas delivery contracts at rates that depend on available capacity on their pipelines and gas futures prices in spot markets (such as the New York Mercantile Exchange); customers are typically utility companies or large industrial units and purchases may be made either directly or through online market places such as the Intercontinental Exchange. Again, RM systems aid managers in pricing such products.

Effective RM tools can have an enormous impact on the revenues and profits of a firm. While there is no dearth of examples across industries, one of the most notable, and perhaps earliest, success stories is the American Airlines RM system “DINAMO”. By carefully controlling the availability of various fare-products on their network via DINAMO, American estimates that they added 1.4 Billion dollars to their bottom line for the period from 1989-92 (see Smith et al. (1992)). This level of success, however, is contingent upon a number of factors. These include: engineering an information processing system that streamlines the sales process and the acquisition of sales data that a manager might find useful, building mathematical models that capture features of the sales process that managers believe to be important, and finally the design of algorithmic tools to support decision making within these models. There are two broad classes of algorithmic problems that arise in this context:

Forecasting and Estimation: A “forecast model” is a black box that predicts demand for products based on factors such as price, popularity and historical demand for that product and similar products, market factors such as competitor prices and so forth. Firms treat their proprietary forecast models as trade secrets. Designing good forecast models is an art and typically requires an experts intuition. Using a forecast model typically requires some manner of statistical estimation – be it regressions against historical data to estimate forecast model parameters, or Bayesian updating to estimate state in a state-space model of demand. As such the design of forecast models and estimators leverages a number of tools from the statistics and econometrics literature.

Optimization: Many revenue management problems are posed as optimization problems; inventory constraints and uncertainty in demand make these problems stochastic and dynamic. These optimization problems are often computationally difficult, the chief difficulty arising from having to jointly manage inventory levels of multiple resources and from the complexities of the processes that describe demand. Optimal solution is rarely possible, and heuristics that relax inventory constraints in certain ways or that make simplifying assumptions about demand are often called for.

This thesis is concerned with issues that arise in the design of *optimization* tools for an RM system. We next introduce two problems that arise in the context of retail and airline RM respectively. These will serve as illustrative examples for our discussion over the remainder of this chapter.

1.2 Problems and Models: Two Examples

Example 1. *Dynamic Pricing for a Single Product*

Consider a retailer selling a single product over some finite sales season. At every point in time, the retailer must post a price should he have inventory of the product on hand. An arriving customer must pay the posted price at that time should he choose to acquire the product. In building an RM system to decide what price to post at each point in time, the retailer might begin with the following modeling assumptions:

1. *Potential customers arrive according to some stochastic point process (say a Poisson process).*

2. *Arriving customers have "reservation prices" and make a purchase if and only if their reservation price exceeds the posted price.*

3. *Reservation prices are themselves random and distributed according to some known distribution.*

With these three modeling assumptions, and assuming the retailer has succeeded in estimating the rate at which customers arrive and their reservation price distribution, one may show that his optimal pricing decision is a function of the time remaining in the sales season and the inventory he has on hand. Dynamic programming may then be used to compute the optimal price function.

Example 2. *Capacity Control for Airline Network RM*

Consider an airline that operates flights across a network of cities. The airline sells fare-products, a given fare-product being completely specified by its associated itinerary, price and restrictions (such as the non-refundability of a canceled ticket). In selling tickets for flights on a particular date in the future, the airline must at suitably frequent intervals of time up to that date decide which fare-products it will make available until the next decision epoch. Of course, the airline can make available a given fare-product only if it has seats available on the flight legs required for that product [1]*. An arriving customer may be interested in a subset of fare-products offered by the airline and must choose among available products (or not purchasing). Again, the retailer might begin with a few modeling assumptions:*

1. *Arriving customers have reservation prices for each of the products available upon their arrival and choose to purchase the product that maximizes their surplus (that is, the difference between their reservation price for the product and its price). A customer buys nothing if no product offers him a positive surplus.*

2. *Customer reservation prices for the products offered are themselves random and jointly distributed according to some known distribution; there are several customer "types" – each type is associated with a specific joint distribution of reservation prices.*

3. *Potential customers of each type arrive according to some stochastic point process (such as a Poisson process) specific to that type.*

Again, assuming that the revenue manager is able to estimate the required arrival rates and reservation price distributions, the optimal capacity allocation decision at every point in time may be shown to be a function of available capacities on each network leg and the time remaining to the end of the sales horizon. The optimal capacity allocation function may be found via dynamic programming, though the compute time required would likely render such an approach impractical.

A seller is ultimately interested in the net proceeds from all sales over the course of a sales season and the sales mechanisms implicit in Examples 1 and 2 are by no means

[1] Although, in general one may also consider over-booking

the only – or most suitable – mechanisms possible. For example, the retailer may choose to sell products via a dynamic auction of some sort. In the case of airline network capacity allocation the notion of a fare-product is somewhat synthetic – at the very least it imposes an unnecessary discretization on the price a consumer is charged. Deciding on the "right" sales mechanism, while important, depends on a large number of factors including laws (such as those forbidding price discrimination) or the infrastructure in place to support the sales mechanism (such as the computer reservation processing system an airline might have in place). As such, many industries including the airline and fashion retail industries have established sales mechanisms in place that have seen few, if any, fundamental changes in decades. We will, in problems we consider in this thesis, not consider these issues and instead treat the sales mechanism as a given.

Example 1 outlines a potential RM system for the retailer's problem. Its success hinges on the validity of the modeling assumptions and the ability of the retailer to estimate customer arrival rate and the reservation price distribution. Having done so, the relevant dynamic program is in fact quite simple. Even so, it is unlikely that accurate estimates of the type needed are available at the start of a sales season, and in practice these estimates might be refined as the sales season unfolds.

The network capacity control problem is more complex. For one, even if the modeler were able to identify customer types, it is unlikely that sufficient data for reliable estimates of joint reservation price distributions for each of these types will be available. Even if these difficulties could somehow be surmounted, the relevant dynamic optimization problem that one must solve to compute optimal capacity allocation decisions suffers from the "curse of dimensionality" and is, for all practical purposes, not amenable to efficient solution. This necessitates the consideration of further specialized models. For example, one may restrict attention to parametric families of reservation price distributions. Assuming that the data to estimate such a model is available one is still left with a difficult dynamic optimization problem. A natural simplification is to assume customer arrival processes are deterministic with rates that agree with observed averages. In doing so, we are left with a far simpler, often tractable, optimization problem. While such simplifications yield tractability, it is clear that models of this nature are likely to provide very crude descriptions of demand and not leverage the full power of a good forecast model's abilities. Coping with these difficulties is something of an art, but typical strategies include repeated model re-estimation and optimization of the re-estimated model.

1.3 "Estimate, Then Optimize"

Consider a vendor of Winter apparel. New items are stocked in the Autumn and sold over several months. Because of significant manufacturing lead times and fixed costs, items are not restocked over this period. Evolving fashion trends generate great uncertainty in the number of customers who will consider purchasing these items. To optimize revenue, the vendor should adjust prices over time. But how should these prices be set as time passes and units are sold? At first sight, this problem appears amenable to the type of modeling that Example 1 suggests. A problematic issue however, is that at the start of the sales season, the retailer is unlikely to have reliable estimates of the customer arrival rate or perhaps even the reservation price distribution required for that model. How might one address this issue?

One pragmatic solution proceeds as follows: Start with a guess for the customer arrival rate and reservation price distribution. There are many reasonable ways of coming up with a good guess. For example, market research or historical sales data for similar products that were sold in the past might be used to estimate these quantities. Given such a guess, solve the necessary dynamic optimization problem to compute an optimal price (assuming the guess is correct). As the sales season unfolds, the retailer accumulates new data on the frequency with which he sees customers purchasing his product at various price levels which in turn allows him to refine his guess of arrival rate and reservation price distribution. He then solves a new dynamic optimization problem based on these revised estimates and proceeds to price based on the solution to this new problem. A procedure of this type could potentially be repeated many times over the course of a sales season with estimates that improve over time. For obvious reasons, we refer to this as the "estimate, then optimize" paradigm.

There are some qualitative differences in the nature of the demand uncertainties faced by a fashion retailer and an airline revenue manger. In particular, the airline is in a position to produce more accurate demand forecasts. For instance, the airline will have access to large quantities of historical data that let the airline calibrate complex forecasting models. Such models might be capable of predicting demand contingent on historical information and various other relevant factors. One mathematical abstraction for such a forecasting model is a Markov-modulated Poisson process with partially observable modulating states. The role of estimation here is then to estimate the underlying state in such a model.

The underlying dynamic optimization problem for the capacity allocation problem is challenging even in the context of simple stochastic demand processes (such as time homogeneous Poisson processes). As such, it is not uncommon in formulating an optimization problem, to assume that demand for a given fare product is deterministic and equal to the expected forecasted demand, and make allocation decisions based on such an assumption. The revenue manager then relies on frequently updating the demand quantities assumed by the optimization algorithm in the hope that this corrects for the fact that the algorithm assumes a model that is a poor description of reality although this clearly fails to utilize all of the forecasting models potential predictive abilities. Although more complex, the "estimate, then optimize" cycle in the case of network capacity control continues to fulfill the same two basic needs as it did in the case of the retailers problem: one is the need to compute updated forecasts, and the second is to update the optimization model inputs so as to compensate for the fact that it assumes a crude model for demand.

The "estimate, then optimize" paradigm is a natural means to addressing the complexities inherent to revenue management in the face of large uncertainties in demand. But it is not without its flaws:

1. **Optimizing assuming the "wrong" model of demand:** For an optimization algorithm to use a cruder model of customer demand than is available is an obvious shortcoming. Over each optimization phase, the retailer assumes that demand is a Poisson process of a deterministically known rate, whereas in reality there is uncertainty in this rate which is unaccounted for; depending on the precise nature of this uncertainty, the retailer may want to adjust prices so as to hedge effectively

against the possibilities of a favorable or unfavorable demand environment. Similar problems are evident with using a crude demand model in the network capacity control case. In particular, consider the following toy forecast model: the model predicts constant demand up to a certain point in time. Beyond that point, demand either goes up by 20 % or falls by 20 % with equal probability. Under the "estimate, then optimize" paradigm, an optimization algorithm that operates based on expected demand forecasts will assume that demand is to remain constant over all time, and updates this assumption only once demand actually does change. Such an algorithm cannot be expected to hedge effectively between the two potential outcomes of demand going up by 20 % or falling by 20 %.

2. **Ignoring the incentive to learn:** The price set by the fashion retailer impacts, in addition to his revenues, the rate at which he is able to learn about customer demand. In particular, price serves to censor demand, so that at high prices the retailer learns slowly thereby potentially wasting precious selling time, whereas at low prices he learns quickly but at the expensive of potentially precious inventory. Clearly a trade-off needs to be made between eliminating uncertainty in demand statistics, exploiting existing demand knowledge and hedging against the possibility that demand for the product is in fact higher than expected. The "estimate, then optimize" paradigm ignores these trade-offs entirely. The same criticism is relevant to the case of network capacity control as well. In particular, the relevant inputs to a forecast model for network capacity control need to be estimated. For example, in the case of Markov-modulated demand with partial observability, the underlying arrival rate modulating state needs to be estimated. A high degree of uncertainty in the underlying state might call for controls that quickly eliminate this uncertainty so as to have accurate forecasts available.

These flaws are by no means subtle. We are naturally led to wonder:

- Would addressing these issues produce a tangible impact on revenues?
- Can these flaws be addressed in a manner that is robust and efficient?

Our intention is to explore these questions.

1.4 Beyond “Estimate, Then Optimize”

This thesis makes an attempt to move beyond the “estimate, then optimize” paradigm and address some of its flaws. The paradigm – like our discussion thus far – is quite general and making progress requires us to specialize our attention in several ways. For one, formulating well posed problems will require us to restrict focus to specific revenue management models. We focus in turn on two models very similar in spirit to those suggested in Examples 1 and 2. Both models we consider are natural generalizations of *optimization* models commonplace in the academic RM literature. The generalizations we incorporate allow for richer forms of demand uncertainty and capture features of real world problems that we believe are generally ignored in the formulation of RM optimization problems for want of tractable solution techniques; they are instead typically dealt with via repeated iterations of estimation followed by re-optimization.

We consider in Chapter 2, a model for one product dynamic pricing where we incorporate uncertainty in demand via the introduction of a prior on customer arrival rate; such models typically call for the solution of high dimensional dynamic programming problems and standard models typically ignore this type of uncertainty. In addition to having several potential real-world applications, our model allows us to understand in a precise way some of the flaws inherent to a scheme based on “estimate, then optimize”. Chapter 2 makes several contributions. Among them:

- We propose “decay balancing” – a simple, new heuristic for dynamic pricing in the face of demand uncertainty. Unlike methods that rely on repeated re-optimization based on revised estimates of expected arrival rate, decay balancing prices implicitly account for the level of uncertainty in making pricing decisions. While being no more complex than a typical “estimate, then optimize” type scheme, decay balancing shows performance improvements of up to about 30% over such schemes.

- We demonstrate performance guarantees for our heuristic, including a uniform performance bound: For Gamma priors on arrival rates and exponential reservation prices, Decay balancing is a 3-approximation algorithm. Such bounds are indicators of robustness across all parameter regimes; computationally observed performance losses are on the order of 1-2%.

- We derive key structural properties that an optimal scheme must possess and show that Decay Balancing inherits these properties.

In the previous section, we were led to ask whether overcoming the flaws inherent to the "estimate, then optimize" paradigm could have a tangible impact on revenues, and whether this could be accomplished via simple schemes. For the one product RM model introduced in Chapter 2, decay balancing provides an affirmative answer to both questions.

Chapter 4 considers a model for network RM wherein customers arrive according to a Markov modulated Poisson process. This is a substantial generalization of the deterministic rate customer arrival processes considered in a majority of the network RM literature. It is an important generalization since it allows for the integration of relatively complex demand forecast models in the optimization process. One might hope that optimal or near optimal solutions to dynamic optimization problems that arise from such models would lead to improvements over schemes that rely on solutions to optimization problems that assume far cruder models of demand (such as expected demand forecasts). Solving such optimization problems however is non-trivial, and Chapter 4 makes several contributions in this direction:

- We develop an approximation algorithm for a dynamic capacity allocation problem arising from a network RM model with Markov modulated arrival rates. Our algorithm is based on the linear programming (LP) approach to approximate dynamic programming (DP).

- Our algorithm demonstrates performance gains of up to 8% over an approach that uses only expected demand forecasts. The approach we compare our performance to is representative of the state of the art in network RM optimization.

- Our algorithm is scalable. In particular, its use requires the one time solution of a single LP that even for large networks could potentially be solved in minutes.

Chapter 4 proposes models and algorithms with a view to allowing for more realistic demand modeling in optimization for network RM. The results in that chapter suggest that doing so is likely to be viable in practice and could in addition to yielding levels of performance superior to the state of the art, substantially reduce dependence on the frequent re-optimization necessary for optimization models that assume crude models of demand.

The "estimate, then optimize" paradigm is applicable to a number of models beyond the scope of this thesis and is an approach that is easily understood and adapted. The schemes we propose on the other hand, while relatively simple to implement, are tailored to specific models. The decay balancing heuristic can be extended to models other than the vanilla one-product dynamic pricing problem in Chapter 2 (see Chapter 3), and the approximate DP approach that drives the algorithms in Chapter 4 is almost certainly applicable to many problems in RM that call for the solution of high dimensional dynamic programs. Nonetheless, moving beyond the "estimate, then optimize" paradigm in general is likely to require some effort on the part of the algorithm designer. In a world where 1-2% gains in revenue have potentially large implications for the profits of a firm, this effort is likely to be well rewarded.

1.5 Further reading

Revenue management is a fairly broad area of research and borrows heavily from fields such as marketing, statistics and stochastic control. The books by Talluri and van Ryzin (2004) and Phillips (2005) are excellent, encyclopedic resources for a broad overview of the area. Little is available in the way of literature on the estimation and forecasting practices in the RM industry, and these texts are among the few thorough treatments of those areas. In contrast, much of the academic RM literature is dedicated to *optimization* problems that arise in various RM contexts.

There are a number of papers on various aspect of airline RM. Smith et al. (1992) provides an overview of the RM heuristics that went into building American Airlines' first successful RM system DINAMO while P.P.Belobaba (2001) surveys industry practice. Papers by Gallego and van Ryzin (1997), Bertsimas and de Boer (2005), van Ryzin and McGill (2000) and Bertsimas and Popescu (2003) are representative of modern algorithmic approaches to the network RM dynamic capacity allocation problem. The main thrust of the work in these papers is designing effective means of dealing with the curse of dimensionality that arises from having to jointly manage multiple resources in network RM problems. RM for the natural gas and fashion industries are of relatively newer vintage; as examples, Talluri and van Ryzin (2004) discuss RM problems that arise in the natural gas industry while Bitran and Mondschein (1997) consider optimal markup/ markdown policies for fashion retail.

Gallego and van Ryzin (1994) undertake a detailed study of a model similar to that in Example 1 and provide optimal and heuristic dynamic pricing policies under the assumption of known arrival rates and reservation price distributions while a subsequent paper (Gallego and van Ryzin (1997)) considers the multi-dimensional (that is, multi-resource, multi-product) generalization to that problem. Alternative sales mechanisms – such as dynamic auctions – have not been throughly studied in an RM context. Vulcano et al. (2002) explores the use of a dynamic auction for a one resource, finite time horizon revenue management problem.

The "estimate, then optimize" paradigm has received little attention in the literature; an exception is Eren and Maglaras (2006) that points out the potential dangers of using such an approach for a certain one product RM problem. Other authors have recognized some of the flaws inherent to the paradigm in the context of problems such as examples 1 and 2 and proposed alternatives; we will discuss that work in subsequent chapters.

Chapter 2

Dynamic Pricing with an Uncertain Market Response

This chapter studies a problem of dynamic pricing faced by a vendor with limited inventory, uncertain about demand, aiming to maximize expected discounted revenue over an infinite time horizon. The vendor learns from purchase data, so his strategy must take into account the impact of price on both revenue and future observations; a key flaw of the "estimate, then optimize" paradigm was its failure to account for this trade-off. We focus on a model in which customers arrive according to a Poisson process, each with an independent, identically distributed reservation price. Upon arrival, a customer purchases a unit of inventory if and only if his reservation price equals or exceeds the vendor's prevailing price.

We propose in this chapter a new heuristic approach to pricing, which we refer to as *decay balancing*. Among other performance bounds, we establish that when reservations prices are exponentially distributed and the vendor begins with a Gamma prior over arrival rates, decay balancing always garners *at least one-third* of the maximum expected discounted revenue. This is the first heuristic for problems of this type for which a uniform performance bound is available. We also establish that changes in inventory and uncertainty in the arrival rate bear appropriate directional impacts on decay balancing prices, in contrast to the recently proposed *certainty equivalent* and *greedy* heuristics. Further, we provide computational results to demonstrate that decay balancing offers significant revenue gains over these alternatives. The decay balancing heuristic may be extended to several interesting related models; we pursue these extensions in the next chapter.

The remainder of this chapter is organized as follows: Section 2.1 introduces the problem and places it in the context of other work in the general area of pricing in the face of demand uncertainty. In Section 2.2, we formulate our model and cast our pricing problem as one of stochastic optimal control. Section 2.3 develops the HJB equation for the optimal pricing problem in the contexts of known and unknown arrival rates. Section 2.4 introduces an existing "estimate, then optimize" style heuristic (the certainty equivalent heuristic) for the problem, while section 2.5 introduces a recently proposed heuristic that attempts to address the flaws inherent to an "estimate, then optimize" based approach. Section 2.6 introduces decay balancing which is the focus of this chapter. This section also discusses structural properties of the decay balancing policy. Section 2.7 is devoted to a performance analysis of the decay balancing heuristic. When the arrival rate is Gamma distributed and reservation prices are exponentially distributed, we prove a uniform performance guarantee for our heuristic. Section 2.8 presents a computational study that compares decay balancing to certainty equivalent and greedy pricing heuristics as well as a clairvoyant algorithm. Finally, in Section 2.9 we conclude with thoughts on future prospects for this work.

2.1 Introduction

In motivating the need to consider moving beyond the "estimate, then optimize" paradigm, the last chapter considered the example of a vendor of winter apparel who needed to adjust prices over time in the face of limited inventory and great uncertainty in the number of customers who might consider purchasing his product. This is representative of problems faced by many vendors of seasonal, fashion, and perishable goods.

There is a substantial literature on pricing strategies for such a vendor (see Talluri and van Ryzin (2004) and references therein). Gallego and van Ryzin (1994), in particular, formulated an elegant model in which the vendor starts with a finite number of identical indivisible units of inventory. Customers arrive according to a Poisson process, with independent, identically distributed reservation prices. In the case of exponentially distributed reservation prices the optimal pricing strategy is easily derived. The analysis of Gallego and van Ryzin (1994) can be used to derive pricing strategies that optimize expected revenue over a finite horizon and is easily extended to the optimization of discounted expected revenue over an infinite horizon. Resulting strategies provide insight into how prices should depend on the arrival rate, expected reservation price, and the length of the horizon or discount rate.

Our focus in this chapter is on an extension of this model in which the arrival rate is uncertain and the vendor learns from sales data. Incorporating such uncertainty is undoubtedly important in many industries that practice revenue management. For instance, in the Winter fashion apparel example, there may be great uncertainty in how the market will respond to the product at the beginning of a sales season; the vendor must take into account how price influences both revenue and future observations from which he can learn.

In this setting, it is important to understand how uncertainty should influence price. However, uncertainty in the arrival rate makes the analysis challenging. Optimal pricing strategies can be characterized by a Hamilton-Jacobi-Bellman (HJB) Equation, but there is no known analytical solution. Further, for arrival rate distributions of interest, grid-based numerical methods require discommoding computational resources and generate strategies that are difficult to interpret. As such researchers have designed and analyzed heuristic approaches.

Aviv and Pazgal (2005) studied a *certainty equivalent* heuristic for exponentially distributed reservation prices which at each point in time computes the conditional expectation of the arrival rate, conditioned on observed sales data, and prices as though the arrival rate is equal to this expectation. This is precisely the "estimate, then optimize" paradigm and it inherits the flaws of that paradigm we pointed out in the introductory chapter. In particular, it uses an incorrect demand model and further ignores the incentive to learn.

In an effort to address these flaws Araman and Caldentey (2005) recently proposed a more sophisticated heuristic that takes arrival rate uncertainty into account when pricing. The idea is to use a strategy that is *greedy* with respect to a particular approximate value function. In this chapter, we propose and analyze *decay balancing*, a new heuristic approach which makes use of the same approximate value function as the greedy approach of Araman and Caldentey (2005).

Several idiosyncrasies distinguish the models studied in Aviv and Pazgal (2005) and Araman and Caldentey (2005). The former models uncertainty in the arrival rate in terms of a Gamma distribution, whereas the latter uses a two-point distribution. The former considers maximization of expected revenue over a finite horizon, whereas the latter considers expected discounted revenue over an infinite horizon. To elucidate relationships among the three heuristic strategies, we study them in the context of a common model. In particular, we take the arrival rate to be distributed according to a finite mixture of Gamma distributions. This is a very general class of priors and can closely approximate any bounded continuous density. We take the objective to be maximization of expected discounted revenue over an infinite horizon. It is worth noting that in the case of exponentially distributed reservation prices such a model is equivalent to one without discounting but where expected reservation prices diminish exponentially over time. This may make it an appropriate model for certain seasonal, fashion, or perishable products. Our modeling choices were made to provide a simple, yet fairly general context for our study. We expect that our results can be extended to other classes of models such as those with finite time horizons, though this is left for future work.

When customer reservation prices are exponential and the arrival rate is Gamma distributed we prove that decay balancing always garners at least 33.3% of the maximum expected discounted revenue. Allowing for a dependance on the number of sales we can show that after four sales decay balancing achieves at least 80% of optimal performance thereafter. It is worth noting that no performance loss bounds (uniform or otherwise) have been established for the certainty equivalent and greedy approaches. Further, our computational results suggest that our theoretical bounds are conservative and also that decay balancing offers substantial increases in revenue relative to certainty equivalent and greedy approaches. Surprisingly, though the two heuristics are based on the same approximate value function, switching from the greedy approach to decay balancing can increase expected discounted revenue by over a factor of three. Further, uncertainty in the arrival rate and changes in inventory bear appropriate directional impacts on decay balancing prices: uncertainty in the arrival rate increases price, while a decrease in inventory increases price. In contrast, uncertainty in the arrival rate has no impact on certainty equivalent prices while greedy prices can increase or decrease with inventory.

Aside from Aviv and Pazgal (2005) and Araman and Caldentey (2005), there is a significant literature on dynamic pricing while learning about demand. Lin (2007) considers a model identical to Aviv and Pazgal (2005) and develops heuristics which are motivated by the behavior of a seller who knows the arrival rate and anticipates all arriving customers. Bertsimas and Perakis (2003) develop several algorithms for a discrete, finite time-horizon problem where demand is an unknown linear function of price plus Gaussian noise. This allows for least-squares based estimation. Lobo and Boyd (2003) study a model similar to Bertsimas and Perakis (2003) and propose a "price-dithering" heuristic that involves the solution of a semi-definite convex program. All of the aforementioned work is experimental; no performance guarantees are provided for the heuristics proposed. Cope (2006) studies a Bayesian approach to pricing where inventory levels are unimportant (this is motivated by sales of on-line services) and there is uncertainty in the distribution of reservation price. His work uses a very general prior distribution (a Dirichlet mixture) on reservation price. Modeling this type of uncertainty within a framework where inventory levels do matter represents an interesting direction for future work. In contrast with the the above work, Burnetas and Smith (1998) and Kleinberg and Leighton (2004) consider non-parametric approaches to pricing with uncertainty in demand. However, those models again do not account for inventory levels. Recently, Besbes and Zeevi (2006) presented a non-parametric algorithm for "blind" pricing; they present a pricing algorithm for pricing multiple products that use multiple resources, similar to the model considered in Gallego and van Ryzin (1997). Their algorithm requires essentially no knowledge of the demand function. While the algorithm is optimal under a certain fluid-limit like scaling, the algorithm requires testing every possible price vector within a multidimensional grid which represents a discretization of the space of price vectors.

2.2 Problem Formulation

We consider a problem faced by a vendor who begins with x_0 identical indivisible units of a product and dynamically adjusts price p_t over time $t \in [0, \infty)$. Customers arrive according to a Poisson process with rate λ. As a convention, we will assume that the arrival process is right continuous with left limits. Each customer's reservation price is an independent random variable with cumulative distribution $F(\cdot)$. A customer purchases a unit of the product if it is available at the time of his arrival at a price no greater than his reservation price; otherwise, the customer permanently leaves the system.

For convenience, we introduce the notation $\overline{F}(p) = 1 - F(p)$ for the tail probability. We place the following restrictions on $F(\cdot)$:

Assumption 1.

1. *$F(\cdot)$ has a density $f(\cdot)$ with support $\mathbb{R}^+$.*
2. *$\frac{\rho(p)}{\overline{F}(p)}$ is an increasing function of p, where $\rho(p) = \frac{f(p)}{\overline{F}(p)}$ is the hazard rate function for F.*
3. *$p - 1/\rho(p)$ is a surjective function of p with $\mathbb{R}^+$ in it's range.*

The first assumption is a regularity assumption. Now, one may think of $\overline{F}(p) \equiv q$ as the expected quantity of the product sold at price p garnering the seller an expected revenue of $R(p) = p\overline{F}(p)$. By the first part of Assumption 1, given a quantity $q \in (0, 1]$, the unique price that achieves this expected quantity is given by $p(q) = F^{-1}(1 - q)$, so that expected revenue can also be thought of as a function of q, $\tilde{R}(q) = R(p(q))$. The marginal revenue to the seller with respect to quantity is then given by $d\tilde{R}/dq = p(q) - 1/\rho(p(q))$. S. Ziya and Foley (2004) note that if f is differentiable, the second assumption is equivalent to the statement that marginal revenue with respect to quantity is increasing in price and equivalently, decreasing in quantity. This is a reasonable economic premise. If the first and second parts of Assumption 1 hold, the assumption that $p - 1/\rho(p)$ is surjective with range $\mathbb{R}^+$ is equivalent to assuming that expected revenue in the presence of a finite non-negative marginal cost c, $(p - c)\overline{F}(p)$, is maximized at some finite price p^*. This too appears to be reasonable. Our assumptions are standard to the revenue management literature (see Talluri and van Ryzin (2004) pp. 315-318) and permit us to use first order optimality conditions to characterize solutions to various optimization problems that will arise in our discussion.

Let t_k denote the time of the kth purchase and $n_t = |\{t_k : t_k \leq t\}|$ denote the number of purchases made by customers arriving at or before time t. The vendor's expected revenue, discounted at a rate of $\alpha > 0$, is given by

$$E\left[\int_{t=0}^{\infty} e^{-\alpha t} p_t dn_t\right].$$

Let $\tau_0 = \inf\{t : x_t = 0\}$ be the time at which the final unit of inventory is sold. For $t \leq \tau_0$, n_t follows a Poisson process with intensity $\lambda \overline{F}(p_t)$. Consequently, one may show that

$$E\left[\int_{t=0}^{\infty} e^{-\alpha t} p_t dn_t\right] = E\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \overline{F}(p_t) dt\right].$$

We now describe the vendor's optimization problem. Because of differences in these two contexts, we first consider the case where the vendor knows λ and later allow for arrival rate uncertainty. In the case with known arrival rate, we consider pricing policies π that are measurable real-valued functions of the inventory level. The price is irrelevant when there is no inventory, and as a convention, we will require that $\pi(0) = \infty$. We denote the set of policies by Π_λ. A vendor who employs pricing policy $\pi \in \Pi_\lambda$ sets price according to $p_t = \pi(x_t)$, where $x_t = x_0 - n_t$, and receives expected discounted revenue

$$J_\lambda^\pi(x) = E_{x,\pi}\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \overline{F}(p_t) dt\right],$$

where the subscripts of the expectation indicate that $x_0 = x$ and $p_t = \pi(x_t)$. The optimal discounted revenue is given by $J_\lambda^*(x) = \sup_{\pi \in \Pi_\lambda} J_\lambda^\pi(x)$, and a policy π is said to be optimal if $J_\lambda^* = J_\lambda^\pi$.

Suppose now that the arrival rate λ is not known, but rather, the vendor starts with a prior on λ that is a finite mixture of Gamma distributions. A kth order mixture of this type is parameterized by vectors a_0, $b_0 \in \mathbb{R}_+^k$ and a vector of k weights $w_0 \in \mathbb{R}_+^k$ that sum to unity. Such a prior is given by:

$$\Pr[\lambda \in d\lambda] = \sum_k w_k \frac{{b_{0,k}}^{a_{0,k}} \lambda^{a_{0,k}-1} e^{-\lambda b_{0,k}}}{\Gamma(a_{0,k})} d\lambda,$$

where Γ denotes the Gamma-function: $\Gamma(x) = \int_{s=0}^{\infty} s^{x-1} e^{-s} ds$. The expectation and variance are $E[\lambda] = \sum_k w_k a_{0,k}/b_{0,k} \sim \mu_0$ and $\mathrm{Var}[\lambda] = \sum_k w_k a_{0,k}(a_{0,k}+1)/b_{0,k}^2 - \mu_0^2$. Any prior on λ with a continuous, bounded density can be approximated to an arbitrary accuracy within such a family (see Dalal and Hall (1983)). Moreover, as we describe below, posteriors on λ continue to remain within this family rendering such a model parsimonious as well as relatively tractable.

The vendor revises his beliefs about λ as sales are observed. In particular, at time t, the vendor obtains a posterior that is a kth order mixture of Gamma distributions with parameters

$$a_{t,k} = a_{0,k} + n_t \quad \text{and} \quad b_{t,k} = b_{0,k} + \int_{\tau=0}^{t} \overline{F}(p_\tau) d\tau.$$

and weights

$$w_{t,k} = w_{0,k} \frac{\Pr(n_0^t | p_0^t, w_{0,k} = 1)}{\Pr(n_0^t | p_0^t)}.$$

Note that the vendor does not observe all customer arrivals but only those that result in sales. Further, lowering price results in more frequent sales and therefore more accurate estimation of the demand rate.

We consider pricing policies π that are measurable real-valued functions of the inventory level and arrival rate distribution parameters. As a convention we require that $\pi(0, a, b, w) = \infty$ for all arrival rate distribution parameters a, b and w. We denote the domain by $\mathcal{S} = \mathbb{N} \times \mathbb{R}_+^k \times \mathbb{R}_+^k \times \mathbb{R}_+^k$ and the set of policies by Π. Let $z_t = (x_t, a_t, b_t, w_t)$. A vendor who employs pricing policy $\pi \in \Pi$ sets price according to $p_t = \pi(z_t)$ and receives expected discounted revenue

$$J^\pi(z) = E_{z,\pi}\left[\int_{t=0}^{\tau_0} e^{-\alpha t} p_t \lambda \overline{F}(p_t) dt\right],$$

where the subscripts of the expectation indicate that $z_0 = z$ and $p_t = \pi(z_t)$. Note that, unlike the case with known arrival rate, λ is a random variable in this expectation. The optimal discounted revenue is given by $J^*(z) = \sup_{\pi \in \Pi} J^\pi(z)$, and a policy π is said to be optimal if $J^* = J^\pi$.

2.3 Optimal Pricing

An optimal pricing policy can be derived from the value function J^*. The value function in turn solves the HJB equation which we develop in this section. Unfortunately direct solution of the HJB equation, either analytically or computationally, does not appear to be a feasible task and one must resort to heuristic policies. With an end to deriving such heuristic policies we characterize optimal solutions to problems with known and unknown arrival rates and discuss some of their properties.

2.3.1 The Case of a Known Arrival Rate

We begin with the case of a known arrival rate. For each $\lambda \geq 0$ and $\pi \in \Pi_\lambda$, define an operator H_λ by

$$(H_\lambda^\pi J)(x) = \lambda \overline{F}(\pi(x))(\pi(x) + J(x-1) - J(x)) - \alpha J(x).$$

Recall that $\pi(0) = \infty$. In this case, we interpret $\overline{F}(\pi(0))\pi(0)$ as a limit, and Assumption 1 (which ensures a finite, unique static revenue maximizing price) implies that $(H_\lambda^\pi J)(0) = -\alpha J(0)$. Further, we define the dynamic programming operator

$$(H_\lambda J)(x) = \sup_{\pi \in \Pi_\lambda} (H_\lambda^\pi J)(x).$$

It is easy to show that J_λ^* is the unique solution to the HJB Equation $H_\lambda J = 0$. The first order optimality condition for prices yields an optimal policy of the form

$$\pi_\lambda^*(x) = 1/\rho(\pi_\lambda^*(x)) + J_\lambda^*(x) - J_\lambda^*(x-1),$$

for $x > 0$. By Assumption 1 and the fact that $J_\lambda^*(x) \geq J_\lambda^*(x-1)$, the above equation always has a solution on $\mathbb{R}_+$.

Now, the HJB Equation implies the recursion

$$\alpha J_\lambda^*(x) = \begin{cases} \sup_{p\geq 0} \lambda \overline{F}(p)(p + J_\lambda^*(x-1) - J_\lambda^*(x)) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Assumption 1 guarantees that $\sup_{p\geq 0} \overline{F}(p)(p+c)$ is an increasing function of c on $\mathbb{R}_+$. This allows one to compute $J_\lambda^*(x)$ given $J_\lambda^*(x-1)$ via bisection. This offers an efficient algorithm that computes $J^*(0), J^*(1), \ldots, J^*(x)$ in x iterations. As a specific concrete example, consider the case where reservation prices are exponentially distributed with mean $r > 0$. We have the HJB equation:

$$\alpha J_\lambda^*(x) = \begin{cases} \lambda r \exp(J_\lambda^*(x-1) - J_\lambda^*(x) - 1) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$J_\lambda^*(x) = W\left(\left(e^{-1}\lambda r/\alpha\right)\exp(J_\lambda^*(x-1))\right) \tag{2.1}$$

for $x > 0$, where $W(\cdot)$ is the Lambert W-function (the inverse of xe^x).

We note that a derivation of the optimal policy for the case of a known arrival rate may also be found in Araman and Caldentey (2005), among other sources.

2.3.2 The Case of an Unknown Arrival Rate

Let $S_{\tilde{x},\tilde{a},\tilde{b}} = \{(x,a,b,w) \in S : a + x = \tilde{a} + \tilde{x}, \tilde{b} \leq b, w \geq 0, 1'w = 1\}$ denote the set of states that might be visited starting at a state with $x_0 = \tilde{x}, a_0 = \tilde{a}, b_0 = \tilde{b}$. Let $\mathcal{J}$ denote the set of functions $J : \mathcal{S} \mapsto \Re$ such that $\sup_{z\in\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}} |J(z)| < \infty$ for all $\tilde{x}$ and $\tilde{b} > 0$ and that have bounded derivatives with respect to the third and fourth arguments. We define $\mu(z)$ to be the expectation for the prior on arrival rate in state z, so that $\mu(z) = \sum_k w_k a_k / b_k$.

For each policy $\pi \in \Pi$, we define an operator

$$(H^\pi J)(z) = \overline{F}(\pi(z))\left(\mu(z)\left(\pi(z) + J(z') - J(z)\right) + (DJ)(z)\right) - \alpha J(z),$$

where $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$, $z = (x, a, b, w)$ and $z' = (x-1, a+1, b, w')$. Here w' is defined according to $w'_k = (w_k a_k)/(b_k \mu(z))$, and D is a differential operator given by:

$$(DJ)(z) = \sum_k w_k \left(\mu(z) - a_k/b_k\right) \frac{d}{dw_k} J(z) + \frac{d}{db_k} J(z).$$

We now define the dynamic programming operator:

$$(HJ)(z) = \sup_\pi (H^\pi J)(z).$$

We then have that the value function J^* solves the HJB Equation in a sense stated precisely by the following Theorem. The proof is somewhat technical and not central to our exposition. It may be found in the appendix for the special case of a Gamma prior and exponential reservation prices which will be the primary context for our performance analysis in later sections.

Theorem 1. *The value function J^* is the unique solution in $\mathcal{J}$ to $HJ = 0$.*

The next Theorem, again proved in the appendix for Gamma priors and exponential reservation prices, offers a necessary and sufficient condition for optimality based on the HJB Equation.

Theorem 2. *A policy $\pi \in \Pi$ is optimal if and only if $H^\pi J^* = 0$.*

Under Assumption 1, we have that for each state z with $x > 0$ and $\mu(z) > 0$, there is a unique price, $\pi^*(z)$, that satisfies the first-order necessary condition for optimality, and it is given by the unique solution to

$$p = \frac{1}{\rho(p)} + J^*(z) - J^*(z') - \frac{1}{\mu(z)}(DJ^*)(z), \tag{2.2}$$

where $z = (x, a, b, w)$ and $z' = (x-1, a+1, b, w')$, w' being defined according to $w'_k = (w_k a_k)/(b_k \mu(z))$

Unfortunately there is no known analytical solution to the HJB Equation when the arrival rate is unknown, even for special cases such as a Gamma or two-point prior with exponential reservation prices. Further, grid-based numerical solution methods require discommoding computational resources and generate strategies that are difficult to interpret. As such, simple effective heuristics are desirable.

2.4 Estimate, Then Optimize

Aviv and Pazgal (2005) studied a *certainty equivalent* heuristic which at each point in time computes the conditional expectation of the arrival rate, conditioned on observed sales data, and prices as though the arrival rate is equal to this expectation. This is effectively ”estimate, then optimize“ for our problem. In our context, the price function for such a heuristic uniquely solves

$$\pi_{\mathrm{ce}}(z) = \frac{1}{\rho(\pi_{\mathrm{ce}}(z))} + J^*_{\mu(z)}(x) - J^*_{\mu(z)}(x-1),$$

for $x > 0$. The existence of a unique solution to this equation is guaranteed by Assumption 1. As derived in the preceding section, this is an optimal policy for the case where the arrival rate is known and equal to $\mu(z)$, which is the expectation of the arrival rate given a prior distribution with parameters a, b and w. The certainty equivalent policy is computationally attractive since J^*_λ is easily computed numerically (and in some cases, even analytically) as discussed in the previous section. As one would expect, prices generated by this heuristic increase as the inventory x decreases. However, arrival rate uncertainty bears no influence on price – the price only depends on the arrival rate distribution through its expectation $\mu(z)$. Hence, this pricing policy is unlikely to appropriately address information acquisition.

2.5 The Greedy Heuristic

We now present another heuristic which was recently proposed by Araman and Caldentey (2005) and does account for arrival rate uncertainty. To do so, we first introduce the notion of a *greedy policy*. A policy π is said to be *greedy* with respect to a function J if $H^\pi J = HJ$. The first-order necessary condition for optimality and Assumption 1 imply that the greedy price is given by the solution to

$$\pi(z) = \left(\frac{1}{\rho(\pi(z))} + J(z) - J(z') - \frac{1}{\mu(z)}(DJ)(z) \right)^+,$$

for $z = (x, a, b, w)$ with $x > 0$ and $z' = (x - 1, a + 1, b, w')$ with $w'_k = (w_k a_k)/(b_k \mu)$.

Perhaps the simplest approximation one might consider to $J^*(z)$ is $J^*_{\mu(z)}(x)$, the value for a problem with known arrival rate $\mu(z)$. One troubling aspect of this approximation is that it ignores the variance (as also higher moments) of the arrival rate. The alternative approximation proposed by Araman and Caldentey takes variance into account. In particular their heuristic employs a greedy policy with respect to the approximate value function $\tilde{J}$ which takes the form

$$\tilde{J}(z) = E[J^*_\lambda(x)],$$

where the expectation is taken over the random variable λ, which is drawn from a Gamma mixture with parameters a, b and w. $\tilde{J}(z)$ can be thought of as the expected optimal value if λ is to be observed at the next time instant.

Since it can only help to know the value of λ, $J^*_\lambda(x) \geq E[J^*(z)|\lambda]$. Taking expectations of both sides of this inequality, we see that $\tilde{J}$ is an upper bound on J^*. The approximation $J^*_{\mu(z)}(x)$ is a looser upper bound on $J^*(z)$. This follows from concavity of J^*_λ in λ, which is established in the proof of the following Lemma whose proof may be found in the appendix:

Lemma 1. *For all $z \in \mathcal{S}, \alpha > 0$*

$$J^*(z) \leq \tilde{J}(z) \leq J^*_{\mu(z)}(x) \leq \frac{\overline{F}(p^*)p^*\mu(z)}{\alpha}.$$

where p^* is the static revenue maximizing price.

The greedy price in state z is thus the solution to

$$\pi_{\mathrm{gp}}(z) = \left(\frac{1}{\rho(\pi_{\mathrm{gp}}(z))} + \tilde{J}(z) - \tilde{J}(z') - \frac{1}{\mu(z)}(D\tilde{J})(z) \right)^+,$$

for $z = (x, a, b, w)$ with $x > 0$ and $z' = (x-1, a+1, b, w')$ with $w'_k = (w_k a_k)/(b_k \mu(z))$.

We have observed through computational experiments (see Section 6) that when reservation prices are exponentially distributed and the vendor begins with a Gamma prior with scalar parameters a and b, greedy prices can increase *or* decrease with the inventory level x, keeping a and b fixed. This is clearly not optimal behavior.

2.6 Decay Balancing

In this section, we describe *decay balancing*, a new heuristic which will be the primary subject of the remainder of this chapter. To motivate the heuristic, we start by deriving an alternative characterization of the optimal pricing policy. The HJB Equation yields

$$\max_{p \geq 0} \overline{F}(p) \left(\mu(z) \left(p + J^*(z') - J^*(z)\right) + (DJ^*)(z)\right) = \alpha J^*(z),$$

for all $z = (x, a, b, w)$ and $z' = (x-1, a+1, b, w')$, with $x > 0$ and $w'_k = (w_k a_k)/(b_k \mu(z))$. This equation can be viewed as a balance condition. The right hand side represents the rate at which value decays over time; if the price were set to infinity so that no sales could take place for a time increment dt but an optimal policy is used thereafter, the current value would become $J^*(x) - \alpha J^*(x) dt$. The left hand side represents the rate at which value is generated from both sales and learning. The equation requires these two rates to balance so that the net value is conserved.

Note that the first order optimality condition implies that if $J(z') - J(z) + \frac{1}{\mu(z)}(DJ)(z) < 0$ (which must necessarily hold for $J = J^*$),

$$\frac{\overline{F}(p^*)}{\rho(p^*)}\mu(z) = \max_{p \geq 0} \overline{F}(p) \left(\mu(z) \left(p + J(z') - J(z)\right) + (DJ)(z)\right),$$

if p^* attains the maximum in the right hand side. Interestingly, the maximum depends on J only through p^*. Hence, the balance equation can alternatively be written in the following simpler form:

$$\frac{\overline{F}(\pi^*(z))}{\rho(\pi^*(z))}\mu(z) = \alpha J^*(z).$$

which implicitly characterizes π^*.

This alternative characterization of π^* makes obvious two properties of optimal prices. Note that $\overline{F}(p)/\rho(p)$ is decreasing in p. Consequently, holding a, b and w fixed, as x decreases, $J^*(z)$ decreases and therefore $\pi^*(z)$ increases. Further, since $J^*(z) \leq J^*_{\mu(z)}(x)$, we see that for a fixed inventory level x and expected arrival rate $\mu(z)$, the optimal price in the presence of uncertainty is higher than in the case where the arrival rate is known exactly.

Like greedy pricing, the decay balancing heuristic relies on an approximate value function. We will use the same approximation $\tilde{J}$. But instead of following a greedy policy with respect to $\tilde{J}$, the decay balancing approach chooses a policy π_{db} that satisfies the balance condition:

$$\frac{\overline{F}(\pi_{\mathrm{db}}(z))}{\rho(\pi_{\mathrm{db}}(z))}\mu(z) = \alpha \tilde{J}(z),$$

with the decay rate approximated using $\tilde{J}(z)$. The following Lemma guarantees that the above balance equation always has a unique solution so that our heuristic is well defined. The proof is omitted; it is a straightforward consequence of Assumption 1 and the fact that $\frac{\overline{F}(p^*)}{\alpha\rho(p^*)}\mu(z) \geq \tilde{J}(z) \geq J^*(z) = \frac{\overline{F}(\pi^*(z))}{\alpha\rho(\pi^*(z))}\mu(z)$ where p^* is the static revenue maximizing price.

Lemma 2. *For all $z \in \mathcal{S}$, there is a unique $p \geq 0$ such that $\frac{\overline{F}(p)}{\rho(p)}\mu(z) = \alpha\tilde{J}(z)$.*

Unlike certainty equivalent and greedy pricing, uncertainty in the arrival rate and changes in inventory level have the correct directional impact on decay balancing prices. Holding a, b and w fixed, as x decreases, $\tilde{J}(z)$ decreases and therefore $\pi_{\mathrm{db}}(z)$ increases. Holding x and the expected arrival rate $\mu(z)$ fixed, $\tilde{J}(z) \leq J^*_{\mu(z)}(x)$, so that the decay balance price with uncertainty in arrival rate is higher than when the arrival rate is known with certainty.

It is frequently possible to express the decay balance price at a state z explicitly, as a function of $\tilde{J}(z)$. Table 1 lists formulas for the decay balance price for several reservation price distributions. This list includes iso-elastic distributions (of the form $\overline{F}(p) = cp^{-\gamma}$) which are frequently used to model reservation prices, but do not satisfy Assumption 1 since they are improper. One may address this technical difficulty by restricting attention to prices in $(\epsilon, \infty]$, so that $\overline{F}(p) = \epsilon^{\gamma}p^{-\gamma}$). Such distributions

Table 2.1: Decay Balance Price Formulas

Distribution	$\overline{F}(p)$	$\pi_{\text{db}}(z)$	**Remarks**
Exponential	$\exp(-p/r)$	$r \log\left(\frac{r\mu(z)}{\alpha\tilde{J}(z)}\right)$	$r > 0$
Logit	$2\frac{\exp(-p/r)}{1+\exp(-p/r)}$	$r \log\left(\frac{2r\mu(z)}{\alpha\tilde{J}(z)}\right)$	$r > 0$
Iso-Elastic	$\epsilon^\gamma p^{-\gamma}$	$\max\left(\left(\frac{\mu(z)\epsilon^\gamma}{\gamma\alpha\tilde{J}(z)}\right)^{\frac{1}{\gamma-1}}, \epsilon\right)$	$\gamma > 2, p \geq \epsilon$

do not satisfy Assumption 1 either since they have no support on $[0, \epsilon)$. Nonetheless, for $\gamma > 2$, it is possible to derive a decay balance equation (which takes the form $\gamma^{-1}\epsilon^\gamma(\pi_{\text{db}}(z))^{1-\gamma}\mu(z) = \min\left(\alpha\tilde{J}(z), \mu(z)\gamma^{-1}\epsilon\right)$) and extend our analysis to such distributions without difficulty.

2.7 Bounds on Performance Loss

For the decay balancing price to be a good approximation to the optimal price at a particular state, one requires only a good approximation to the value function at that state (and *not* its derivatives). This section characterizes the quality of our approximation to J^* and uses such a characterization to ultimately bound the performance loss incurred by decay balancing relative to optimal pricing. Our analysis will focus primarily on the case of a Gamma prior and exponential reservation prices (although we will also provide performance guarantees for other types of reservation price distributions). We will show that in this case, decay balancing captures at least 33.3% of the expected revenue earned by the optimal algorithm for all choices of $x_0 > 1, a_0 > 0, b_0 > 0, \alpha > 0$ and $r > 0$ when reservation prices are exponentially distributed with mean $r > 0$. Such a bound is an indicator of robustness across all parameter regimes. Decay balancing is the first heuristic for problems of this type for which a uniform performance guarantee is available. Further, by allowing for a dependence on the number of sales we can guarantee that after four sales decay balancing achieves a level of performance that is within 80% of optimal thereafter.

Before we launch into the proof of our performance bound, we present an overview of the analysis. Since our analysis will focus on a gamma prior we will suppress the state variable w in our notation, and a and b will be understood to be scalars. Without loss of generality, we will restrict attention to problems with $\alpha = e^{-1}$; in particular, the value function exhibits the following invariance where the notation $J^{*,\alpha}$ makes the dependence on α explicit (see the appendix for a proof):

Lemma 3. *For all* $z \in \mathcal{S}, \alpha > 0$, $J^{*,\alpha}(z) = J^{*,1}(x, a, \alpha b)$.

As a natural first step, we attempt to find upper and lower bounds on $\pi_{\text{db}}(z)/\pi^*(z)$, the ratio of the decay balancing price in a particular state to the optimal price in that state. We are able to show that $1 \geq J^*(z)/\tilde{J}(z) \geq 1/\kappa(a)$ where $\kappa(\cdot)$ is a certain decreasing function. Under an additional assumption on reservation prices, this suffices to establish that:

$$\frac{1}{\kappa(a)} \leq \frac{\pi_{\text{db}}(z)}{\pi^*(z)} \leq 1$$

By considering a certain system under which revenue is higher than the optimal revenue, we then use the bound above and a dynamic programming argument to show that:

$$\frac{1}{\kappa(a)} \leq \frac{J^{\pi_{\text{db}}}(z)}{J^*(z)} \leq 1$$

where $J^{\pi_{\text{db}}}(z)$ denotes the expected revenue earned by the decay balancing heuristic starting in state z. If z is a state reached after i sales then $a = a_0 + i > i$, so that the above bound guarantees that the decay balancing heuristic will demonstrate performance that is within a factor of $\kappa(i)$ of optimal moving forward after i sales.

Our general performance bound can be strengthened to a uniform bound in the special case of exponential reservation prices. In particular, a coupling argument that uses a refinement of the general bound above along with an analysis of the maximal loss in revenue up to the first sale for exponential reservation prices, establishes the uniform bound

$$\frac{1}{3} \leq \frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \leq 1$$

We begin our proof with a simple dynamic programming result that we will have several opportunities to use. The proof is essentially a consequence of Dynkin's formula and may be found in the appendix:

Lemma 4. *Let* $J \in \mathcal{J}$ *satisfy* $J(0,a,b) = 0$. *Let* $\tau = \inf\{t : J(z_t) = 0\}$. *Let* $z_0 \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$. *Then,*

$$E\left[\int_0^\tau e^{-\alpha t} H^\pi J(z_t) dt\right] = J^\pi(z_0) - J(z_0)$$

Let $J : \mathbb{N} \to \mathbb{R}$ *be bounded and satisfy* $J(0) = 0$. *Let* $\tau = \inf\{t : J(x_t) = 0\}$. *Let* $x_0 \in \mathbb{N}$. *Then,*

$$E\left[\int_0^\tau e^{-\alpha t} H_\lambda^\pi J(x_t) dt\right] = J_\lambda^\pi(x_0) - J(x_0)$$

2.7.1 Decay Balancing Versus Optimal Prices

As discussed in the preceding outline, we will establish a lower bound on $J^*(z)/\tilde{J}(z)$ in order to establish a lower bound on $\pi_{\mathrm{db}}(z)/\pi^*(z)$. Let $J^{nl}(z)$ be the expected revenue garnered by a pricing scheme that does not learn, upon starting in state z. Delaying a precise description of this scheme for just a moment, we will have $J^{nl}(z) \leq J^*(z) \leq \tilde{J}(z) \leq J^*_{a/b}(x)$. It follows that $J^{nl}(z)/J^*_{a/b}(x) \leq J^*(z)/\tilde{J}(z)$, so that a lower bound on $J^{nl}(z)/J^*_{a/b}(x)$ is also a lower bound on $J^*(z)/\tilde{J}(z)$. We will focus on developing a lower bound on $J^{nl}(z)/J^*_{a/b}(x)$.

Upon starting in state z, the "no-learning" scheme assumes that $\lambda = a/b = \mu$ and does not update this estimate over time. Assuming we begin with a prior of mean μ, such a scheme would use a pricing policy given implicitly by:

$$\pi^{nl}(z) = \pi^*_\mu(x) = 1/\rho(\pi^*_\mu(x)) + J^*_\mu(x) - J^*_\mu(x-1). \tag{2.3}$$

Some simplification yields

$$H_\lambda^{\pi^{nl}} J^*_\mu(x) = (\lambda/\rho - 1)\alpha J^*_\mu(x).$$

The following two results are then essentially immediate consequences of Lemma 4:

Lemma 5. *If* $\lambda < \mu$, $J_\lambda^{\pi^{nl}}(x) \geq \lambda/\mu J_\mu^*(x)$ *for all* $x \in \mathbb{N}$.

Lemma 6. *If* $\lambda \geq \mu$, $J_\lambda^{\pi^{nl}}(x) \geq J_\mu^*(x)$ *for all* $x \in \mathbb{N}$.

Armed with these two results we can establish a lower bound on $J^{nl}(z)/J_{a/b}^*(x)$:

Theorem 3. *For all* $z \in \mathcal{S}$,

$$\frac{J^{nl}(z)}{J_{a/b}^*(x)} \geq \frac{\Gamma(a+1) - \Gamma(a+1,a) + a\Gamma(a,a)}{a\Gamma(a)} \equiv 1/\kappa(a)$$

Proof: Proof: We have

$$\begin{aligned} J^{nl}(z) &= E_\lambda\left[J_\lambda^{\pi^{nl}}(x)\right] \\ &\geq E_\lambda\left[1_{\lambda<\mu}\lambda/\mu J_\mu^*(x) + 1_{\lambda\geq\mu} J_\mu^*(x)\right] \\ &= \frac{\Gamma(a+1) - \Gamma(a+1,a) + a\Gamma(a,a)}{a\Gamma(a)} J_\mu^*(x) \end{aligned}$$

where the inequality follows from the two preceding Lemmas. $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function and is given by $\Gamma(x,y) = \int_y^\infty s^{x-1}e^{-s}ds$. □

The decay balance equation allows one to use the above bound on the quality of our approximation $\tilde{J}$ to compute bounds on the decay balance price relative to the optimal price at a given state. Within the class of reservation price distributions specified by the Assumption 1 and the following assumption, this bound depends merely on the state variable a:

Assumption 2.

1. $\frac{\rho(p)}{\overline{F}(p)}$ *is a differentiable, convex function of* p *with support* $\mathbb{R}_+$.

2. *There exists a unique static revenue maximizing price* $p^* > 0$ *with* $\frac{d}{dp}\frac{\rho(p)}{\overline{F}(p)}\Big|_{p=p^*} \geq 1/\overline{F}(p^*)p^{*2}$.

While Assumption 2 is not a typical assumption within the revenue management literature, it is easily verified that it includes several interesting distributions with support

Table 2.2: Decay Balance vs. Optimal Prices

Distribution	$\overline{F}(p)$	$\pi^{\mathrm{db}}(z)/\pi^*(z) \geq$
Exponential	$\exp(-p/r)$	$\frac{1}{1+\log\kappa(a)}$
Logit	$2\frac{\exp(-p/r)}{1+\exp(-p/r)}$	$\frac{1.27}{1.27+\log\kappa(a)}$
Iso-Elastic	$\epsilon^\gamma p^{-\gamma}$	$\left(\frac{1}{\kappa(a)}\right)^{\frac{1}{\gamma-1}}$

on $\mathbb{R}_+$ including the exponential and the Logit. It does not include constant elasticity distributions, but the bound in Corollary 1 can nevertheless be verified directly for such distributions. For reservation price distributions satisfying Assumptions 1 and 2 the following is a simple corollary to Theorem 3 and is proved in the appendix.

Corollary 1. *For all $z \in \mathcal{S}$, and reservation price distributions satisfying Assumptions 1 and 2*

$$\frac{1}{\kappa(a)} \leq \frac{\pi_{\mathrm{db}}(z)}{\pi^*(z)} \leq 1$$

The lower bound in Corollary 1 can often be strengthened for specific reservation price distributions (via a finer analysis of their respective decay balance equations). Table 2 lists the bounds that can be derived for exponential, Logit and iso-elastic distributions.

2.7.2 An Upper Bound on Performance Loss

We now proceed with proving a general performance bound for reservation price distributions satisfying Assumptions 1 and 2 . In particular, we will establish a lower bound on $J^{\pi_{\mathrm{db}}}(z)/J^*(z)$ that will depend on the coefficient of variation of the prior on λ, $1/\sqrt{a}$.

Let

$$R^{\mathrm{db}}(z) = \sum_{k:t_k \leq \tau} e^{-e^{-1}t_k} \pi_{\mathrm{db}}(z_{t_k^-})$$

be the revenue under the decay balancing policy for a particular sample path of the sales process, starting in state z and define

$$R^{\mathrm{ub}}(z) = \sum_{k:t_k \le \tau} e^{-e^{-1}t_k} \pi^*(z_{t_k^-}).$$

This describes a system whose state evolution is identical to that under the decay balancing policy but whose revenues on a sale correspond to those that would be earned if the price set prior to the sale was that of the optimal pricing algorithm.

Of course, $J^{\pi_{\mathrm{db}}}(z) = E_z[R^{\mathrm{db}}(z)]$. Define $J^{\mathrm{ub}}(z) = E_z[R^{\mathrm{ub}}(z)]$, where the expectation is over $\{t_k\}$ and assumes that an arriving consumer at time t_k makes a purchase with probability $\overline{F}(\pi_{\mathrm{db}}(z_{t_k}))$. That is, the expectation is understood to be according to the dynamics of the system controlled by π_{db}. The following result should be intuitive given our construction of the upper-bounding system and the fact that since $\pi_{\mathrm{db}}(z) \le \pi^*(z)$, the probability that a customer arriving in state z chooses to purchase is higher in the system controlled by the decay balancing policy. The proof uses the monotonicity of the dynamic programming operator and may be found in the appendix.

Lemma 7. *For all $z \in \mathcal{S}$, and reservation price distributions satisfying Assumptions 1 and 2,*

$$J^{\mathrm{ub}}(z) \ge J^*(z)$$

Now observe that since $\kappa(a)$ is decreasing in a, we have from Corollary 1 that

$$\frac{1}{\kappa(a)} \le \frac{R^{\mathrm{db}}(z)}{R^{\mathrm{ub}}(z)} \le 1.$$

Taking expectations, and employing Lemma 7, we then immediately have:

Theorem 4. *For all $z \in \mathcal{S}$, and reservation price distributions satisfying Assumptions 1 and 2,*

$$\frac{1}{\kappa(a)} \le \frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \le 1$$

We note that upon using the reservation price distribution specific bounds on $\pi^{\mathrm{db}}(z)/\pi^*(z)$ in Table 2 as opposed to the general bound in Corollary 1, one recovers correspondingly sharper, reservation price distribution specific bounds on $\frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)}$. See Figure 2.1 for an illustration of these bounds.

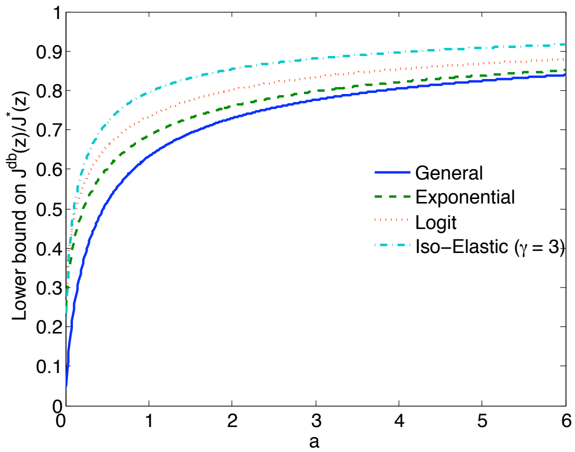

Figure 2.1: Lower bound on Decay Balancing performance

It is worth pausing to reflect on the bound we have just established. Our performance bound does *not* depend on x, b, α or the parameters of the reservation price distribution. It is valid for a number of reservation price distributions including the exponential, Logit, and iso-elastic (or power law) distributions. There are several ways to interpret the bound. For one, since the coefficient of variation of a Gamma prior with parameters a and b is given by $1/\sqrt{a}$, the bound illustrates how decay balancing performance approaches optimal performance as the coefficient of variation of the initial prior is decreased. For example, for coefficients of variation smaller than 0.5, decay balancing is guaranteed to capture at least 80% of the optimal performance. The fact that after after i sales, we must be in a state with $a > i$, yields a second interpretation: after merely four sales, we are guaranteed performance that is within 80% of optimal from that point onwards. Next, we further specialize our attention to exponential reservation price distributions and present a uniform performance guarantee for that case.

2.7.3 A Uniform Performance Bound for Exponential Reservation Prices

We now consider the case of exponentially distributed prices. In particular, we have $\overline{F}(p) = \exp(-p/r)$ where $r > 0$. In light of the following Lemma (where the notation

$J^{*,r}$ makes the dependence of the value function on r explicit), we can assume without loss that the mean reservation price, r, is 1; see the appendix for a proof.

Lemma 8. *For all $z \in \mathcal{S}, r > 0$, $J^{*,r}(z) = rJ^{*,1}(z)$.*

Using the stronger bounds on $\pi^{\mathrm{db}}(z)/\pi^*(z)$ for exponential reservation prices from Table 2, it is easy to see that the result of Theorem 4 can be strengthened further to:

Theorem 5. *Assume $\overline{F}(p) = \exp(-p)$. Then, for all $z \in \mathcal{S}$,*

$$\frac{1}{1 + \log \kappa(a)} \leq \frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \leq 1$$

Our proof of a uniform performance bound will use Theorem 5 along with a coupling argument to bound performance loss up to the time of the first sale.

Begin by considering the following coupling (A superscript "db" on a variable indicates that the variable is relevant to a system controlled by the π_{db} policy): For an arbitrary policy $\pi(\cdot) \in \Pi$, the sales processes n_t^{db} and n_t^{π} are coupled in the following sense: Denote by $\{t_k\}$ the points of the Poisson process corresponding to customer *arrivals* (not sales) to both systems. Assume $\pi_{\mathrm{db}t_k^-} \leq \pi_{t_k^-}$. Then a jump in n^π at time t_k can occur if and only if a jump occurs in n^{db} at time t_k. Further, conditioned on a jump in n^{db} at t_k, the jump in n^π occurs with probability $\exp(-(\pi_{t_k^-} - \pi_{\mathrm{db}t_k^-}))$. The situation is reversed if $\pi_{\mathrm{db}t_k^-} > \pi_{t_k^-}$. Let τ denote the time of the first sale for the π_{db} system i.e. $\tau = \inf\{t : n_t^{\mathrm{db}} = 1\}$. In the context of this coupling consider the optimal (i.e. π^*) and π_{db} controlled systems. We then have:

Lemma 9. *For all $z \in \mathcal{S}$,*

$$J^*(z|\tau) \leq$$
$$e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\mathrm{db}})}\left[\pi^* + J^*(x-1, a+1, b_\tau^{\mathrm{db}})\right] + (1 - e^{-(\pi^*-\pi_{\mathrm{db}})})J^*(x, a+1, b_\tau^{\mathrm{db}})\right)$$

where $\pi^* = \pi^*(x, a, b_\tau^*)$ and $\pi_{\mathrm{db}} = \pi_{\mathrm{db}}(x, a, b_\tau^{\mathrm{db}})$.

The result above is essentially a consequence of the fact that it is never the case that the π^* controlled system sells it's first item before the π_{db} system, and moreover, that conditioning on τ, and the information available in both systems up to τ^- yields a posterior with shape parameter $a+1$ and scale parameter b_τ^{db}. A formal proof may be found in the appendix.

We will also find the following technical Lemma, whose proof is in the appendix, useful:

Lemma 10. *For $x > 1, a > 1, b > 0$, $J^*(x,a,b) \leq 2.05 J^*(x-1,a,b)$.*

The result above is intuitive; it would follow, for example, from decreasing returns to an additional unit of inventory. It is unfortunate that we aren't able to show such a "decreasing returns" property directly. In particular, should such a property hold for general reservation price distributions, the uniform bound we establish in the following theorem could be extended to reservation price distributions satisfying Assumptions 1 and 2 as also Iso-elastic reservation price distributions. We are now poised to prove a uniform (over $x > 1$) performance bound for our pricing scheme:

Theorem 6. *For all $z \in \mathcal{S}$ with $x > 1$,*

$$\frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \geq 1/3.$$

Proof: Proof: In Lemma 9 we showed:

$$J^*(z) \leq$$
$$E\left[e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\mathrm{db}})}\left[\pi^* + J^*(x-1,a+1,b_\tau^{\mathrm{db}})\right] + (1-e^{-(\pi^*-\pi_{\mathrm{db}})})J^*(x,a+1,b_\tau^{\mathrm{db}})\right)\right]$$

Now,

$$\begin{aligned}
&e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\mathrm{db}})}\left[\pi^* + J^*(x-1,a+1,b_\tau^{\mathrm{db}})\right] + (1-e^{-(\pi^*-\pi_{\mathrm{db}})})J^*(x,a+1,b_\tau^{\mathrm{db}})\right) \\
&\leq e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\mathrm{db}})}\pi^* + J^*(x,a+1,b_\tau^{\mathrm{db}})\right) \\
&\leq e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\mathrm{db}})}\pi^* + 2.05\ J^*(x-1,a+1,b_\tau^{\mathrm{db}})\right) \\
&\leq e^{-e^{-1}\tau}\left(\pi_{\mathrm{db}} + 2.05(1+\log\kappa(a+1))J^{\pi_{\mathrm{db}}}(x-1,a+1,b_\tau^{\mathrm{db}})\right) \\
&\leq e^{-e^{-1}\tau}2.05(1+\log\kappa(a+1))\left(\pi_{\mathrm{db}} + J^{\pi_{\mathrm{db}}}(x-1,a+1,b_\tau^{\mathrm{db}})\right)
\end{aligned}$$

where the first inequality is because J^* is non-decreasing in x. The second inequality follows from Lemma 10. The third inequality follows from the fact that $\pi^* \geq \pi_{\mathrm{db}} \geq 1$ so that $\pi^* e^{-\pi^*} \leq \pi_{\mathrm{db}} e^{-\pi_{\mathrm{db}}}$ and from Theorem 4. Finally, taking expectations of both sides we get:

$$\frac{J^{\pi_{\mathrm{db}}}(z)}{J^*(z)} \geq \frac{1}{2.05(1+\log\kappa(1))} \geq 1/3.$$

□

2.8 Computational Study

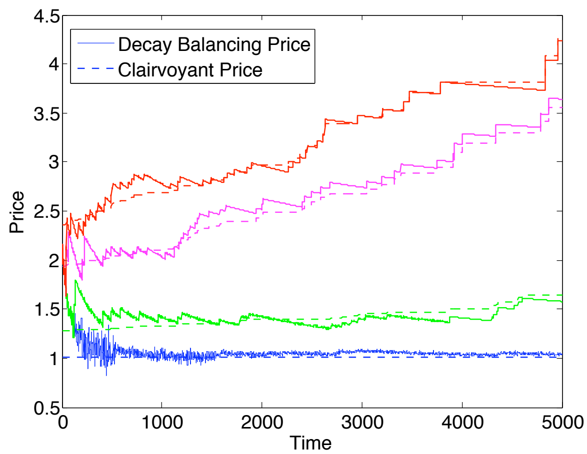

Figure 2.2: Decay Balancing price trajectories vs. Clairvoyant price trajectories

This section will present computational results that highlight the performance of the decay balancing heuristic. We consider exponentially distributed reservation prices and gamma priors. Further, we will only consider problem instances where $\alpha = e^{-1}$ and $r = 1$; in light of the invariances discussed in Lemmas 3 and 8, this is not restrictive.

Consider a "clairvoyant" algorithm that has access to the realization of λ at $t = 0$ and subsequently uses the pricing policy π^*_λ. The expected revenue garnered by such a pricing policy upon starting in state z is simply $E[J^*_\lambda(x)] = \tilde{J}(z)$ which, by Lemma 1, is an upper bound on $J^*(z)$. Our first experiment measures the average revenue earned using decay balancing with that earned using such a clairvoyant algorithm. The results are summarized in Table 1; λ here is drawn from a distribution with shape parameter

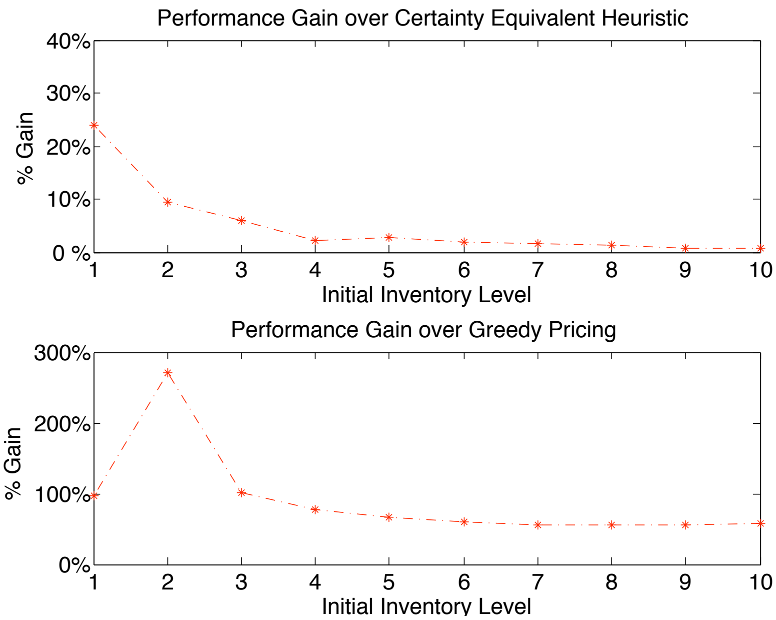

Figure 2.3: Performance gain over the Certainty Equivalent and Greedy Pricing heuristics at various inventory levels

$a = 0.04$ and scale parameter $b = 0.001$ which corresponds to a mean of 40 and a coefficient of variation of 5. These parameter values are representative of a high level of uncertainty in λ. As is seen in Table 1, the performance of the decay balancing heuristic is surprisingly close to that of the clairvoyant algorithm.

We next record price trajectories realized using decay balancing for several widely different realizations of λ. We compare these trajectories with the prices set by a clairvoyant algorithm. In particular, for a given state trajectory realized under the use of the decay balancing heuristic, z^t, we plot $\pi_{\rm db}(z_t)$ along side $\pi^*_\lambda(z_t)$. The decay balancing heuristic begins with a prior with mean 50 and coefficient of variation 4.47. Figure 2.2 plots decay balancing price trajectories for $\lambda = 10$, 50, 200, and 400 along side the optimal clairvoyant price for each trajectory. From lowest to highest, the curves correspond respectively to $\lambda = 10$, 50, 200, and 400. First, it is satisfying to note that the decay balancing price tracks the clairvoyant price closely. Second, we note that each of the price trajectories have relatively stable trends which is desirable from an operational perspective. In contrast, price trajectories realized by a non-parametric pricing algorithm

Table 2.3: Performance vs. a Clairvoyant algorithm

Inventory Level	$x_0 = 1$	$x_0 = 2$	$x_0 = 5$	$x_0 = 10$	$x_0 = 20$	$x_0 = 40$
Performance Gain i.e. $(J^{\pi_{\rm db}} - \tilde{J})/\tilde{J}$	-13%	-10%	-6%	-3.7%	-2%	-0.5%

such as Besbes and Zeevi (2006) will vary substantially over periods corresponding to price-experimentation since such an algorithm needs to experiment with essentially every possible price. Finally, these trajectories are also an indicator of robustness to our choice of prior; $\pi_{\rm db}$ when $\lambda = 400$ (the topmost solid curve) tracks π^*_{400} (the topmost dashed curve) in spite of starting with a prior mean that is smaller than the actual arrival rate by an order of magnitude!

We finally turn to comparing the performance of decay balancing against that of the certainty equivalent and greedy policies. We first compare gain in performance over the certainty equivalent heuristic and the greedy policy for inventory levels between 1 and 10. We set $a = 0.05$ and $b = 0.001$ for all heuristics, corresponding to a mean of 50 and a coefficient of variation of 4.47. Now $E[\lambda]/\alpha$ can be interpreted as the expected number of potential customers encountered over the sales season. Consequently, $E[\lambda]/\alpha$ divided by the inventory level may be viewed as the "load factor" here. Figure 2.3 indicates that we offer a substantial gain in performance over the certainty equivalent and greedy pricing heuristics at higher load factors (approximately 2 for these experiments) which is representative of a regime wherein judiciously managing inventory levels is crucial. The greedy policy performs particularly poorly. In addition, as discussed earlier, that policy exhibits qualitative behavior that is clearly suboptimal: for a problem with mean reservation price 1 and discount factor e^{-1}, we compute $\pi_{\rm gp}(1, 0.1, 0.1)(= 1.26) < \pi_{\rm gp}(4, 0.1, 0.1)(= 1.61) > \pi_{\rm gp}(10, 0.1, 0.1)(= 1.25)$ so that, all other factors remaining the same, prices may increase or decrease with an increase in inventory level. Our gain in performance falls at lower load factors. This is not surprising; intuitively, the control problem at hand is simpler at low load factors since we are essentially allowed to sacrifice a few units of inventory early on so as to learn quickly without incurring much of a penalty. We next present a lower bound on highest potential performance gain for various coefficients of variation of the initial prior on λ. See Figure 2.4 wherein the data point for each coefficient of variation c, corresponds to an experiment with $a = 1/c^2$, $b = 0.001$ (which corresponds to a mean of $1000/c^2$), and an inventory level of 1 and 2 for the certainty equivalent and greedy pricing heuristics respectively. These experiments indicate that the potential gain from using decay balancing increases with increasing uncertainty in λ, and that in fact the gain over certainty equivalence can be as much as a factor of 1.3, and that over the greedy policy can be as much as a factor of 3.

To summarize our computational experience, the performance of decay balancing is surprisingly competitive with that of a clairvoyant algorithm even at high levels of uncertainty, so that we suspect decay balancing is near optimal. In comparison with other available heuristics (namely certainty equivalent and greedy) decay balancing offers substantial performance gains especially at high load factors. Finally, the performance gains offered by decay balancing over these competing heuristics increase with increasing uncertainty in customer arrival rate.

2.9 Discussion and Conclusions

The dynamic pricing model proposed by Gallego and van Ryzin (1994) is central to a large body of the revenue management literature. This chapter considered an important extension to that model. In particular, we considered incorporating uncertainty in the customer arrival rate or "market response" which is without doubt important in many industries that practice revenue management, but whose effects are nonetheless ignored in "estimate, then optimize" style pricing schemes.

We proposed and analyzed a simple new heuristic for this problem: decay balancing. Decay balancing is computationally efficient and leverages the solution to problems with no uncertainty in market response. Our computational experiments (which focused on gamma priors and exponentially distributed reservation prices) suggest that decay balancing achieves near-optimal performance even on problems with high levels of uncertainty in market response. Pricing policies generated by decay balancing have the appealing structural property that, all other factors remaining the same, the price in the presence of uncertainty in market response is higher than the corresponding price with no uncertainty . This is reasonable from an operational perspective and is in fact a property possessed by the optimal policy. Uncertainty in market response could qualitatively induce two potential price movements: whereas on the one hand, the revenue manager may wish to price lower than the no-uncertainty price in an effort to eliminate uncertainty as quickly as possible, he may on the other hand want to hedge against the possibility that market response is in fact stronger than expected by pricing higher than the no-unceratinty price and protecting inventory for this eventuality. It is interesting to note that this second hedging effect dominates irrespective of the level of uncertainty. Finally, our analysis demonstrated a uniform performance guarantee for decay balancing when reservation prices are exponentially distributed, which is an indicator of robustness.

Two heuristics proposed for problems of this nature prior to our work were the certainty equivalent heuristic (by Aviv and Pazgal (2005)) and the greedy pricing heuristic (by Araman and Caldentey (2005)). Our computational results suggest that decay balancing offers significant performance advantages over these heuristics. These advantages are especially clear at high levels of uncertainty in market response which is arguably the regime of greatest interest. Decay balancing relies only on a good approximation to the value of an optimal policy at a given state. This is in contrast with greedy pricing that requires not only a good approximation to value but further to derivatives of value with respect to the scale parameter. At the same time, uncertainty in arrival rate and changes in inventory levels bear the appropriate directional impact on decay balancing prices: uncertainty in the arrival rate calls for higher prices than in corresponding situations with no uncertainty, while a decrease in inventory calls for an increase in prices. In contrast, uncertainty in the arrival rate has no impact on certainty equivalent prices while greedy prices can increase or decrease with decreasing inventory.

Our computational study and performance analysis were focused on exponentially distributed reservation prices and gamma priors, but we expect favorable performance for other distributions as well. In particular, the analysis of Theorem 3 can be extended to mixtures of Gamma priors yielding encouraging estimates on the quality of approximation provided by $\tilde{J}$. Since the decay balancing price at state z is likely to be a good approximation to the optimal price at z if $\tilde{J}(z)$ is a good approximation to $J^*(z)$, this suggests that decay balancing is likely to do a good job of approximating the optimal price for general reservation price distributions and priors on arrival rate, which in turn should lead to superior performance.

There is ample room for further work in the general area of pricing with uncertainty in market response and other factors that impact demand. One direction is considering more complex models. The next chapter makes a foray in this direction. In particular, that chapter will consider two interesting models closely related to the model we focused on here, that allows for the modeling of joint learning from sales at multiple locations and product "versioning" respectively. There are other models one might hope to consider. For example, the multi product model proposed by Gallego and van Ryzin (1997). Another potential direction is exploring new approximations to the value function beyond the approximation considered here and applying such approximations with either the greedy pricing or decay balancing heuristics. Finally, it would be interesting to extend the approaches in this chapter to problems with uncertainty in other factors that impact demand such as price elasticity.

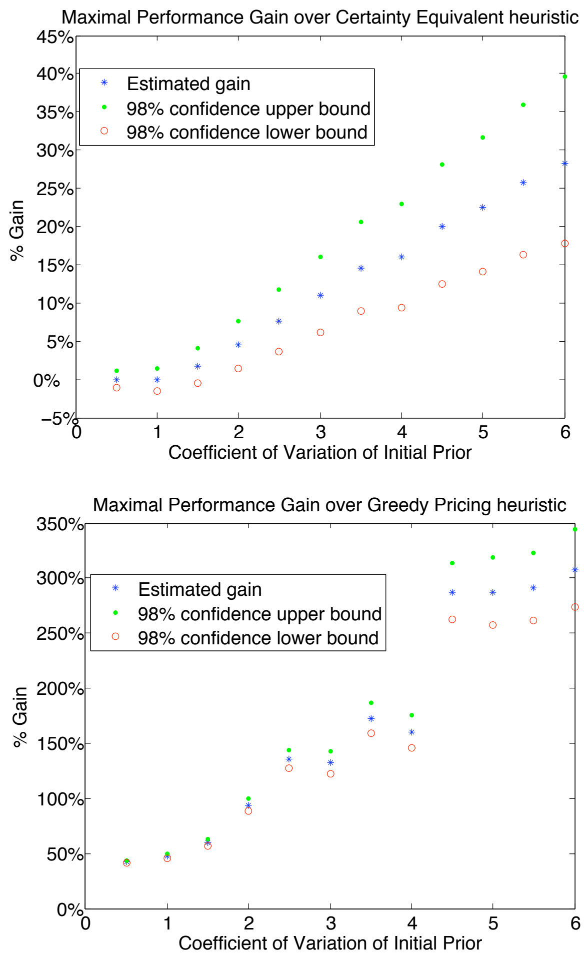

Figure 2.4: Maximal performance gain over the Certainty Equivalent Heuristic (left) and Greedy Heuristic (right) for various coefficients of variation

Chapter 3

Decay Balancing Extensions

The previous chapter developed a "decay balancing" heuristic in the context of a simple one product dynamic pricing problem. This chapter explores generalizing that heuristic to two closely related problems of dynamic pricing with demand uncertainty. In particular, we consider a problem of dynamic pricing across multiple stores with uninterchangeable inventories, where stores attempt to learn from each others' purchase data. We also consider a problem of dynamic pricing with product "versioning": a single product may be sold in multiple versions. Customers arrive at an uncertain rate and must choose from among these versions according to a pre-specified demand model. The revenue manager would like to set prices for all of the product versions over time so as to maximize expected discounted revenues.

3.1 Multiple Stores and Consumer Segments

We consider in this section a model with multiple stores and consumer segments. We do not attempt to extend our performance analysis to this more general model but instead present numerical experiments, the goal being to show that decay balancing demonstrates the same qualitative behavior as in the one store, one customer segment case we have studied to this point.

More formally, we consider a model with N stores and M consumer segments. Each store is endowed with an initial inventory $x_{0,i}$ for $i \in \{1, \ldots, N\}$. Customers from class

j, for $j \in \{1, \ldots, M\}$ arrive according to a Poisson process of rate λ_j where λ_j is a Gamma distributed random variable with shape parameter $a_{0,j}$ and scale parameter $b_{0,j}$. An arriving segment j customer considers visiting a single store and will consider store i with probability α_{ij}. Consequently, each store i sees a Poisson stream of customers having rate $\sum_j \alpha_{ij}\lambda_j$. We assume without loss of generality that $\sum_i \alpha_{ij} = 1$. We assume that customers in each segment have exponential reservation price distributions with mean r and moreover that upon a purchase the store has a mechanism in place to identify what segment the purchasing customer belongs to.

Let $p_t \in \mathbb{R}^N, t \in [0, \infty)$ represent the process of prices charged at the stores over time. Let $n_{t,i}^j$ represent the total number of type j customers served at store i up to time t and let $n_t^j = \sum_i n_{t,i}^j$. The parameter vectors a and b are then updated according to:

$$a_{t,j} = a_{0,j} + n_{t,j} \quad \text{and} \quad b_{t,j} = b_{0,j} + \int_{\tau=0}^{t} \sum_i e^{-p_{\tau,i}/r} d\tau$$

Our state at time t is now $z_t = (x_t, a_t, b_t)$. As before, we will consider prices generated by policies π that are measurable, non-negative vector-valued functions of state, so that $p_t = \pi(z_t) \geq 0$. Letting Π denote the set of all such policies, our objective will be to identify a policy $\pi^* \in \Pi$ that maximizes

$$\tilde{J}^{\pi}(z) = E_{z,\pi}\left[\sum_i \int_0^{\tau^i} \hat{\rho}_{t,i} e^{-p_{t,i}/r} dt\right]$$

where $\tau^i = \inf\{t : \sum_j n_{t,i}^j = x_{0,i}\}$ and $\hat{\rho}_i = \sum_j \alpha_{i,j}(a_j/b_j)$. We define the operator

$$\begin{aligned} &(\tilde{H}^{\pi} J)(z) = \\ &\sum_i \left[\hat{\rho}_i e^{-\pi(z)_i/r}\left(\pi(z)_i + \sum_j \frac{\alpha_{i,j}(a_j/b_j)}{\hat{\rho}_i} J(x - e_i, a + e_j, b) - J(z)\right)\right. \\ &\left. + \sum_j e^{-\pi(z)_i/r} \frac{d}{db_j} J(z)\right] - \alpha J(z). \end{aligned}$$

where e_k is the vector that is 1 in the kth coordinate and 0 in other coordinates. One may show that $\tilde{J}^* = \tilde{J}^{\pi^*}$ is the unique solution to

$$\sup_{\pi \in \Pi} (\tilde{H}^\pi J)(z) = 0 \ \ \forall z$$

satisfying $\tilde{J}^*(0, a, b) = 0$, and that the corresponding optimal policy for $x_i > 0$ is given by

$$(\pi^*(z))_i = \\ r + \tilde{J}^*(z) - \sum_j \frac{\alpha_{i,j}(a_j/b_j)}{\hat{\rho}_i} \tilde{J}^*(x - e_i, a + e_j, b) - \frac{1}{\hat{\rho}_i} \sum_j \frac{d}{db_j} \tilde{J}^*(z). \tag{3.1}$$

Now, assuming that the λ_j's are known perfectly a-priori, it is easy to see that the control problem decomposes across stores. In particular, the optimal strategy simply involves store i using as it's pricing policy

$$p_{t,i} = \pi^*_{\rho_i}(x_{t,i})$$

where $\rho_i = \sum_j \alpha_{i,j} \lambda_j$. Consequently, a certainty equivalent policy would use the pricing policy

$$(\pi_{CE}(z))_i = \pi^*_{\hat{\rho}_i}(x_i)$$

We can also consider as an approximation to $\tilde{J}^*$, the following upper bound (which is in the spirit of the upper bound we derived in Section 3):

$$\overline{J}(z) = E\left[\sum_i \tilde{J}^*_{\rho_i}(x_i)\right].$$

The analogous greedy pricing policy $\pi_{\rm gp}$ is then given by (3.1) upon substituting $\overline{J}(\cdot)$ for $\tilde{J}^*(\cdot)$ in that expression.

Motivated by the decay balancing policy derived for the single store case we consider using the following pricing policy at each store:

$$(\pi_{\rm db}(z))_i = r \log\left(\frac{r\hat{\rho}_i}{\alpha E[J^*_{\rho_i}(x_i)]}\right).$$

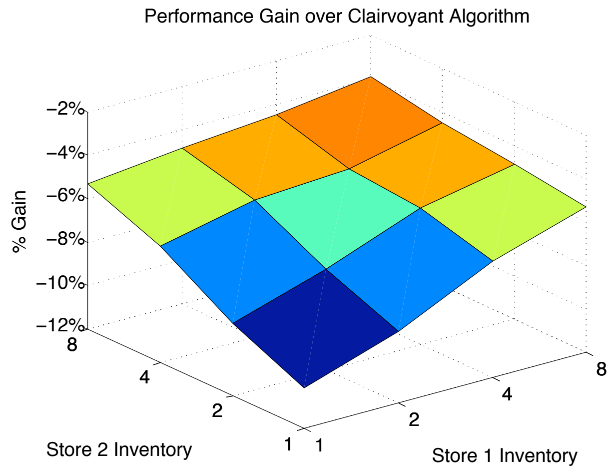

Figure 3.1: Performance relative to a clairvoyant algorithm

Such a heuristic involves joint learning across stores and continues to account for the level of uncertainty in the pricing process. Further, the structural properties discussed in the single store case are retained. The joint learning under this heuristic does not however account for inventory levels across stores.

We now present computational results for the three heuristics. Our experiments will use the following model parameters. We take $N = 2, M = 2$ and assume $\alpha_{i,j} = 1/2$ for all i, j, and further that we begin with prior parameters $a_1 = a_2 = 0.04$, and $b_1 = b_2 = 0.001$ (which corresponds to a mean of 40 and a coefficient of variation of 5). As usual, $\alpha = e^{-1}, r = 1$. Our first set of results (Figure 3.1) compares the decay balancing heuristic's performance against that of a clairvoyant algorithm which as in Section 2.8 has perfect a-priori knowledge of λ. As in the $N = 1, M = 1$ case, our performance is quite close to that of the clairvoyant algorithm. Figure 3.2 compares decay balancing performance to the certainty equivalent heuristic and the greedy heuristic. Figure 3.2 is indicative of performance that is qualitatively similar to that observed for the $N = 1, M = 1$ case; there is a significant gain over certainty equivalence at lower inventory levels, but this gain shrinks as inventory level increases. The performance of the greedy heuristic is particularly dismal, one explanation for which is that $\sum_j \frac{d}{db_j}\overline{J}(z)$ is a potentially poor approximation to $\sum_j \frac{d}{db_j}\tilde{J}^*(z)$.

3.2 Product Versioning

Firms often create and sell several "versions" of a given product as a means of third degree price discrimination. For example an airline may sell several versions of a particular itinerary for air travel. While the various versions might require identical airline network resources, they are distinguished from each other by their respective restrictions. For example, while one version might allow for ticket cancellations at any point prior to departure with no penalties, another version might charge penalties for such a cancellation, and yet another version might forbid cancellation altogether. We would like to consider an extension to the model in Chapter 2 that allows the vendor to sell several versions of his product simultaneously. The question we would like to ask is how such a vendor should adjust prices *of each version* over time so as to maximize total discounted revenues when faced with uncertainty in net customer arrival rate.

Discrete choice models offer a typically tractable means of modeling demand when a customer is faced with a choice from among several versions of a product, and in particular are able to capture to a first order the substitution effects that having several versions of a given product give rise to. More formally, assuming a set of N products indexed by $i \in \{1, \ldots, N\}$, where product i is associated with differentiating features and a price p_i, we will treat a discrete choice model as a mapping $\mathcal{P}$ from prices $p \in \mathbb{R}^N_+$ to purchase probability $\Pr(p) \in [0,1]^N$ satisfying $\sum_i \Pr(p)_i < 1$. $\mathcal{P}(p)_i$ is thus the probability that an arriving customer would choose to purchase product version i, when the prices for the various versions on offer are set according to p. An example of such a model is the multinomial Logit choice model which is specified by a set of $2N$ parameters, $\alpha \in \mathbb{R}^N_+$ and $\beta \in \mathbb{R}^N$, and is given by:

$$\mathcal{P}_i(p) = \frac{e^{-\alpha_i p_i + \beta_j}}{1 + \sum_j e^{-\alpha_j p_j + \beta_j}}$$

Models of customer choice such as the multinomial Logit above may seem somewhat arbitrary at first glance; nonetheless many prevalent customer models have interesting economic justifications. For example, the multinomial Logit models customer choice when customers have independent random reservation prices (of the form $u_j + \eta_j$ where

η_j is a certain extreme valued random variable) for the various product versions, and make choices that maximize their surplus. Ben-Akiva and Lerman (1985) provides an interesting discussion of various choice models that may be used in modeling situations such as the product versioning example under consideration.

3.2.1 Problem Formulation

We consider the following extension to the model with uncertain arrival rates of Chapter 2. The vendor is endowed with a finite initial inventory $x_0 \in \mathbb{Z}_+$ of a product and chooses to sell N versions of the product. Customers arrive at a rate λ where λ is a Gamma distributed random variable with parameters a and b. Upon arrival the customer faces a choice among the N product versions priced according to $p \in \mathbb{R}_+^N$. The customer chooses at most one product version for purchase; version i is chosen with probability $\mathcal{P}(p)_i$, and we denote by $I(p)$ the random index of the chosen version. The vendor sets a vector of prices $p_t \in \mathbb{R}_+^N$ over time $t \in [0, \infty)$. Letting n_t denote the total number of sales up to time t, the vendor updates the parameters of his prior on arrival rate according to

$$a_t = a_0 + n_t \text{ and } b_t = b_0 + \int_0^t \theta_\tau d\tau$$

where $\theta_\tau = \sum_i \mathcal{P}(p_\tau)_i$.

From a control standpoint, our state is now $z_t = (x_t, a_t, b_t)$, and as in Chapter 2 we will consider prices generated by policies π that are measurable, non-negative, vector valued functions of state, so that $p_t = \pi(z_t) \geq 0$. Letting Π denote the set of all such policies, our objective will be to identify a policy $\pi^* \in \Pi$ that maximizes

$$\begin{aligned} J^\pi(z) &= E_{z,\pi}\left[\int_0^{\tau_0} p_{t,I(p_t)} dn_t\right] \\ &= E_{z,\pi}\left[\int_0^{\tau_0} \sum_i \mathcal{P}(p_t)_i p_{t,i} \lambda dt\right] \end{aligned}$$

where we use the fact that the sales process for product version i is Poisson of instantaneous rate $\lambda \mathcal{P}(p_t)_i$.

3.2.2 Optimal pricing

For each policy $\pi \in \Pi$, we define an operator

$$(H^{\pi}J)(z) = \sum_i \mu(z)\mathcal{P}(\pi(z))_i \left(\pi(z)_i + J(z') - J(z) + \frac{1}{\mu(z)}\frac{d}{db}J(z)\right) - \alpha J(z)$$

where $z' = (x-1, a+1, b)$. Assuming that there exists a solution $J^* \in \mathcal{J}$ to

$$\sup_{\pi \in \Pi}(H^{\pi}J)(z) = 0 \; \forall z$$

it is straight forward to show that this solution is unique and that $J^* = J^{\pi^*}$. Further, the optimal policy requires

$$\pi^*(z) \in \mathrm{argmax}_{\pi(z)} \sum_i \mu(z)\mathcal{P}(\pi(z))_i \left(\pi(z)_i + J(z') - J(z) + \frac{1}{\mu(z)}\frac{d}{db}J(z)\right)$$

We note that in the absence of uncertainty in λ, the corresponding HJB equation is in fact reduced to a simple recursion via which computing J^*_{λ} is a computationally tractable task.

While the preceding formulation is natural, it has the disadvantage of requiring the solution to a control problem with an N-dimensional action space. An alternative, equivalent view of the optimal control problem at hand with a one-dimensional action space is the following: At each point in time, the vendor picks $\theta_t \in [0, \beta]$, the fraction of arriving customers that will make a purchase of some version. Here $\beta = \max_{p \in \mathbb{R}^N_+} \sum_i \mathcal{P}(p)_i$, the maximal arrival rate that the seller can induce. One may then assume without loss that he should set prices $p_t = p^*(\theta_t) \in \mathrm{argmax}_{p: \sum_i \mathcal{P}(p)_i = \theta_t} \sum_i \mathcal{P}(p)_i p_i$. We may consequently view a policy $\tilde{\pi}$ as a measurable mapping from state to the interval $[0, \beta]$ and we denote by $\tilde{\Pi}$ the set of all such policies.

Define the expected revenue from a sale upon taking action θ as $R(\theta) = \sum_i \frac{p^*(\theta)_i \mathcal{P}(p^*(\theta))_i}{\sum_i \mathcal{P}(p^*(\theta))_i}$. We make the following assumptions on the functions $R(\cdot)$ and $\cdot R(\cdot)$:

Assumption 3. *$R(\cdot)$ is a bounded, differentiable function and $\cdot R(\cdot)$ is a strictly concave* *function on the interval* $[0, \beta]$.

We next define for each policy $\tilde{\pi}$, the operator:

$$(\tilde{H}^{\tilde{\pi}} J)(z) = \mu(z)\tilde{\pi}(z) \left(R(\tilde{\pi}(z)) + J(z') - J(z) + \frac{1}{\mu(z)} \frac{d}{db} J(z) \right) - \alpha J(z).$$

Now, assuming that there exists a solution $J^* \in \mathcal{J}$ to

$$\sup_{\tilde{\pi} \in \tilde{\Pi}} (\tilde{H}^{\tilde{\pi}} J)(z) = 0 \; \forall z$$

it is straight forward to show that this solution is unique and that $J^* = J^{\pi^*}$.

3.2.3 Decay Balancing

A solution J^* to the HJB equation (for our second formulation) must satisfy:

$$\sup_{\theta \in [0,\beta]} \mu(z)\theta \left(R(\theta) + J^*(z') - J^*(z) + \frac{1}{\mu(z)} \frac{d}{db} J^*(z) \right) = \alpha J^*(z) \tag{3.2}$$

Now, if θ^* attains the maximum in the equation above, and assuming that this maximum is always attained in $(0, \beta)$, the first order optimality conditions are by Assumption 3 necessary and sufficient, and imply that

$$-(\theta^*)^2 R'(\theta^*) = \frac{\alpha J^*(z)}{\mu(z)}$$

This is a decay balance equation for the dynamic pricing problem with product versioning. Now by Assumption 3, the optimal control problem at hand has a unique optimizing policy in $\tilde{\Pi}$ (that is, the maximizing set for the left hand side of (3.2) is a singleton), so that the decay balance equation at state z has a unique solution $\tilde{\theta}(z)$ on $(0, \beta)$ with $\tilde{\theta}(z) = \theta^*(z)$. $p^*(\tilde{\theta}(z))$ is then the optimal price vector in state z.

Of course, since we do not know J^* we could consider using the same approximation to J^* as in the previous section, $\tilde{J}(z) = E[J^*_\lambda(z)]$. The decay balance heuristic price in state z is then $p^*(\theta_{\mathrm{db}}(z))$ where $\theta_{\mathrm{db}}(z)$ solves

$$-(\theta_{\mathrm{db}}(z))^2 R'(\theta_{\mathrm{db}}(z)) = \frac{\alpha \tilde{J}(z)}{\mu(z)}$$

If $\tilde{J}(z)$ is a good approximation to $J^*(z)$, $p^*(\theta_{\mathrm{db}}(z))$ is likely to be a good approximation to $p^*(\theta^*(z))$ and one would expect similar levels of performance as in the one product case studied in Chapter 2.

3.3 Discussion

Decay balancing type heuristics should typically be useful in deriving pricing policies for problems with uncertainty in a single parameter and a one dimensional control. This chapter attempted to generalize the heuristic to other situations. The product versioning problem nominally requires an N-dimensional control, but our development took advantage of the fact that the optimal control problem may be reduced to one with a one dimensional control. For the problem with multiple stores and consumer segments, our use of the decay balancing heuristic was ad-hoc and motivated by the single store solution of Chapter 2, but nonetheless yielded promising computational results. There are several other formulations for which we believe decay balancing might well yield useful pricing heuristics including problems with uncertainty in demand elasticity (as opposed to market response), or the finite time horizon version of the problem in Chapter 2.

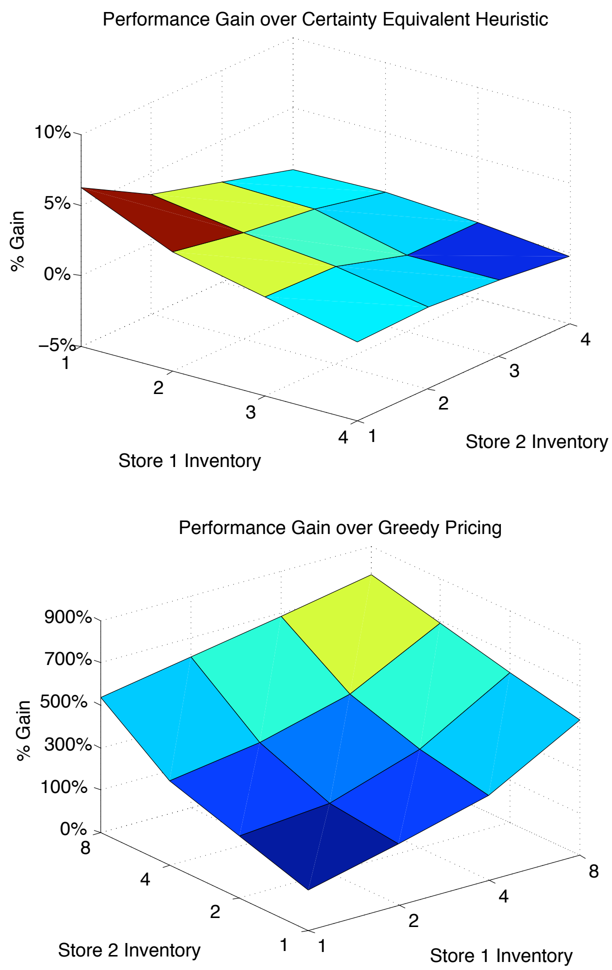

Figure 3.2: Performance relative to the Certainty Equivalent (left) heuristic and Greedy Pricing heuristic (right)

Chapter 4

Network-RM with Forecast Models

This chapter focuses on a problem of airline Network-RM. In contrast to the dynamic pricing problems studied in Chapters 2 and 3, demand in the airline industry is far more predictable and fairly well modeled. Most airlines have access to complex forecast models estimated on the basis of large quantities of relevant historical data. The "Estimate, Then Optimize" heuristic in this context entails the use of optimization algorithms that in the interest of tractability make simplifying assumptions of demand. In doing so, the revenue manager is typically unable to harness all the predictive capabilities of the forecast models available to him.

We develop in this chapter an approximation algorithm for a dynamic capacity allocation problem with Markov modulated customer arrival rates (the typical simplifying assumption made in solving such optimization problems is that demand is deterministic and equal to expected forecasted demand). For each time period and each state of the modulating process, the algorithm approximates the dynamic programming value function using a concave function that is separable across resource inventory levels. We establish via computational experiments that our algorithm increases expected revenue, in some cases by close to 8%, relative to a deterministic linear program that is widely used for bid-price control.

4.1 Introduction

Network revenue management refers to the activity of a vendor who is endowed with limited quantities of multiple resources and sells products, each composed of a bundle of resources, controlling their availability and/or prices over time with an aim to maximize revenue. The airline industry is perhaps the most notable source for such problems. An airline typically operates flights on each leg of a network of cities and offers for sale "fare products" composed of seats on one or more of these legs. Each fare product is associated, among other things, with some fixed price which the airline receives upon its sale. Since demand for fare products is stochastic and capacity on each leg limited, the airline's problem becomes one of deciding which of its fare products to offer for sale at each point in time over a finite sales period so as to maximize expected revenues. This chapter presents a new algorithm for this widely studied problem.

For most models of interest, the dynamic capacity allocation problem we have described can be cast as a dynamic program, albeit one with a computationally intractable state-space even for networks of moderate size. As such, revenue management techniques have typically resorted to heuristic control strategies. Early heuristics for the problem were based primarily on the solutions to a set of single resource problems solved for each leg. Today's state of the art techniques involve "bid-price" control. A generic bid-price control scheme might work as follows: At each point in time the scheme generates a bid-price for a seat or unit of capacity on each leg of the network. A request for a particular fare product at that point in time is then accepted if and only if the revenue garnered from the sale is no smaller than the sum of the bid prices of the resources or seats that constitute that fare product. There is a vast array of available algorithms that may be used in the generation of bid-prices. There are two important dimensions along which such an algorithm must be evaluated. One, of course, is revenues generated from the strategy. Since bid-prices must be generated in real time, a second important dimension is the efficiency of the procedure used to generate them. A simple approach to this problem which has found wide-spread acceptance involves the solution of a single linear program referred to as the deterministic LP (DLP). This approach and associated bid-price techniques have found widespread use in modern revenue management systems and are believed to have generated incremental revenues on the order of 1-2% greater than previously used "fare-class" level heuristics (see P.P.Belobaba and Lee (2000), P.P.Belobaba (2001)).

The algorithm we present applies to models with Markov-modulated customer arrival rates. This represents a substantial generalization of the deterministic arrival rate arrival process models generally considered in the literature and accommodates a broad class of demand forecasting models. We demonstrate via a sequence of computational examples that our algorithm consistently produces higher revenues than a strategy using bid-prices computed via re-solution of the DLP *at each time step*. While the performance gain relative to the DLP is modest ($\sim 1\%$) for a model with time homogeneous arrivals, this gain increases significantly when arrival rates vary stochastically. Even for a simple arrival process in which the modulating process has three states, we report relative performance gains of up to about 8% over a DLP approach suitably modified to account for the stochasticity in arrival rates.

Our algorithm is based on a linear programming approach to approximate dynamic programming (de Farias and Van Roy (2003), de Farias and Van Roy (2004)). A linear program is solved to produce for each modulating process state and each time an approximation to the optimal value function that is separable across resource inventory levels. A heuristic is then given by the greedy policy with respect to this approximate value function. This policy can be interpreted in terms of bid-price control for which bid prices are generated at each point in time via a table look-up, which takes far less time than solving the DLP.

The ALP has as many constraints as the size of the state space and practical solution requires a constraint sampling procedure. We exploit the structural properties afforded by our specific approximation architecture to derive a significantly simpler alternative (the rALP) for which the number of constraints grows linearly with maximal capacity on each network leg. The rALP generates a feasible solution to the ALP. We show that this solution is optimal for affine approximations. While we aren't able to prove that this solution is optimal for concave approximations, the rALP generates optimal ALP solutions in all of our computational experiments with that architecture as well. The rALP thus significantly enhances the scalability of our approach.

The literature on both general dynamic capacity allocation heuristics, as well as bid-price controls is vast and predominantly computational; Talluri and van Ryzin (2004) provides an excellent review. Closest to this work is the paper by Adelman (2005), which also proposes an approximate DP approach to computing bid prices via an *affine* approximation to the value function. The rALP we propose allows an exponential reduction in the number of constraints for the ALP with affine approximation. However, in spite of affine approximation being a computationally attractive approximation architecture, our computational experiments suggest that affine approximations are not competitive with an approach that uses bid-prices computed via re-solution of the DLP at each time step.

Our approach might be viewed as a means of generating bid-prices. There have been a number of algorithms and heuristics proposed for this purpose. One class of schemes is based on mathematical programming formulations of essentially static versions of the problem that make the simplifying assumption that demand is deterministic and equal to its mean. The DLP approach is representative of this class and apparently the method of choice in practical applications (Talluri and van Ryzin (2004)). We compare the performance of our approach to such a scheme. Highly realistic simulations in P.P.Belobaba (2001) suggest that this class of approaches generates incremental revenues of approximately 1-2% over earlier leg-based RM techniques. There are alternatives to the use of bid price controls, the most prominent among them being "virtual nesting" schemes such as the displacement adjusted (DAVN) scheme (see Talluri and van Ryzin (2004)). We do not consider our performance relative to such schemes; a subjective view (E. A. Boyd (2005)) is that these schemes are consistently outperformed by bid-price based schemes in practice.

An important thrust of our work is the incorporation of Markov-modulated customer arrival processes. There is an emerging literature on optimization techniques for models that incorporate demand processes where arrival rates are correlated in time. A recent example is the paper by de Miguel and Mishra (2006), that evaluates various multi-stage stochastic programming techniques for a linear (with additive noise) model of demand evolution. These approaches rely on building "scenario-trees" based on simulations of demand trajectories. While they can be applied to Markov-modulated arrival processes, scenario trees and their associated computational requirements typically grow exponentially in the horizon.

The remainder of this chapter is organized as follows: In section 2, we formally specify a model for the dynamic capacity allocation problem. In section 3 we review the benchmark DLP heuristic. Section 4 presents an ADP approach to the dynamic capacity allocation problem and specifies our approximation architecture. That section also discusses some simple structural properties possessed by our approximation to the value function. Section 5 presents a series of computational examples comparing the performance of our algorithm with the DLP approach as also an approach based on an affine approximation to the value function. Section 6 studies a simple scalable alternative to the ALP, the rALP, and discusses computational experience with that program. Section 7 concludes.

4.2 Model

We consider an airline operating L flight legs. The airline may offer up to F fare products for sale at each point in time. Each fare product f is associated with a price p_f and requires seats on one or more legs. A matrix $A \in \mathbb{Z}_+^{L \times F}$ encodes the capacity on each leg consumed by each fare product: $A_{l,f} = k$ if and only if fare product f requires k seats on leg l. For concreteness we will restrict attention to the situation wherein a given fare product can consume at most 1 seat on any given leg although our discussion and algorithms carry over without any change to the more general case. Initial capacity on each leg is given by a vector $x_0 \in \mathbb{Z}_+^L$. Time is discrete. We assume an N period horizon with at most one customer arrival in a single period. A customer for fare product f arrives in the nth period with probability $\lambda_f(m_n)$. Here $m_n \in \mathcal{M}$ (a finite set) and represents the current demand "mode". m_n evolves according to a discrete time Markov process on $\mathcal{M}$ with transition kernel P_n. We note that the discrete time arrival process model we have described may be viewed as a uniformization of an appropriately defined continuous time arrival process. At the start of the nth period the airline must decide which subset of fare products from the set $\{f : A_f \preceq x_n\}$ it will offer for sale; an arriving customer for fare product f is assigned that fare product should it be available, the airline receives p_f, and

$x_{n+1} = x_n - A_f$.

We define the *state-space* $\mathcal{S} = \{x : x \in \mathbb{Z}_+^L, x \preceq x_0\} \times \{0, 1, 2, \ldots, N\} \times \mathcal{M}$. Encoding the fare products offered for sale at time n by a vector in $\{0,1\}^F \equiv \mathcal{A}$, a control policy is a mapping $\pi : \mathcal{S} \to \mathcal{A}$ satisfying $A\pi(s) \leq x(s)$ for all $s \in \mathcal{S}$. Let Π represent the set of all such policies. Let $R(s,a)$ be a random variable representing revenue generated by the airline in state $s \in \mathcal{S}$ when fare products $a \in \mathcal{A}$ are offered for sale, and define for $s \in \mathcal{S}$,

$$J^\pi(s) = E_\pi\left[\sum_{n'=n(s)}^{N-1} R(s_{n'}, \pi(s_{n'}))\big| s_{n(s)} = s\right].$$

We let $J^*(s) = \max_{\pi\in\Pi} J^\pi(s)$, denote the expected revenue under the optimal policy π^* upon starting in state s.

J^* and π^* can, in principle, be computed via Dynamic Programming. In particular, define the dynamic programming operator T for $s \in \{s' : n(s') < N-1\}$ according to

$$\begin{aligned}(TJ)(s) = &\sum_{f:A_f \leq x(s)} \lambda_f(m(s)) \max\left[p_f + E\left[J\left(S'_f\right)\right], E\left[J\left(S'\right)\right]\right] \\ &+ \Big(1 - \sum_{f:A_f \leq x(s)} \lambda_f(m(s))\Big) E\left[J(S')\right].\end{aligned} \tag{4.1}$$

where $S'_f = (x(s) - A_f, n(s)+1, m_{n(s)+1})$ and $S' = (x(s), n(s)+1, m_{n(s)+1})$. For $s \in \{s' : n(s') = N-1\}$ we define $(TJ)(s) = \sum_{f:A_f\leq x(s)} \lambda_f(m(s))p_f$. We define $(TJ)(s) = 0$ for all $s \in \{s' : n(s') = N\}$. J^* may then be identified as the unique solution to the fixed point equation $TJ = J$. π^* is then the policy that achieves the maximum in (4.1); in particular, $\pi^*(s)_f = 0$ iff $p_f + E\big[J(S'_f)\big] < E\left[J\left(S'\right)\right]$ and $n(s) < N-1$.

We will focus on three special cases of the above model, with N assumed even for notational convenience:

- (M1) Time homogeneous arrivals: Here we have $|\mathcal{M}| = 1$. That is the arrival rate of customers for the various fare products is constant over time and the arrival process is un-correlated in time.

- (M2) Multiple demand modes, deterministic transition time: Here we consider a model with $\mathcal{M} = \{\text{med,hi,lo}\}$. We have $m_n = \text{med}$ for $n \leq N/2$. With probability $\tilde{p}$, $m_n = \text{lo}$ for all $n > N/2$ and with probability $1 - \tilde{p}$ and $m_n = \text{hi}$ for all $n > N/2$. This is representative of a situation where there is likely to be a change in arrival rates at some known point during the sales season. The revenue manager has a probabilistic model of what the new arrival rates are likely to be.

- (M3) Multiple demand modes, random transition time: Here we consider a model with $\mathcal{M} = \{\text{med, hi,lo}\}$, with the transition kernel P_t defined according to

$$P_n(m_{n+1} = y | m_n = x) = \begin{bmatrix} 1-q & q\tilde{p} & q(1-\tilde{p}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{xy}.$$

where $q, \tilde{p} \in (0, 1)$. This arrival model is similar to the second with the exception that instead of a change in demand modes occurring at precisely $n = N/2$, there is now uncertainty in when this transition will occur. In particular, the transition time is now a geometric random variable with expectation $1/q$.

The above models were chosen since they are simple and yet serve to illustrate the relative merits of our approach for Markov-modulated demand processes.

4.3 Benchmark Heuristic: The Deterministic LP (DLP)

The Dynamic Programming problem we have formulated is computationally intractable and so one must resort to various sub-optimal control strategies. We review the DLP-heuristic for generating bid prices. This heuristic makes the simplifying assumption that demand is deterministic and equal to its expectation. In doing so, the resulting control problem reduces to the solution of a simple LP (the DLP) and the optimal control policy is static. In particular, if demand for fare product f over a $N-n$ period sales season, $D_{n,f}$, were deterministic and equal to expected demand, $\mathbb{E}[D_{n,f} | m_n]$, the maximal revenue that one may generate with an initial capacity $x(s)$ is given by the optimal solution to the DLP:

$$\begin{aligned} DLP(s): \quad & \max && p'z \\ & \text{s. t.} && Az \le x(s) \\ & && 0 \le z \le \mathbb{E}[D_{n(s)} | m_{n(s)} = m(s)] \end{aligned}$$

Denote by $r^*(s)$ a vector of optimal shadow prices corresponding to the constraint $Az \le x(s)$ in $DLP(s)$. The bid price control policy based on the DLP solution is then given by:

$$\pi^{\mathrm{DLP}}(s)_f = \begin{cases} 1 & \text{if } A_f' r^*(s) \le p_f \text{ and } A_f \le x(s) \\ 0 & \text{otherwise} \end{cases}$$

The above description of the DLP heuristic assumes that the shadow prices r^* are recomputed at each time step. While this many not always be the case, a general computational observation according to Talluri and van Ryzin (2004) is that frequent recomputation of r^* improves performance. This is consistent with our computational experience.

In the case of model M2, one might correctly point out that a simple modification of the DLP is likely to have superior performance. In particular, one may consider retaining the probabilistic structure of the demand mode transition model and solving a multi-stage stochastic program with recourse variables for capacity allocation in the event of a transition to the hi and lo demand modes respectively. We do not consider such a stochastic programming approach as it is intractable except for very simple models (such as M2); for a general Markov-modulated demand model with at least two demand modes, the number of recourse variables grows exponentially with horizon length.

4.4 Bid Price Heuristics via Approximate DP

Given a component-wise positive vector c, the optimal value function J^* may be identified as the optimal solution to the following LP:

$$\begin{array}{ll} \min & c'J \\ \text{s.t.} & (TJ)(s) \leq J(s) \quad \forall s \in \mathcal{S} \end{array}$$

The linear programming approach to approximate DP entails adding to the above LP, the further constraint that the value function J lie in the linear span of some set of basis functions $\phi_i : \mathcal{S} \to \mathbb{R},\ i = 1, 2, \ldots, k$. Encoding these functions as a matrix $\Phi \in \mathbb{R}^{|\mathcal{S}| \times k}$, the approximate LP (ALP) computes a vector of weights $r \in \mathbb{R}^k$ that optimally solve:

$$\begin{array}{ll} \min & c'\Phi r \\ \text{s.t.} & (T\Phi r)(s) \leq (\Phi r)(s) \quad \forall s \in \mathcal{S} \end{array}$$

Given a solution r^* to the ALP (assuming it is feasible), one then uses a policy that is greedy with respect to Φr^*. Of course, the success of this approach depends crucially upon the choice of the set of basis functions Φ. In the next two subsections we examine affine and concave approximation architectures. The affine approximation architecture for the network RM problem was proposed by Adelman (2005) in the context of the M1 model. The concave architecture is the focus of this work. In the sequel we assume that $c_{s_0} = 1$ and that all other components of c are 0.

4.4.1 Separable Affine Approximation

Adelman (2005) considers the use of affine basis functions in the M1 model. In particular, Adelman (2005) explores the use of the following set of $(L+1)N$ basis functions defined according to

$$\phi_{l,n}(x, n') = \begin{cases} x_l & \text{if } l \leq L \text{ and } n = n' \\ 1 & \text{if } l = L + 1 \text{ and } n = n' \\ 0 & \text{otherwise} \end{cases}$$

The ALP here consequently has $\Theta(LN)$ variables but $\Theta(\bar{x}^L NF)$ constraints. Adelman (2005) proposes the use of a column generation procedure to solve the ALP. We show in Section 6 that the ALP can be reduced to an LP with $\Theta(LN)$ variables and $\Theta(\bar{x}2^L LNF)$ constraints making practical solution of the ALP to optimality possible for relatively large networks (including, for instance, the largest examples in Adelman (2005)).

In spite of being a computationally attractive approximation architecture, affine approximations have an obvious weakness: the greedy policy with respect to an affine approximation to the value function is insensitive to intermediate capacity levels so that the set of fare products offered for sale at any intermediate point in time depends only upon the time left until the sales season ends. In particular the greedy policy with respect to an affine approximation, π^{aff}, will satisfy $\pi^{\text{aff}}(x, n) = \pi^{\text{aff}}(\tilde{x}, n)$ provided x and $\tilde{x}$ are positive in identical components. We observe in computational experiments that a policy that is greedy with respect to an affine approximation to the value function is in fact not competitive with a policy based on re-computation of bid-prices at each time step via the DLP. While one possible approach to consider is frequent re-solution of the ALP with affine approximation, this is not a feasible option given that bid-prices must often be generated in real time. It is simple to show (using, for example, the monotonicity of the T operator) for any vector $e \in \{0, 1\}^L$ that is positive in a single component, that $J^*(x + e, n) - J^*(x, n)$ is non-increasing in x. Affine approximations are incapable of capturing this concavity of J^* in inventory level. This motivates us to consider a separable concave approximation architecture which is the focus of this chapter.

4.4.2 Separable Concave Approximation

Consider the following set of basis functions, $\phi_{l,n,i,m}$, defined for integers $l \in [1, L]; n \in [0, N], i \in [0, (x_0)_l]$, and $m \in \mathcal{M}$ according to:

$$\phi_{l,n,i,m}(x', n', m') = \begin{cases} 1 & \text{if } x'_l \geq i \text{ , } n = n' \text{ and } m = m' \\ 0 & \text{otherwise} \end{cases}$$

The ALP in this case will have $\Theta(\bar{x}LN|\mathcal{M}|)$ variables and $\Theta(\bar{x}^L NF|\mathcal{M}|)$ constraints. Note that optimal solution is intractable since $|\mathcal{S}|$ is exponentially large. One remedy is the constraint sampling procedure in de Farias and Van Roy (2004) which suggests sampling constraints from $\mathcal{S}$ according to the state-distribution induced by an optimal policy. Assuming a sales season of N periods and an initial inventory of x_0, we propose using the following procedure with parameter K:

1. Simulate a bid price control policy starting at state $s_0 = (x_0, 0, m_0)$, using bid prices generated by re-solving the DLP at each time step. Let $\mathcal{X}$ be the set of states visited over the course of several simulations. We generate a set with $|\mathcal{X}| = K$

2. Solve the following Relaxed LP (RLP):

$$\begin{array}{lll} \min & (\Phi r)(s_0) & \\ \mathrm{s.\,t.} & (T\Phi r)(s) \leq (\Phi r)(s) & \text{for } s \in \mathcal{X} \\ & r_{l,n,i,m} \geq r_{l,n,i+1,m} & \forall i > 0, l, n, m \end{array}$$

3. Given a solution r^* to the RLP, use the following control policy over the actual sales season:

$$\pi^{\text{con}}(s)_f = \begin{cases} 1 & \text{if } \sum_{l:A_{l,f}=1} r^*_{l,n(s),x(s)_l,m(s)} \leq p_f \text{ and } A_f \leq x(s) \\ 0 & \text{otherwise} \end{cases}$$

Several comments on the above procedure are in order. Step 1 in the procedure entails choosing a suitable number of samples K; de Farias and Van Roy (2004) provides some guidance on this choice. Our choice of K was heuristic and is described in the next section. Step 2 of the procedure entails solving the RLP whose constraints are samples of the original ALP. We will shortly mention several simple structural properties that an optimal solution to the ALP must posses. Adding these constraints to the RLP strengthens the quality of our solution. Also, note that the inequality constraints on the weights enforce concavity of the approximation. Finally note that the greedy policy with respect to the our approximation to J^* takes the form of a bid price policy as in the case of affine approximation. However, unlike affine approximation the resulting policy decisions depend on available capacity as well as time.

4.4.3 ALP Solution Properties

The optimal solution to the DLP provides an upper bound to the true value function J^*, i.e. $DLP(s) \geq J^*(s)$. There are several proofs of this fact for the time homogeneous model M1. For example, see Gallego and van Ryzin (1997) or Adelman (2005). The DLP continues to be an upper bound to the true value function for the more general model we study here (via a simple concavity argument and the use of Jensen's inequality). We can show that the ALP with separable concave approximation provides a tighter upper bound than does the DLP for model M2, and generalizations to M2 which allow for more than a single branching time. The same result for time homogeneous arrival rates (i.e. for model M1) follows as a corollary. We are at present unable to establish such a result for the general model.

Lemma 11. *For model M2 with initial state s, $J^*(s) \leq ALP(s) \leq DLP(s)$*

The proof of the lemma can be found in the appendix. The above result is not entirely conclusive. In particular, while it is indeed desirable to have a good approximation to the true value function, a tighter approximation does not guarantee an improved policy. Nonetheless, stronger approximations to the true value function imply stronger bounds on policy performance. Finally, solutions to the ALP must satisfy simple structural properties. For example, in the case of model M1 it is clear that we must have $\sum_{i=0}^{x} r_{l,n,i} \geq 0$ for all $x, l, n < N$ and further $\sum_{i=0}^{x} r_{l,n,i} \leq \sum_{i=0}^{x} r_{l,n-1,i}$ for all $l, x, 0 < n < N$. We explicitly enforce these constraints in our computational experiments.

4.5 Computational Results

It is difficult to establish theoretical performance guarantees for our algorithm. Indeed, we are unaware of any algorithm for the dynamic capacity allocation problem for which non-asymptotic theoretical performance guarantees are available. As such, we will establish performance merits for our algorithm via a computational study. We will consider two simple test networks each with a single "hub" and either three or four spoke cities. This topology is representative of actual airline network topologies. Each leg in our network represents two separate aircraft (one in each direction) making for a total of $f = 15$ itineraries on the 3 spoke network and $f = 24$ itineraries on the 4 spoke network. Arrival rates for each itinerary, demand mode i.e. (f, m) pair were picked randomly from the

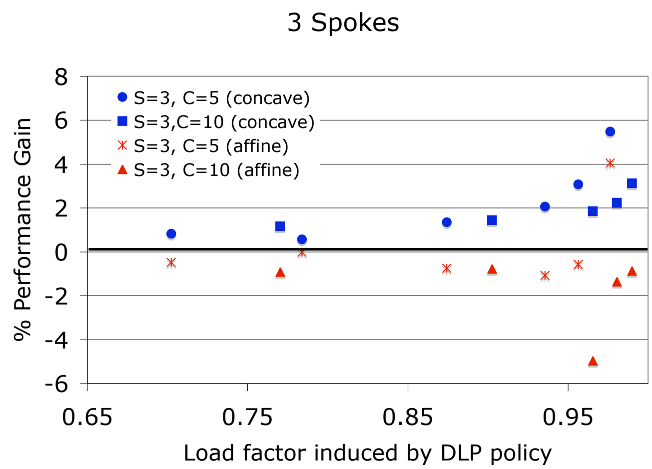

Figure 4.1: Performance relative to the DLP for model M1

unit f-dimensional simplex and suitably normalized. Route prices were generated uniformly in the interval $[50, 150]$ for single leg routes and $[50, 250]$ for two leg routes. We consider a random instantiation of arrival rates and probabilities for each network topology and for each instantiation measure policy performance upon varying initial capacity levels and sales horizon. We compare performance against the DLP with re-solution at each time step. In the case of model M1, we also include policies generated via the separable affine approximation architecture in our experiments. We solve RLPs with $50,000$ sampled states, this number being determined by memory constraints. We now describe in detail our experiments and results for each of the three models.

4.5.1 Time homogeneous arrivals (M1)

We consider three and four spoke models. The arrival probabilities for each fare product were drawn uniformly at random on the unit simplex and normalized so that the probability of no customer arrival in each period was 0.7. For both models, we consider fixed capacities (of 5 and 10 for 3-spoke networks, 10 and 20 for four spoke networks) on each network leg and vary the sales horizon N. For each value of N we record the average load-factor (i.e. the average fraction of seats sold) under the DLP policy; we select values of N so that this induced load factor is > 0.7. We plot in Figures 1 and 2 the performance

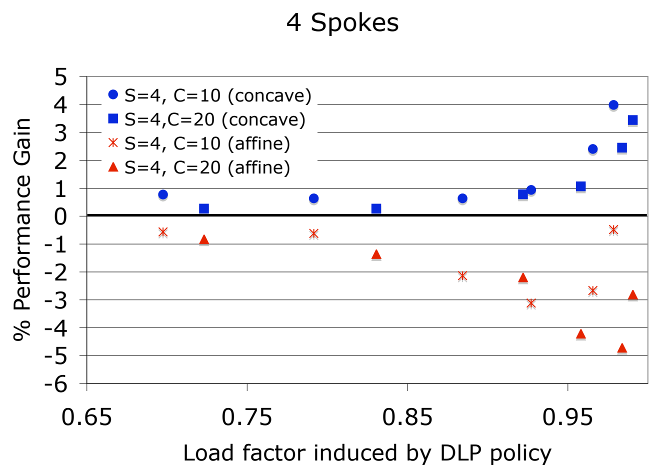

Figure 4.2: Performance relative to the DLP for model M1

of the the ADP based approaches with affine and separable approximations relative to the DLP heuristic for two different initial capacity levels. The x-intercept for a data point in both plots is the average load factor induced by the DLP heuristic for the problem data in question at that point.

The plots suggest a few broad trends. The affine approximation architecture is almost uniformly dominated by the DLP heuristic when the DLP is re-solved at every time step, while the separable architecture uniformly dominates both heuristics in every problem instance. We note that since a bid price computation in the ADP approach is simply a lookup it is far quicker than solving the DLP, so that together these facts support the plausibility of using an ADP approach with separable approximation. Another trend is performance gain. This is actually quite low at low induced load factors ($< 0.5\%$) but can be as high as 5% at high load factors. At moderate load factors (that are at least nominally representative) the performance gain is on the order of 1%. We anticipate the gain to be larger for more complex networks.

It is difficult to expect higher performance gains than we have observed for the M1 demand model. In particular, at low load factors, the problem is trivialized (since it is optimal to accept all requests). Moreover, it is well known (see Gallego and van Ryzin (1997)) that in a certain fluid scaling (which involves scaling both initial capacity x_0 and sales horizon N by some scaling factor $\tilde{N}$), the DLP heuristic is optimal as $\tilde{N}$ gets large.

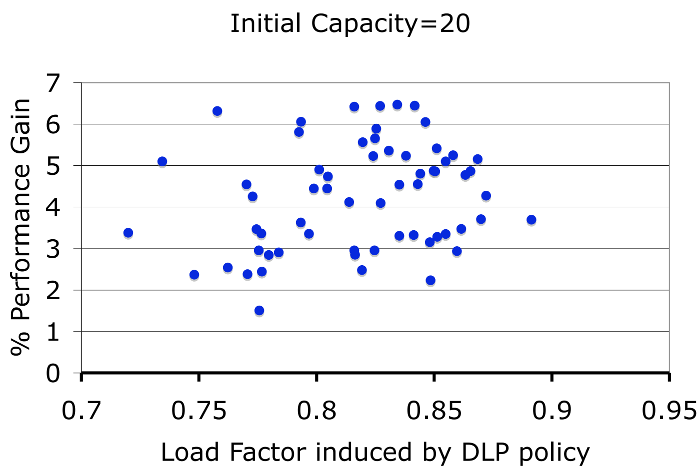

Figure 4.3: Performance relative to the DLP for model M2

The purpose of our experiments with this model is to illustrate the fact that the separable concave approximations we employ are robust in this simple demand setting.

4.5.2 Multiple demand modes (M2, M3)

Model M1 is potentially a poor representation of reality. This leads us to consider incorporating a demand forecasting model such as that in models M2 and M3. In our experiments, the arrival probabilities for each demand mode were drawn uniformly at random on the unit 24-dimensional simplex and normalized so that the probability of no customer arrival in each period was 0.55 for the "med" demand mode, 0.7 for the "lo" mode, and 0.1 in the "hi" mode. The probability of transitioning from the med to lo demand mode, p, was set to 0.5 in both models, and we set $m_0 =$ med.The probability of transitioning out of the med demand state, q, was set to $2/N$ in model M3. The sales horizon N was varied so that the load-factor induced by the DLP policy was approximately between 0.8 and 0.9. We generate a random ensemble of 40 such problems for a network with 4 spokes and consider initial capacity levels of 20 seats and 40 seats. We measure the performance gain of our ADP with separable concave approximation derived bid price control over the DLP. The DLP is resolved at every time step so that it may recompute expected total remaining demand for each fare product conditioned on

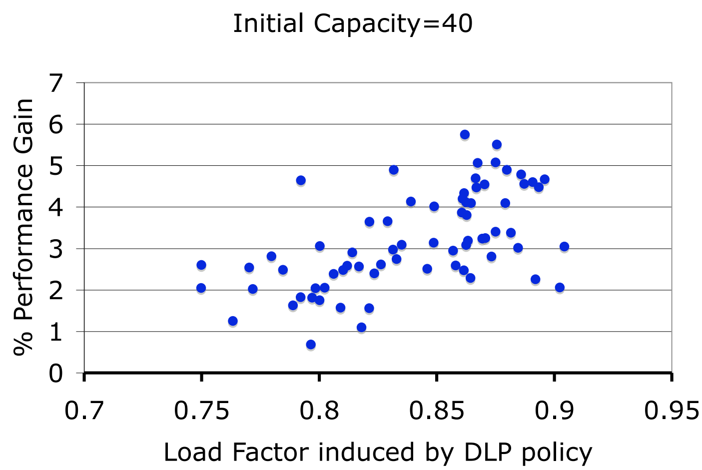

Figure 4.4: Performance relative to the DLP for model M2

the current demand mode.

For model M2, we plot in Figures 3 and 4 the performance of the the ADP based approach with separable concave approximation relative to the DLP heuristic with initial capacity levels of 20 and 40 respectively. We note that the relative performance gain here is significant (up to about 8%) in a realistic operating regime. In the case of model M3, Figure 5 illustrates similar performance trends.

We see that the approximate DP approach with concave approximation offers substantiative gains over the use of the DLP even with very simple stochastic variation in arrival rates. We anticipate that these gains will be further amplified for more complex models of arrival rate variability (for example in models with a larger number of demand modes etc.).

4.6 Towards scalability: A simpler ALP

Assuming maximal capacities of $\bar{x}$ on each of L legs, a time horizon N, and F fare products, the ALP with separable concave approximations has $\Theta(\bar{x}^L NF)$ constraints. In this section we will demonstrate a program - the relaxed ALP (rALP) - with $O(\bar{x}NLF2^L)$ constraints that generates a feasible solution to the ALP. The rALP has the same decision variables as the ALP, and a small number of additional auxiliary variables. The

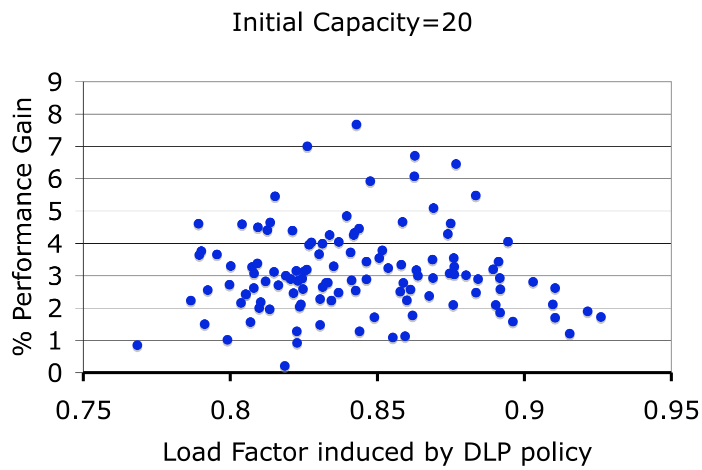

Figure 4.5: Performance relative to the DLP for model M3

rALP is consequently a significantly simpler program than the ALP. In the case of affine approximation, the rALP generates the optimal solution to the ALP. In the case of separable concave approximations, the rALP generates a feasible solution to the ALP whose quality we demonstrate through computational experiments to be excellent. The rALP solution in fact coincides with the ALP solution in all of our experiments. Our presentation will assume that an itinerary can consist of at most 2 flight legs and will be in the context of model M1 for simplicity; extending the program to more general arrival process model is straightforward.

4.6.1 The rALP

In what follows, we understand that for a state $s \in \mathcal{S}$, $s \equiv (x(s), n(s))$. Let us partition $\mathcal{S}$ into sets of the form $\mathcal{S}_y = \{s : s \in \mathcal{S}, x(s)_i = 0 \iff y_i = 0\}$ for all $y \in \{0,1\}^L$. Clearly, $\mathcal{S}$ can be expressed as the disjoint union of all such sets $\mathcal{S}_y$. Also define the subset of fare products $\mathcal{F}_y$ according to $\mathcal{F}_y = \{f : A_f \leq y\}$ and assume $\mathcal{F}_y \neq \emptyset$ for all y. For some $y \in \{0,1\}^L$, consider the set of constraints

$$(T\Phi r)(s) \leq (\Phi r)(s) \text{ for all } s \in \mathcal{S}_y. \tag{4.2}$$

We will approximate the feasible region specified by this set of constraints by the following set of constraints in r and the auxiliary variables m:

$$\begin{array}{ll} LP_{y,n}(r,m) \leq 0 & \forall n < N \\ m^f_{l,n,i} \geq r_{l,n,i} & \forall\, l,n \leq N, i \in \{1,\ldots,\bar{x}\}, f \in \mathcal{F}_y \\ \sum_{l:A_{l,f}=1} m^f_{l,n,\bar{x}} \geq p_f & \forall n \leq N, f \in \mathcal{F}_y \\ m^f_{l,n,i+1} \leq m^f_{l,n,i} & \forall l, n \leq N, i \in \{1,\ldots,\bar{x}-1\}, f \in \mathcal{F}_y \end{array} \tag{4.3}$$

where $LP_{y,n}(r,m)$ refers to a certain linear program with decision variables $x \in \mathbb{R}^{LN(\bar{x}+1)}$. We will now proceed to describe this linear program as also discuss how the constraint $LP_{y,n}(r,m) \leq 0$ may itself be described by a set of linear constraints in r, m and certain additional auxiliary variables.

Let us define:

$$c_{y,n}(r,m)'x = \sum_l \sum_{i=0}^{\bar{x}} (r_{l,n+1,i} - r_{l,n,i}) x_{l,n,i} \quad +$$
$$\sum_{f \in \mathcal{F}_y} \lambda_f \sum_{l:A_{l,f}=1} \left(\sum_{i=1}^{\bar{x}-1} (m^f_{l,n+1,i} - r_{l,n+1,i})(x_{l,n,i} - x_{l,n,i+1}) + (m^f_{l,n+1,\bar{x}} - r_{l,n+1,\bar{x}}) x_{l,n,\bar{x}} \right)$$

Implicit in this definition, the vector $c_{y,n}(r,m)$ has components that are themselves linear functions of r and m. Delaying a precise description for a moment, our goal is to employ the approximation

$$\sum_{l:A_{l,f}=1} m^f_{l,n,\tilde{x}_l} \sim \max((\Phi r)(\tilde{x}, n) - (\Phi r)(\tilde{x} - A_f, t), p_f),$$

for all $f \in \mathcal{F}_y$, so that $c_{y,n}(r,m)'x$ will serve as our approximation to $(T\Phi r)(s) - (\Phi r)(s)$ when $s \in \mathcal{S}_y, n(s) = n$ and $x(s)_l = \sum_i x_{l,n,i}$.

We next define the linear program $LP_{y,n}(r,m)$.

$$\begin{array}{lll} LP_{y,n}(r,m): & \max & c_{y,n}(r,m)'x \\ & \text{s.t.} & x_{l,n,0} = 1 \quad \forall l \\ & & x_{l,n,1} = 1 \quad \forall l \text{ s.t. } y_l = 1 \\ & & x_{l,n,1} = 0 \quad \forall l \text{ s.t. } y_l = 0 \\ & & x_{l,n,i+1} \leq x_{l,n,i} \quad \forall l,n,i \geq 1 \\ & & 0 \leq x_{l,n,i} \quad \forall l,n,i \geq 1 \end{array}$$

The constraint set for $LP_{y,n}(r,m)$ may be written in the form $\{x : Cx \leq b, x \geq 0\}$ where C and b have entries in $\{0,1,-1\}$. The dual to $LP_{y,n}(r,m)$ is then given by:

$$\begin{array}{ll} \min & b'z_{y,n} \\ \text{s.t.} & C'z_{y,n} \geq c_{y,n}(r,m) \\ & z_{y,n} \geq 0 \end{array}$$

so that by strong duality, our approximation to the set of constraints (4.2), i.e. (4.3), may equivalently be written as the following set of linear constraints in the variables r, m and z_y:

$$\begin{array}{ll} b'z_{y,n} \leq 0 & \forall n < N \\ C'z_{y,n} \geq c_{y,n}(r,m) & \forall n < N \\ z_{y,n} \geq 0 & \forall n < N \\ m^f_{l,n,i} \geq r_{l,n,i} & \forall\, l, n \leq N, i \in \{1,\ldots,\bar{x}\}, f \in \mathcal{F}_y \\ \sum_{l:A_{l,f}=1} m^f_{l,n,\bar{x}} \geq p_f & \forall n \leq N, f \in \mathcal{F}_y \\ m^f_{l,n,i+1} \leq m^f_{l,n,i} & \forall l, n \leq N, i \in \{1,\ldots,\bar{x}-1\}, f \in \mathcal{F}_y \end{array} \tag{4.4}$$

Assuming a starting state $s_0 = (\bar{x}, 0)$, we thus propose to minimize $(\Phi r)(s_0)$ subject to the set of constraints (4.4) for all $y \in \{0,1\}^L$ and

$$\begin{array}{ll} r_{l,n,i,m} \geq r_{l,n,i+1,m} & \forall l,n,i,m \\ r_{l,n,i,m} = 0 & \forall i,l,m; n = N \end{array}$$

in order to compute our approximation to the value function. We will refer to this program as $rALP(s_0)$.

4.6.2 Quality of Approximation

We have proposed approximating the feasible region specified by the set of constraints

$$(T\Phi r)(s) \leq (\Phi r)(s) \text{ for all } s \in \mathcal{S}$$

which has size that is $\Theta(\bar{x}^L NF)$ by a set of linear constraints of size $O(\bar{x}NLF2^L)$. There are two potential sources of error for this approximation: For one, we would ideally like to enforce the constraint $c_{y,n}(r,m)'x \leq 0$ only for $x_{\cdot,n,\cdot}$ in $\{0,1\}^{(\bar{x}+1)L}$, whereas in fact we allow $x_{\cdot,n,\cdot}$ to take values in $[0,1]^{(\bar{x}+1)L}$. It turns out that this relaxation introduces no error to the approximation, simply because the vertices of $LP_{y,n}(r,m)$ are integral. That is, the optimal solutions always satisfy $x^*_{l,n,i} \in \{0,1\}$. This is simple to verify; $LP_{y,n}(r,m)$ may be rewritten as a min-cost flow problem on a certain graph with integral supplies at the sources and sinks.

The second source of approximation error arises from the fact that we approximate $\max((\Phi r)(x,n) - (\Phi r)(x - A_f, n), p_f)$ by $\sum_{l:A_{l,f}=1} m^f_{l,n,x_l}$. In particular, we have:

$$\sum_{l:A_{l,f}=1} m^f_{l,n(s),x(s)_l} \geq \max((\Phi r)(s) - (\Phi r)(x(s) - A_f, n(s)), p_f) \tag{4.5}$$

This yields the following Lemma. A proof may be found in the appendix.

Lemma 12. $rALP(s_0) \geq ALP(s_0)$. *Moreover if* (r^{rALP}, m) *is a feasible solution to the rALP then* r^{rALP} *is a feasible solution to the ALP.*

In the case of affine approximations the reverse is true as well. That is, we have:

Lemma 13. *For affine approximations,* $rALP(s_0) \leq ALP(s_0)$. *Moreover if* r^{ALP} *is a feasible solution to the ALP then there exists a feasible solution to the rALP,* (r^{rALP}, m) *satisfying* $r^{rALP} = r^{ALP}$.

Consequently, the rALP yields the optimal solution to the ALP for affine approximations. In the case of separable concave approximations, the rALP will in general yield suboptimal solutions to the ALP. One may however show that there exists an optimal solution to the rALP satisfying for all $s \in \mathcal{S}$:

$$\sum_{l:A_{l,f}=1} m^{*,f}_{l,n(s),x(s)_l} \geq \max((\Phi r^*)(s) - (\Phi r^*)(x(s) - A_f, n(s)), p_f) \geq \frac{1}{2} \sum_{l:A_{l,f}=1} m^{*,f}_{l,n(s),x(s)_l}, \tag{4.6}$$

so that heuristically we might expect the rALP to provide solutions to the ALP that are of reasonable quality. In fact, as our computational experiments in the next subsection illustrate, the rALP appears to yield the optimal ALP solution in the case of separable concave approximations as well.

4.6.3 Computational experience with the rALP

We consider problems with 3, 4, and 8 flights with problem data generated as in the computational experiments in Section 5. Table 1 illustrates the solution objective and solution time for the rALP and ALP for each of these problems. For the 3 and 4 dimensional problems, we consider instances small enough so that it is possible to solve the ALP exactly. We see in these instances that the rALP delivers the same solution as the ALP in a far shorter time. The ALP for the 8 dimensional instance cannot be stored - let alone solved - on most conventional computers; the rALP for that problem on the other hand is relatively easy to solve and yields a near optimal solution (the comparison here being with the optimal solution of an RLP with 100,000 sampled constraints; recall that the RLP solution is a lower bound on the ALP).

In practice we envision the rALP being used in conjunction with constraint sampling. In particular, consider the following alternative to the RLP of Section 4: Let $\mathcal{X}$ be the set of sampled states one might use for the RLP. We then include in the rALP the set of constraints (4.3) for only those (y, n) such that there exists a sampled state $(x, n) \in \mathcal{X}$ with $x(s) \in \mathcal{S}_y$. The sampled rALP will have $O(\bar{x}LFK(\mathcal{X}))$ constraints where $K(\mathcal{X}) = |\{(y, n) : \exists s \in \mathcal{X} \text{ s.t. } s \in \mathcal{S}_y, n(s) = n\}|$. Since a majority of sampled states are likely to be in $\mathcal{S}_e$ where e is the vector of all ones (indicating that all fare products can potentially be serviced), one may expect $K(\mathcal{X})$ to be *far* smaller than $|\mathcal{X}|$, making the sampled rALP a significantly simpler program than the RLP. Moreover, since the sampled rALP

Dimension	$\bar{x}$	T	$ALP(s_0)$	$rALP(s_0)$	t_{ALP}	t_{rALP}
3	10	50	925.46	925.46	99.33	2.674
	20	50	1035.23	1035.23	2611.89	11.46
4	10	10	218.28	218.28	161.67	0.41
	10	30	653.37	653.37	2771.0	2.45
8	10	100	5019.90*	5028.15	849.2*	1251.92
	10	100	5019.90*	5028.12**	849.2*	177.77 **

Table 4.1: Solution quality and computation time for the rALP and ALP. * indicates values for an ***RLP*** with 100,000 constraints (recall that the RLP provides a lower bound on the ALP). ** indicates values for the sampled rALP described in section 6.3 using the same sample set as that in the computation of the corresponding RLP. Computation time reported in seconds for the CPLEX barrier LP optimizer running on a workstation with a 64 bit AMD processor and 8GB of RAM.

attempts to enforce $(T\Phi r)(s) \leq (\Phi r)(s)$ for a collection states that are a superset of the states in $\mathcal{X}$, we might expect it to provide a stronger approximation as well. While a thorough exploration of the sampled rALP is beyond the scope of this chapter, the last row of Table 1 provides encouraging supporting evidence.

4.7 Discussion and Conclusions

We have explored the use of separable concave functions for the approximation of the optimal value function for the dynamic capacity allocation problem. The approximation architecture is quite flexible and we have illustrated how it might be employed in the context of a general arrival process model wherein arrival rates vary stochastically according to a Markov process. Our computational experiments indicate that the use of the LP approach to Approximate DP along with this approximation architecture can yield significant performance gains over the DLP (of up to about 8%) , even when re-computation of DLP bid prices is allowed at every time step. Moreover, our control policy is a bid price policy where policy execution requires a table look-up at each epoch making the methodology ideally suited to real time implementation. State of the art heuristics for the dynamic capacity allocation problem typically resort to using point estimates of demand in conjunction with a model that assumes simple time homogeneous arrival processes in order to make capacity allocation decisions dynamically. As such, our algorithm may be viewed as a viable approach to moving beyond the use of point estimates and instead integrating forecasting and optimization. The approach we propose is also scalable. For example, the sampled rALP proposed in section 6 may be solved in a few minutes for quite large problems.

Our approximate DP approach offered only a marginal performance improvement relative to the DLP in the case where demand for the various fare classes were time homogenous Poisson processes. In fact, the DLP is asymptotically optimal for such demand processes (when one scales the time horizon T and the starting inventory level x_0). This is not surprising, and is essentially the consequence of an averaging effect: in particular, if demand for a fare class is a Poisson processes, then expected demand for that fare class is $\Theta(T)$ (in the time horizon T), and realized demand is with high probability within an additive factor of $\Theta(\sqrt{T})$ of expected demand. Of course, the Poisson process is not unique in this regard, and the DLP is likely to yield close to optimal performance if demand for fare classes demonstrates averaging behavior of this manner on a time-scale comparable to the sales horizon T. Demand in models $M1$ and $M2$ (for which we show a significant improvement over the DLP) *does not* demonstrate this averaging effect due to a random transition in arrival rates at $\sim T/2$. In summary, we note that if demand for fare classes is likely to experience significant shocks on a time scale that is slow relative to the time horizon, the DLP is unlikely to be near optimal and our approximate DP methods are likely to yield significant improvements.

Several issues remain to be resolved. For example, in the interest of very large-scale implementations, it would be useful to explore the use of simpler basis functions that are nonetheless capable of capturing the concavity of the true value function. In computational experiments, the rALP produced optimal solutions to the ALP; it would be interesting to establish that the programs are equivalent (as we have in the case of affine approximations). The ALP produces a tighter approximation to the true value function than does the DLP but it remains to show that the ALP policy dominates the DLP policy as well if this is at all true. Finally, a computational exploration of our approach with a highly realistic simulator such as that used by P.P.Belobaba (2001) would give a better sense for the gains that one may hope to achieve via the use of this approach in practice.

Concluding Remarks

The "estimate, then optimize" paradigm is widely used in practice. Its use stems primarily from the speed and modularity requirement for modern RM systems. Yet, it is not without its flaws. This thesis set out to address some of those flaws. In particular, we asked two questions:

- Would addressing these flaws produce a tangible impact on revenues?
- Could these flaws be addressed in a manner that is robust and efficient?

While definitive answers to either question can only be provided through real world tests and implementations, we have over the course of this thesis provided support for affirmative answers to both. In particular, we saw that accounting for the incentive to learn (in the case of the one product dynamic pricing problem) and optimization that attempts to harness all the predictive capabilities of a forecast model (for problems of network-RM), could potentially yield large performance gains. This perhaps isn't so surprising. What is noteworthy is that we were able to accomplish these goals via tractable schemes that required no more computational effort than existing approaches to these problems.

RM optimization research has typically focused on highly simplified models of customer demand. In part this is because there is a dearth of information on "true" demand models used in practice (these are highly proprietary), and in part because many of these optimization problems are hard in spite of such simplifying assumptions. Hopefully, the insights we have gained through this study encourage the use of optimization techniques that attempt to incorporate more realistic models of demand.

Appendix A

Proofs for Chapter 2

A.1 Proofs of Theorems 1 and 2

A.1.1 Existence of Solutions to the HJB Equation

Our proofs to both Theorems 1 and 2 will rely on showing the existence of a bounded solution to the HJB Equation $(HJ)(z) = 0$ for $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$. We will demonstrate the existence of a solution to the HJB Equation wherein price is restricted to some bounded interval. We will later show that the solution obtained is in fact a solution to the original HJB Equation.

Define $B = r + \frac{r}{\tilde{b}}\left(1 + \frac{e^{-1}(\tilde{a}+\tilde{x})}{\tilde{a}\alpha} + \frac{e^{-1}(\tilde{a}+\tilde{x})}{\alpha}\right)$. Let Π_B be the set of admissible price functions bounded by B, and define the Dynamic programming operator

$$(H^B J)(z) = \sup_{\pi \in \Pi_B} (H^\pi J)(z)$$

We will first illustrate the existence of a bounded solution to the HJB Equation:

$$(H^B J)(z) = 0 \tag{A.1}$$

for $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$.

For some arbitrary $N > \tilde{b}$ let us first construct a solution on the compact set $\mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}} \equiv \{(x, a, b) \in \mathcal{S} : x + a = \tilde{x} + \tilde{a}; \tilde{b} \leq b \leq N\}$ with the boundary conditions $J(x, a, N) = 0$

and $J(0, a, b) = 0$.

Lemma 14. (A.1) *has a unique bounded solution on* $\mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}}$ *satisfying* $J(x, a, N) = 0$ *and* $J(0, a, b) = 0$.

In the interest of brevity, the proof to the Lemma is omitted. Nonetheless, we provide a sketch: Upon setting $J(0, a, b) = 0$, (A.1) can be interpreted as an initial value problem of the form $\dot{J} = f(J, b)$ with $J(N) = 0$, in the space $\mathbb{R}^{\tilde{x}-1}$ equipped with the $\max$-norm. It is then routine to check the requirements for the application of a local existence Theorem for ODE's in a Banach space (such as Theorem 11.19 in Jost (2000)).

The following two Lemma's construct a solution to (A.1) on $\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$ using solutions constructed on $\mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}}$.

Lemma 15. *Let* J^N *be the unique solution to* (A.1) *on* $\mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}}$ *with* $J(x, a, N) = 0$ *and* $J(0, a, b) = 0$. *Moreover, let* $J^{N'}$ *be the unique solution to* (A.1) *on* $\mathcal{S}^{N'}_{\tilde{x},\tilde{a},\tilde{b}}$ *for some* $N' > N$ *with* $J(x, a, N') = 0$ *and* $J(0, a, b) = 0$. *Then, for* $(x, a, b) \in \mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}}$,

$$|J^N(x, a, b) - J^{N'}(x, a, b)| \leq r\frac{\tilde{a} + \tilde{x}}{\tilde{b}} \exp(-\alpha(N - b))$$

Moreover, $J^N(x, a, b) \leq \frac{re^{-1}(\tilde{a}+\tilde{x})}{\alpha\tilde{b}}$

Proof: Define $\tau_N = \inf\{t : n_t = x\} \wedge \inf\{t : b_t = N\}$. Similarly, define $\tau_{N'}$. Let $\pi^{*,N}(\cdot)$, defined on $\mathcal{S}^N_{\tilde{x},\tilde{a},\tilde{b}}$, be the greedy price with respect to J^N. Finally, define the "revenue" function $r^{*,N}_t = \frac{a_t e^{-\pi^{*,N}_t/r}\pi^{*,N}_t}{b_t}$. We then have, via an application of Lemma 4,

$$\begin{aligned} J^N(x, a, b) &= E_{z,\pi^{*,N}}\left[\int_0^{\tau_N} e^{-\alpha t} r^{*,N}_t dt\right] + E_{z,\pi^{*,N}}\left[e^{-\alpha\tau_N} J^N(x_{\tau_N}, a_{\tau_N}, b_{\tau_N})\right] \\ &= E_{z,\pi^{*,N}}\left[\int_0^{\tau_N} e^{-\alpha t} r^{*,N}_t dt\right] \end{aligned}$$

Note that this immediately yields:

$$J^N(x, a, b) \leq J^*(x, a, b) \leq J^*_{a/b}(x) \leq \frac{re^{-1}(\tilde{a} + \tilde{x})}{\alpha\tilde{b}}.$$

Now, for an arbitrary $\pi \in \Pi^B$, and the corresponding revenue function r, we have (again, via Lemma 4)

$$\begin{aligned} J^{N'}(x,a,b) &\geq E_{z,\pi}\left[\int_0^{\tau_{N'}} e^{-\alpha t} r_t dt\right] + E_{z,\pi}\left[e^{-\alpha \tau_{N'}} J^{N'}(x_{\tau_{N'}}, a_{\tau_{N'}}, b_{\tau_{N'}})\right] \\ &= E_{z,\pi}\left[\int_0^{\tau_{N'}} e^{-\alpha t} r_t dt\right] \end{aligned}$$

In particular, using the price function $\pi = \pi^{*,N}$ for $b \leq N$ and 0 otherwise, yields,

$$J^{N'}(x,a,b) \geq E_{z,\pi^{*,N}}\left[\int_0^{\tau_N} e^{-\alpha t} r_t^{*,N} dt\right] = J^N(x,a,b) \tag{A.2}$$

The same argument, applied to J^N, with the price function $\pi^{*,N'}$, yields

$$E_{z,\pi^{*,N'}}\left[\int_0^{\tau_N} e^{-\alpha t} r_t^{*,N'} dt\right] \leq J^N(x,a,b)$$

Finally, noting that on $\{\tau_{N'} > \tau_N\}$, $\tau_N \geq N - b$, we have

$$E_{z,\pi^{*,N'}}\left[\int_{\tau_N}^{\tau_{N'}} e^{-\alpha t} r_t^{*,N'} dt\right] \leq r\frac{\tilde{a}+\tilde{x}}{\tilde{b}} \exp(-\alpha(N-b))$$

Adding the two preceding inequalities, yields

$$J^{N'}(x,a,b) - r\frac{\tilde{a}+\tilde{x}}{\tilde{b}} \exp(-\alpha(N-b)) \leq J^N(x,a,b).$$

Since $J^{N'}(x,a,b) \geq J^N(x,a,b)$ by (A.2), the result follows.

□

Lemma 16. $\lim_{N\to\infty} J^N$ *exists on* $\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$*, is bounded, and solves system* (A.1)

The key step here is showing $\lim_N \frac{d}{db} J^N = \frac{d}{db} \lim_N J^N$ for all $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$; this is routine analysis given the result of the preceding Lemma and is omitted for brevity. The previous Lemma constructs a bounded solution to (A.1). We now show that this solution is in fact a solution to the original HJB Equation $(HJ)(z) = 0$ for $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$.

Lemma 17. *Let $\tilde{J}$ be a bounded solution to* (A.1)*. Then, $\tilde{J}$ is a solution to $(HJ)(z) = 0$ for $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$.*

Proof: We show the claim by demonstrating that the greedy price (in Π^B) with respect to $\tilde{J}$ is in fact attained in $[0, B)$. We begin by proving a bound on such a greedy price. Let $\pi_{\text{db}} \in \Pi^B$ be the greedy price with respect to $\tilde{J}$, and $\tau = \inf\{t : N_t = x_0\}$. We have, via Lemma 4,

$$\begin{aligned}
\tilde{J}(z) &= E_{z,\pi_{\text{db}}}\left[\int_0^\tau e^{-\alpha t}\tilde{r}_t dt\right] + E_{z,\pi_{\text{db}}}\left[e^{-\alpha\tau}\tilde{J}(z_\tau)\right] \\
&= E_{z,\pi_{\text{db}}}\left[\int_0^\tau e^{-\alpha t}\tilde{r}_t dt\right] \\
&\leq J^*(z) \\
&\leq \frac{re^{-1}(\tilde{a}+\tilde{x})}{\alpha\tilde{b}}.
\end{aligned}$$

Now let $\tilde{J}^\delta$ be the solution to (A.1) when the discount factor is $\alpha(1+\delta/b)$. Let π^δ_{db} be the corresponding greedy price. We then have from Lemma 4 and using the fact that $\tilde{J}(x, a, b+\delta) = \tilde{J}^\delta(x, a, b)$,

$$\begin{aligned}
\tilde{J}(x, a, b+\delta) &= E_{z,\pi^\delta_{\text{db}}}\left[\int_0^\tau e^{-\alpha(1+\delta/b)t}\tilde{r}^\delta_t dt\right] \\
&\geq E_{z,\pi_{\text{db}}}\left[\int_0^\tau e^{-\alpha(1+\delta/b)t}\tilde{r}_t dt\right]
\end{aligned}$$

It follows that

$$\begin{aligned}
\tilde{J}(z) - \tilde{J}(x, a, b+\delta) &\leq E_{z,\pi_{\text{db}}}\left[\int_0^\tau (e^{-\alpha t} - e^{-\alpha(1+\delta/b)t})\tilde{r}_t dt\right] \\
&\leq \int_0^\infty (e^{-\alpha t} - e^{-\alpha(1+\delta/b)t})\frac{re^{-1}(a+x)}{b}dt
\end{aligned}$$

so that

$$\frac{d}{db}\tilde{J}(z) \geq -\frac{r\alpha}{b}\frac{e^{-1}(a+x)}{b\alpha^2}$$

Putting the two bounds together yields

$$\tilde{J}(x-1,a+1,b)-\tilde{J}(z)+\frac{b}{a}\frac{d}{db}\tilde{J}(z)\geq-\frac{re^{-1}(\tilde{a}+\tilde{x})}{\alpha\tilde{b}}-\frac{re^{-1}(\tilde{a}+\tilde{x})}{\tilde{a}\tilde{b}\alpha} \tag{A.3}$$

Now observe that the greedy price $\pi_{\mathrm{db}}\in\Pi$ with respect to $\tilde{J}$ is given by

$$p=\left(r-\tilde{J}(x-1,a+1,b)+\tilde{J}(z)-\frac{b}{a}\frac{d}{db}\tilde{J}(z)\right)^{+}$$

which by (A.3) is in $[0,B)$, so that we have that $\tilde{J}$ is, in fact, a solution to $(HJ)(z)=0$ for $z\in\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$. □

A.1.2 Proofs for Theorems 1 and 2

Lemma 18.

$$\mathcal{A}_{\pi,z}J(z)=e^{-\pi(z)/r}\frac{a}{b}\left(J(z')-J(z)+\frac{b}{a}\frac{d}{db}J(z)\right)-\alpha J(z)$$

Proof: As in Theorem T1 in Section VII.2 of Bremaud (1981), one may show for $J\in\mathcal{J}$, and an arbitrary $z_0\in\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$,

$$\begin{aligned}J(z_t)=&J(z_0)+\int_0^t\left[\frac{b}{a}\frac{d}{db}J(z_s)+J(x_s-1,a_s+1,b_s)-J(z_s)\right]\frac{a_s}{b_s}e^{-p_s/r}ds\\&+\int_0^t\left[J(x_{s^-}-1,a_{s^-}+1,b_{s^-})-J(z_{s^-})\right](dN_s-\frac{a_s}{b_s}e^{-p_s/r}ds)\end{aligned}$$

It is not hard to show that that $N_s-\frac{a_s}{b_s}e^{-p_s/r}$ is a zero-mean $\sigma(z^s,p^s)$ martingale, so that we may conclude

$$\begin{aligned}&e^{-\alpha t}E[J(z_t)]-J(z_0)=\\&e^{-\alpha t}E\left[\int_0^t\left[\frac{d}{db}J(z_s)+J(x_s-1,a_s+1,b_s)-J(z_s)\right]\frac{a_s}{b_s}e^{-p_s/r}ds\right]+(e^{-\alpha t}-1)J(z_0)\end{aligned}$$

Dividing by t and taking a limit as $t\to0$ yields, via bounded convergence, the result. □

Lemma 19. *(Verification Lemma) If there exists a solution,* $\tilde{J} \in \mathcal{J}$ *to*

$$(HJ)(z) = 0$$

for all $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$*, we have:*

1. $\tilde{J}(\cdot) = J^*(\cdot)$

2. *Let* $\pi^*(\cdot)$ *be the greedy policy with respect to* $\tilde{J}$*. Then* $\pi^*(\cdot)$ *is an optimal policy.*

Proof:

Let $\pi \in \Pi$ be arbitrary. By Lemma 4

$$\begin{aligned} J^\pi(z_0) - \tilde{J}(z_0) =& E\left[\int_0^{\tau_0} e^{-\alpha s} H^\pi J(z_s) ds\right] \\ \leq& 0 \end{aligned} \tag{A.4}$$

with equality for $\pi^*(\cdot)$, since $H^{\pi^*}\tilde{J}(z) = (H\tilde{J})(z) = 0$ for all $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$. □

Now we have shown the existence of a bounded solution, $\tilde{J}$ to $(HJ)(z) = 0$ on $\mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$ in the previous section, so that the first conclusion of the Verification Lemma gives

Theorem 1. *The value function* J^* *is the unique solution in* $\mathcal{J}$ *to* $HJ = 0$*.*

The second conclusion and (A.4) in the Verification Lemma give

Theorem 2. *A policy* $\pi \in \Pi$ *is optimal if and only if* $H^\pi J^* = 0$*.*

A.2 Proofs for Section 2.5

Lemma 1. *For all* $z \in \mathcal{S}, \alpha > 0$

$$J^*(z) \leq \tilde{J}(z) \leq J^*_{\mu(z)}(x) \leq \frac{\overline{F}(p^*)p^*\mu(z)}{\alpha}.$$

where p^* *is the static revenue maximizing price.*

Proof: We begin with showing that $J^*_\lambda(\cdot)$ is concave in λ. Consider maximizing the sum of revenues from two independent systems, both of which have an initial inventory x and arrival rates λ_1 and λ_2 respectively. It is clear that the revenue maximizing policy is one which charges $\pi^*_{\lambda_1}(x_1(t))$ in the λ_1 system and $\pi^*_{\lambda_2}(x_2(t))$ in the λ_2 system. Now consider a system with no inventory constraint that at every time t must post a price for both the λ_1 and λ_2 streams, but registers sales and receives revenues from only one of the streams (w.p. $1/2$ for each). The system cannot register any sales after a total of $2x$ sales (both registered and unregistered) have occurred. Note that irrespective of the policy employed the system will register X sales where X is a Binomial$(2x, 1/2)$ random variable. Moreover it is possible to show that given X, one may generate arrivals so that the relevant arrival stream continues to be Poisson$((\lambda_1+\lambda_2)/2$. Consider a policy that charges $\pi^*_{\lambda_1}(x_1(t))$ in the λ_1 system and $\pi^*_{\lambda_2}(x_2(t))$ in the λ_2 system. The expected revenue under such a policy is precisely $(J^*_{\lambda_1}(x)+J^*_{\lambda_2}(x))/2$. Moreover, it is clear that the expected revenue for such a system under the optimal policy is $E[J^*_{(\lambda_1+\lambda_2)/2}(X)]$ where X is a Binomial$(2x, 1/2)$ random variable. A simple induction using the monotonicity of the H_λ operator establishes that $J^*_\lambda(x) - J^*_\lambda(x-1)$ is non-increasing in x so that we have by Jensen's inequality that $E[J^*_{(\lambda_1+\lambda_2)/2}(X)] \leq J^*_{(\lambda_1+\lambda_2)/2}(x)$. The concavity of $J^*_\lambda(\cdot)$ in λ follows.

Now since $J^*_\lambda(x)$ is concave in λ, Jensen's inequality gives us that $J^*_{a/b}(x) = J^*_{E[\lambda]}(x) \geq E[J^*_\lambda(x)] = \tilde{J}(z)$. Note that $J^*_\lambda(x)$ is bounded above by the value of a system with customer arrival rate λ but without a finite capacity constraint. The optimal policy in such a system is simply to charge the static revenue maximizing price, p^*, garnering a value of $\frac{\overline{F}(p^*)p^*\lambda}{\alpha}$ yielding $J^*_\lambda(x) \leq \frac{\overline{F}(p^*)p^*\lambda}{\alpha}$. □

A.3 Proofs for Section 2.7

Lemma 3. *For all* $z \in \mathcal{S}, \alpha > 0$, $J^{*,\alpha}(z) = J^{*,1}(x, a, \alpha b)$.

Proof: Consider the following coupling of the α system starting at state (x, a, b), and of the 1 system starting at state $(x, a, \alpha b)$. Let us assume that the first system is controlled by the price function $\pi_1(\cdot)$ while the second is controlled by the price function $\pi_2(\cdot)$ where $\pi_2(x, a, b) = \pi_1(x, a, b/\alpha)$. Consider the evolution of system 1 under a sample path with arrivals at $\{t_k\}$ and a corresponding binary valued sequence $\{\psi_k\}$ and of system 2 with arrivals $\{t'_k\} = \{\alpha t_k\}$ and the same binary valued sequence $\{\psi_k\}$. Now let $\{t_k \ , k \leq x\}$ be distributed as the first x points of a Poisson(λ) process where $\lambda \sim \Gamma(a, b)$. Then it is easy to verify that $\{\alpha t_k \ , k \leq x\}$ is distributed as the first x points of a Poisson(λ) process where $\lambda \sim \Gamma(a.\alpha b)$. It immediately follows that:

$$\begin{aligned} J^{\pi_1,\alpha}(x,a,b) &= E_{\lambda\sim\Gamma(a.b)}\left[\sum_{k=1}^{x}\psi_k\pi_1(t_k^-)\exp(-\alpha(t_k))\right] \\ &= E_{\lambda\sim\Gamma(a.b)}\left[\sum_{k=1}^{x}\psi_k\pi_2(\alpha t_k^-)\exp(-(\alpha t_k))\right] \\ &= E_{\lambda\sim\Gamma(a.\alpha b)}\left[\sum_{k=1}^{x}\psi_k\pi_2(t_k^-)\exp(-t_k)\right] \\ &= J^{\pi_2,1}(x,a,\alpha b) \end{aligned}$$

The result follows by taking a supremum over all price functions π_1. □

Lemma 4. *Let $J \in \mathcal{J}$ satisfy $J(0,a,b) = 0$. Let $\tau = \inf\{t : J(z_t) = 0\}$. Let $z_0 \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$. Then,*

$$E\left[\int_0^\tau e^{-\alpha t} H^\pi J(z_t)dt\right] = J^\pi(z_0) - J(z_0)$$

Let $J : \mathbb{N} \to \mathbb{R}$ be bounded and satisfy $J(0) = 0$. Let $\tau = \inf\{t : J(x_t) = 0\}$. Let $x_0 \in \mathbb{N}$. Then,

$$E\left[\int_0^\tau e^{-\alpha t} H^\pi_\lambda J(x_t)dt\right] = J^\pi_\lambda(x_0) - J(x_0)$$

Proof: Define for $J \in \mathcal{J}$, and $\pi \in \Pi$,

$$\mathcal{A}_{\pi,z}J(z) = \lim_{t>0,t\to 0} = \frac{e^{-\alpha t}E_{z,\pi}[J(z(t))] - J(z)}{t}$$

Define

$$H^\pi J(z) = \overline{F}(\pi(z))\frac{a}{b}\pi(z) + \mathcal{A}_{\pi,z}J(z)$$

Lemma 18 verifies that this definition is in agreement with our previous definition provided $J \in \mathcal{J}$. Let τ be a stopping time of the filtration $\sigma(z^t)$. We then have:

$$\begin{aligned} E\left[\int_0^\tau e^{-\alpha t} H^\pi J(z_t) dt\right] &= E\left[\int_0^\tau e^{-\alpha t}\left(\overline{F}(\pi(z_t))\frac{a}{b}\pi(z_t) + \mathcal{A}_{\pi,z}J(z_t)\right)dt\right] \\ &= J^\pi(z_0) + E_{z_0}\left[e^{-\alpha\tau}J(z_\tau)\right] - J(z_0) \\ &= J^\pi(z_0) - J(z_0) \end{aligned}$$

where the second equality follows from Dynkin's formula. The proof of the second statement is analogous. □

Lemma 5. *If* $\lambda < \mu$, $J_\lambda^{\pi^{nl}}(x) \geq \lambda/\mu J_\mu^*(x)$ *for all* $x \in \mathbb{N}$.

Proof: Letting $\tau = \inf\{t : n_t = x_0\}$ as usual, we have

$$\begin{aligned} -E\left[\int_0^\tau e^{-\alpha t} H_\lambda^{\pi^{nl}} J_\rho^*(x_t) dt\right] &= E\left[\int_0^\tau e^{-\alpha t}(1-\lambda/\rho)\alpha J_\rho^*(x_t)dt\right] \\ &\leq E\left[\int_0^\tau e^{-\alpha t}(1-\lambda/\rho)\alpha J_\rho^*(x_0)dt\right] \\ &\leq (1-\lambda/\rho)J_\rho^*(x_0) \end{aligned}$$

where the inequality follows from the fact that $J_\rho^*(x)$ is decreasing in x and since $\lambda < \rho$ here. So, from Lemma 4, we immediately have:

$$J_\rho^*(x_0) - J_\lambda^{\pi^{nl}}(x_0) \leq (1-\lambda/\rho)J_\rho^*(x_0)$$

which is the result. □

Lemma 6. *If* $\lambda \geq \mu$, $J_\lambda^{\pi^{nl}}(x) \geq J_\mu^*(x)$ *for all* $x \in \mathbb{N}$.

Proof: Here,

$$-E\left[\int_0^\tau e^{-\alpha t} H J_{\pi_0}^*(x(t))dt\right] \leq 0$$

so the result follows immediately from Lemma 4. □

Corollary 1. *For all* $z \in \mathcal{S}$*, and reservation price distributions satisfying Assumptions 1* *and 2*

$$\frac{1}{\kappa(a)} \leq \frac{\pi_{\text{db}}(z)}{\pi^*(z)} \leq 1$$

Proof: Recall that the decay balance equation implies that $\frac{\overline{F}(p^*)p^*\rho(\pi^*(z))}{\overline{F}(\pi^*(z))} = \frac{\overline{F}(p^*)p^* a}{J^*(z)b\alpha} = r^*$. Let $\frac{\overline{F}(p^*)p^*(a/b\alpha)}{\tilde{J}(z)} = \tilde{r}$. Lemma 1 implies that $r^* \geq \tilde{r} \geq 1$. It is simple to check that when the right hand side is 1, the equation is satisfied uniquely by $p = p^*$, the static revenue maximizing price. Now we have that $\pi^{\text{db}}(z) = p^* + \frac{\pi^{\text{db}}(z)-p^*}{\tilde{r}-1}(\tilde{r}-1)$ and by part 1 of Assumption 2, $\pi^*(z) \leq p^* + \frac{\pi^{\text{db}}(z)-p^*}{\tilde{r}-1}(r^*-1)$. Consequently,

$$\begin{aligned}
\frac{\pi_{\text{db}}(z)}{\pi^*(z)} &\geq \frac{p^* + \frac{\pi_{\text{db}}(z)-p^*}{\tilde{r}-1}(\tilde{r}-1)}{p^* + \frac{\pi_{\text{db}}(z)-p^*}{\tilde{r}-1}(r^*-1)} \\
&\geq \frac{p^* + \frac{\pi_{\text{db}}(z)-p^*}{\tilde{r}-1}(\tilde{r}-1)}{p^* + \frac{\pi_{\text{db}}(z)-p^*}{\tilde{r}-1}(\kappa(a)\tilde{r}-1)} \\
&\geq \frac{p^* + (\tilde{r}-1)/(\overline{F}(p^*)p^* \frac{d}{dp}\frac{\rho(p)}{\overline{F}(p)}\big|_{p=p^*})}{p^* + (\kappa(a)\tilde{r}-1)/(\overline{F}(p^*)p^* \frac{d}{dp}\frac{\rho(p)}{\overline{F}(p)}\big|_{p=p^*})} \\
&\geq \frac{1}{\kappa(a)}
\end{aligned}$$

where the second inequality follows from Theorem 3, the third inequality follows from the convexity assumed in part 1 of Assumption 2, and the final inequality follows from part 2 of Assumption 2. The upper bound is immediate from the fact that $J^*(z) \leq \tilde{J}(z)$.
□

Lemma 7. *For all $z \in \mathcal{S}$, and reservation price distributions satisfying Assumptions 1 and 2,*

$$J^{\text{ub}}(z) \geq J^*(z)$$

Proof: Define the operator:

$$(H^{\text{ub}}J)(z) = \overline{F}(\pi_{\text{db}}(z))\left(\frac{a}{b}\left(\pi^*(z) + J(z') - J(z)\right) + \frac{d}{db}J(z)\right) - e^{-1}J(z).$$

Analogous to the proof of Theorem 1, one may verify that J^{ub} is the unique bounded solution to $(H^{\text{ub}}J)(z) = 0$ for all $z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$ satisfying $J^{\text{ub}}(0,a,b) = 0$. Identically to the proof of Lemma 4, we can then show for $J \in \mathcal{J}$ satisfying $J(0,a,b) = 0$, and $z_0 \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$ that

$$E\left[\int_0^\tau e^{-\alpha t} H^{\mathrm{ub}} J(z_t) dt\right] = J^{\mathrm{ub}}(z_0) - J(z_0) \tag{A.5}$$

Now, observe that for $x > 0$,

$$\begin{aligned}
&(H^{\mathrm{ub}} J^*)(z) \\
&= \overline{F}(\pi_{\mathrm{db}}(z)) \left(\frac{a}{b} \left(\pi^*(z) + J^*(z') - J^*(z)\right) + \frac{d}{db} J^*(z) \right) - e^{-1} J^*(z) \\
&\geq \overline{F}(\pi^*(z)) \left(\frac{a}{b} \left(\pi^*(z) + J^*(z') - J^*(z)\right) + \frac{d}{db} J^*(z) \right) - e^{-1} J^*(z) \\
&= 0
\end{aligned}$$

where for the inequality, we use the fact that

$$\pi^*(z) + J^*(z') - J^*(z) + \frac{b}{a}\frac{d}{db} J^*(z) = 1/\rho(\pi^*(z)) \geq 0$$

and that $\pi_{\mathrm{db}}(z) \leq \pi^*(z)$ from Corollary 1. The equality is simply Theorem 1. We consequently have

$$H^{\mathrm{ub}} J^*(z) \geq 0 \;\; \forall z \in \mathcal{S}_{\tilde{x},\tilde{a},\tilde{b}}$$

so that (A.5) applied to J^* immediately gives:

$$J^{\mathrm{ub}}(x,a,b) \geq J^*(x,a,b)$$

□

Lemma 8. *For all $z \in \mathcal{S}, r > 0$, $J^{*,r}(z) = r J^{*,1}(z)$.*

Proof: Consider the following coupling of the r system starting at state $z = (x,a,b)$, and of the 1 system starting at state z. Let us assume that the first system is controlled by the price function $\pi_1(\cdot)$ while the second is controlled by the price function $\pi_2(\cdot) = (1/r)\pi_1(\cdot)$. Consider the evolution of both systems under a sample path with arrivals at $\{t_k\}$ and a corresponding binary valued sequence $\{\psi_k\}$ indicating whether or not the consumer chose to make a purchase. Let $E[\cdot]$ be a joint expectation over $\{t_k, \psi_k; k \leq x\}$

assuming $\{t_k\}$ are the points of a Poisson(λ) process where $\lambda \sim \Gamma(a,b)$, and ψ_k is a Bernoulli random variable with parameter $\exp(-\pi_1(t_k^-)/r) = \exp(-\pi_2(t_k^-))$. We then have:

$$\begin{aligned} J^{\pi_1,r}(z) &= E\left[\sum_{k=1}^{x} \psi_k \pi_1(t_k^-)\exp(-\alpha(t_k))\right] \\ &= rE\left[\sum_{k=1}^{x} \psi_k \pi_2(t_k^-)\exp(-\alpha(t_k))\right] \\ &= rJ^{\pi_2,1}(z) \end{aligned}$$

The result follows by taking a supremum over all price functions π_1. □

Lemma 9. *For all $z \in \mathcal{S}$,*

$$\begin{aligned} &J^*(z|\tau) \leq \\ &e^{-e^{-1}\tau}\left(e^{-(\pi^*-\pi_{\rm db})}\left[\pi^* + J^*(x-1, a+1, b_\tau^{\rm db})\right] + (1-e^{-(\pi^*-\pi_{\rm db})})J^*(x, a+1, b_\tau^{\rm db})\right) \end{aligned}$$

where $\pi^* = \pi^*(x, a, b_\tau^*)$ and $\pi_{\rm db} = \pi_{\rm db}(x, a, b_\tau^{\rm db})$.

Proof: Define $\mathcal{F}_t^{\rm db} = \sigma(z_t^{\rm db})$ and $\mathcal{F}_t^* = \sigma(z_t^*)$ Then,

$$\begin{aligned} &J^*(z|\tau) \\ &= E\left[\sum_{k=1}^{x} e^{-e^{-1}t_k}\pi^*_{t_k^-}\middle|\sigma(\mathcal{F}^{\rm db}_{\tau^-} \cup \mathcal{F}^*_{\tau^-})\right] \\ &= E\left[\sum_{k=1}^{x} e^{-e^{-1}t_k}\pi^*_{t_k^-}\middle|\lambda|\sigma(\mathcal{F}^{\rm db}_{\tau^-} \cup \mathcal{F}^*_{\tau^-}), x\right] \\ &\leq e^{-e^{-1}\tau}\left[e^{-(\pi^*-\pi_{\rm db})}[\pi^* + J^*(x-1, a+1, b_\tau^{\rm db})] + (1-e^{-(\pi^*-\pi_{\rm db})})J^*(x, a+1, b_\tau^{\rm db})\right] \end{aligned}$$

The second equality is from conditional independence of the past given the distribution of $\lambda|\sigma(\mathcal{F}_t^{\rm db} \cup \mathcal{F}_t^*)$ and x_t. For the third inequality, we note that since $\pi^*(\cdot) \geq \pi_{\rm db}(\cdot)$, and further since $\pi_{\rm db}(\cdot)$ is decreasing in b, we must have that $\pi_t^* \geq \pi_{{\rm db}t}$ on $t < \tau$. Consequently, we must have that, $b_t^* \leq b_t^{\rm db}$, on $t < \tau$, so that, $\lambda|\sigma(\mathcal{F}^{\rm db}_{\tau^-} \cup \mathcal{F}^*_{\tau^-})$ is a Gamma random variable with shape parameter $a+1$ and scale parameter, $b_\tau^{\rm db}$ so that we have,

$$E\left[\sum_{k=1}^{x} e^{-e^{-1}t_k}\pi^*_{t_k^-}\middle|\lambda|\sigma(\mathcal{F}^{\mathrm{db}}_{\tau^-}\cup\mathcal{F}^*_{\tau^-}),x\right]$$
$$\leq \sup_{\{\pi:\pi_t=\pi^*_t \text{on } t<\tau\}} E\left[\sum_{k=1}^{x} e^{-e^{-1}t_k}\pi_{t_k^-}\middle|\lambda|\sigma(\mathcal{F}^{\mathrm{db}}_{\tau^-}\cup\mathcal{F}^*_{\tau^-}),x\right]$$

which on $\{n^*_\tau = 1\}$ is equal to $e^{-e^{-1}\tau}(\pi^* + J^*(x-1, a+1, b^{\mathrm{db}}_\tau))$ and on $\{n^*_\tau = 0\}$ is equal to $e^{-e^{-1}\tau}J^*(x, a+1, b^{\mathrm{db}}_\tau)$. The inequality follows from the fact that π^* ignores the information in $\mathcal{F}^{\mathrm{db}}_{\tau^-}$. □

Lemma 10. *For $x > 1, a > 1, b > 0$, $J^*(x, a, b) \leq 2.05J^*(x-1, a, b)$.*

Proof: Let $\tau_1 = \inf\{t : n^*(t) = x-1\}$, and define

$$J^{*,\tau_1}(z) = E_{z,\pi^*}\left[\sum_{k=1}^{x-1} e^{-e^{-1}t_k}\pi_{t_k^-}\right].$$

Now,

$$J^*(z) = J^{*,\tau_1}(z) + E\left[e^{-e^{-1}\tau_1}J^*(1, a+x-1, b_{\tau_1})\right] \tag{A.6}$$

We will show that $E\left[e^{-e^{-1}\tau_1}J^*(1, a+x-1, b_{\tau_1})\right] \leq 1.05J^*(x-1, a, b)$. Since we know by definition that $J^*(x-1, a, b) \geq J^{*,\tau_1}(z)$, the result will then follow immediately from (A.6).

To show $E\left[e^{-e^{-1}\tau_1}J^*(1, a+x-1, b_{\tau_1})\right] \leq 1.05J^*(x-1, a, b)$, we will first establish a lower bound on

$$\pi^*(2, a+x-2, b_{\tau_1})/J^*(1, a+x-1, b_{\tau_1}).$$

Let $a+x-2 \equiv k, a+x-1 \equiv k'$. Certainly, $k' \leq 2k$ since $a > 1$. Now,

$$\pi^*(2, k, b) = 1 + \log k/b - \log J^*(2, k, b) \geq 1 + \log k/b - \log J^*_{k/b}(2)$$

and $J^*(1,k',b) \leq J^*(1,2k,b) \leq J^*_{2k/b}(1)$ so that

$$\frac{\pi^*(2,k,b)}{J^*(1,k',b)} \geq \frac{1+\log k - \log J^*_k(2)}{J^*_{2k}(1)}$$

But,

$$\inf_{k\in(0,\infty)} \frac{1+\log k - \log J^*_k(2)}{J^*_{2k}(1)} = \inf_{k\in(0,\infty)} \frac{1+\log k - \log W(ke^{W(k)})}{W(2k)} \geq 0.96$$

so that

$$\frac{\pi^*(2,a+x-2,b_{\tau_1})}{J^*(1,a+x-1,b_{\tau_1})} \geq 0.96$$

It follows that

$$\begin{aligned} J^*(x-1,a,b) &\geq J^{*,\tau_1}(z) \\ &\geq E[e^{-e^{-1}\tau_1}\pi^*(2,a+x-2,b_{\tau_1})] \\ &\geq 0.96\; E[e^{-e^{-1}\tau_1} J^*(1,a+x-1,b_{\tau_1})] \end{aligned}$$

Substituting in (A.6), we have the result. □

Appendix B

Proofs for Chapter 4

B.1 Proofs for Section 4.4

Lemma 11. *For model M2 with initial state s, $J^*(s) \leq ALP(s) \leq DLP(s)$*

Proof: We assume for notational convenience that N is even. Consider the following linear program:

$$
\begin{array}{lll}
sDLP(s): & \max & p'z_0 + \Pr(s_{N/2} = \text{lo})p'z_1 + \Pr(s_{N/2} = \text{hi})p'z_2 \\
 & \text{s. t.} & A(z_0 + z_1) \leq x(s) \\
 & & A(z_0 + z_2) \leq x(s) \\
 & & 0 \leq z_0 \leq \mathbb{E}[D_0] - E[D_{N/2}] \\
 & & 0 \leq z_1 \leq \mathbb{E}[D_{N/2} | s_{N/2} = \text{lo}] \\
 & & 0 \leq z_2 \leq \mathbb{E}[D_{N/2} | s_{N/2} = \text{hi}]
\end{array}
$$

It is clear that $sDLP(s) \leq DLP(s)$. This is because $z_0 + \Pr(s_{N/2} = \text{lo})z_1 + \Pr(s_{N/2} = \text{hi})z_2$ is a feasible solution to $DLP(s)$ of the same value as $sDLP(s)$. We will first show that $ALP(s) \leq sDLP(s)$. The dual to $sDLP(s)$ is given by:

$$\begin{array}{ll} \min & x(s)'y_{1,1} + x(s)'y_{1,2} + \tilde{D}_0' y_{2,0} + \tilde{D}_1' y_{2,1} + \tilde{D}_2' y_{2,2} \\ \text{s. t.} & A'(y_{1,1} + y_{1,2}) + y_{2,0} \geq p \\ & A'y_{1,1} + y_{2,1} \geq p\Pr(s_{N/2} = \text{lo}) \\ & A'y_{1,2} + y_{2,2} \geq p\Pr(s_{N/2} = \text{hi}) \\ & y_{1,1}, y_{1,2}, y_{2,0}, y_{2,1}, y_{2,2} \geq 0 \end{array}$$

where $\tilde{D}_0 = \mathbb{E}[D_0] - E[D_{N/2}]$, $\tilde{D}_1 = \mathbb{E}[D_{N/2}|s_{N/2} = \text{lo}]$ and $\tilde{D}_2 = \mathbb{E}[D_{N/2}|s_{N/2} = \text{hi}]$. Consider the following solution to the ALP for M2: Set

$$r^*_{l,n,i,med} \begin{cases} = (y^*_{1,1})_l + (y^*_{1,2})_l \\ \qquad \text{for } i > 0, n < N/2 \\ = \tilde{D}'_{0,n} y^*_{2,0} + \Pr(s_{N/2} = \text{lo}) r^*_{1,N/2,0,lo} + \Pr(s_{N/2} = \text{hi}) r^*_{1,N/2,0,hi} \\ \qquad \text{for } i = 0, l = 1, n < N/2 \\ = 0 \qquad \text{otherwise} \end{cases}$$

$$r^*_{l,n,i,lo} \begin{cases} = (y^*_{1,1})_l / \Pr(s_{N/2} = \text{lo}) & \text{for } i > 0, N > n \geq N/2 \\ = \tilde{D}'_{1,n} y^*_{2,1} / \Pr(s_{N/2} = \text{lo}) & \text{for } i = 0, l = 1, N > n \geq N/2 \\ = 0 & \text{otherwise} \end{cases}$$

$$r^*_{l,n,i,hi} \begin{cases} = (y^*_{1,2})_l / \Pr(s_{N/2} = \text{hi}) & \text{for } i > 0, N > n \geq N/2 \\ = \tilde{D}'_{2,n} y^*_{2,2} / \Pr(s_{N/2} = \text{hi}) & \text{for } i = 0, l = 1, N > n \geq N/2 \\ = 0 & \text{otherwise} \end{cases}$$

where $\tilde{D}_{0,n} = \mathbb{E}[D_n] - E[D_{N/2}]$ for $n < N/2$, $\tilde{D}_{1,n} = \mathbb{E}[D_n|s_{N/2} = \text{lo}]$ and $\tilde{D}_{2,n} = \mathbb{E}[D_n|s_{N/2} = \text{hi}]$ for $n \geq N/2$. It is routinely verified that this solution is in fact feasible for the ALP and has value equal to $sDLP(s)$. The fact that $ALP(s) \geq J^*(s)$ follows from the monotonicity of the T operator and the fact that J^* is the unique fixed point of T. This completes the proof.

B.2 Proofs for Section 4.6

Lemma 12. $rALP(s_0) \geq ALP(s_0)$. *Moreover if* (r^{rALP}, m) *is a feasible solution to the rALP then* r^{rALP} *is a feasible solution to the ALP.*

Proof: Let r^{rALP} be optimal weights from a solution to the rALP, and consider an arbitrary state $s \in \mathcal{S}$. We have

$$\begin{aligned}
& (\Phi r^{\text{rALP}})(s) \\
\geq \quad & \sum_{f: A_f \leq x(s)} \lambda_f \left((\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1) + \sum_{l: A_{l,f}=1} m^f_{l,n(s)+1,x(s)_l} \right) \\
& + \left(1 - \sum_{f: A_f \leq x(s)} \lambda_f \right) (\Phi r^{\text{dALP}})(x(s), n(s)+1) \\
\geq \quad & \sum_{f: A_f \leq x(s)} \lambda_f \Big((\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1) \\
& \quad + \max\big((\Phi r^{\text{rALP}})(x(s), n(s)+1) - (\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1), p_f \big) \Big) \\
& + \left(1 - \sum_{f: A_f \leq x(s)} \lambda_f \right) (\Phi r^{\text{rALP}})(x(s), n(s)+1) \\
= \quad & (T \Phi r^{\text{rALP}})(s)
\end{aligned}$$

where the first inequality is by the feasibility of r^{rALP}, m^* for the rALP and the second inequality is enforced by the fourth through sixth constraints in (4.4). This yields the result. □

Lemma 13. *For affine approximations,* $rALP(s_0) \leq ALP(s_0)$. *Moreover if* (r^{ALP}) *is a feasible solution to the ALP then there exists a feasible solution to the rALP,* (r^{rALP}, m) *satisfying* $r^{rALP} = r^{ALP}$.

Proof: Let r^* be the optimal solution to the ALP. For each $i \geq 1, l, f, n \leq N$, define

$$m^{*,f}_{l,n,i} = r_{l,n,i} + \left(\max \left(\sum_{l: A_{l,f}=1} r^*_{l,n,1}, p_f \right) - \sum_{l: A_{l,f}=1} r^*_{l,n,1} \right) / L(f)$$

where $L(f) = |\{l : A_{l,f} = 1\}|$. Since we are considering affine approximations, $r^*_{l,n,i'} = r^*_{l,n,i''}$ for $i', i'' > 0$. Consequently, our definition implies that for every f, n,

$$\sum_{l:A_{l,f}=1} m^{*,f}_{l,n,i} = \max\left(\sum_{l:A_{l,f}=1} r^*_{l,n,i}, p_f\right)$$

so that the feasibility of r^* for the ALP implies that $LP_{y,n}(r^*, m^*) \leq 0$ for each $y \in \{0,1\}^L, n < N$. Moreover, m^* clearly satisfies the second through fourth constraints of (4.3). This completes the proof. □

Bibliography

Adelman, D. 2005. Dynamic bid price in revenue management. To appear in Operations Research.

Araman, V., R. Caldentey. 2005. Dynamic pricing for non-perishable products with demand learning. *Submitted* .

Aviv, Y., A. Pazgal. 2005. Pricing of short life-cycle products through active learning. *Working paper* .

Ben-Akiva, Moshe, Steven Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press.

Bertsimas, D., S de Boer. 2005. Simulation-based booking limits for airline revenue management. *Operations Research* **53**(1) 90–106.

Bertsimas, D., G. Perakis. 2003. Dynamic pricing: A learning approach. *to appear in Models for Congestion Charging/Network Pricing* .

Bertsimas, D., I Popescu. 2003. Revenue management in a dynamic network environment. *Transportation Science* **37**(3) 257–277.

Besbes, O., A. Zeevi. 2006. Blind nonparametric revenue management: Asymptotic optimality of a joint learning and pricing method. *Submitted* .

Bitran, G., Susana V. Mondschein. 1997. Periodic pricing of seasonal products in retailing. *Management Science* **43**(1) 64–79.

Bremaud, P. 1981. *Point Processes and Queues: Martingale Dynamics*. 1st ed. Springer-Verlag.

Burnetas, A. N., C. E. Smith. 1998. Adaptive ordering and pricing for perishable products. *Operations Research* **48**(3) 436–443.

Cope, E. 2006. Bayesian strategies for dynamic pricing in E-commerce. *Naval Research Logistics* **54**(3) 265–281.

Dalal, S. R., W. J. Hall. 1983. Approximating priors by mixtures of natural conjugate priors. *J. R. Statist. Soc. B.* **45**(2) 278–286.

de Farias, D.P., B. Van Roy. 2003. The linear programming approach to approximate dynamic programming. *Operations Research* **51:6** 850–865.

de Farias, D.P., B. Van Roy. 2004. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research* **29:3** 462–478.

de Miguel, V, N. Mishra. 2006. What multistage stochastic programming can do for network revenue management. London Business School working paper.

E. A. Boyd, I. C. Bilegan. 2005. Revenue management and E-commerce. Pros-RM white paper. URL: `http://www.prospricing.com`.

Eren, S., C. Maglaras. 2006. Revenue management heuristics under limited market information: A maximum entropy approach. *Working paper* .

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* **40**(8) 999–1020.

Gallego, G., G. van Ryzin. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research* **45**(1) 24–41.

Jost, J. 2000. *Postmodern Analysis*. 3rd ed. Springer-Verlag.

Kleinberg, R., F. T. Leighton. 2004. The value of knowing a demand curve: Bounds on regret for on-line posted-price auctions. *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS 2003)*. 594–605.

Lin, K. Y. 2007. Dynamic pricing with real-time demand learning. *European Journal of Operational Research (forthcoming)* .

Lobo, M. S., S. Boyd. 2003. Pricing and learning with uncertain demand. *Submitted* .

Phillips, Robert. 2005. *Pricing and Revenue Optimization*. Stanford Business Books.

P.P.Belobaba. 2001. Revenue and competitive impacts of O-D control: Summary of PODS results. *First Annual INFORMS Revenue Management Section Meeting*. New York, NY.

P.P.Belobaba, S. Lee. 2000. PODS update: Large network O-D control results. *Proc. 2000 AGIFORS Reservations and Yield Management Study Group Symposium*. New York, NY.

S. Ziya, H. Ayhan, R. D. Foley. 2004. Relationships among three assumptions in revenue management. *Operations Research* **52**(5) 804–809.

Smith, Barry C., John F. Leikuhler, Ross M. Darrow. 1992. Yield management at American Airlines. (special issue: Franz Edelman award papers). *Interfaces* **22**(1) 8–32.

Talluri, Kalyan T., Garrett J. van Ryzin. 2004. *The Theory and Practice of Revenue Management*. Springer Science+Business Media.

van Ryzin, Garrett, Jeff McGill. 2000. Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels. *Management Science* **46**(6) 760–775.

Vulcano, Gustavo, Garrett van Ryzin, Costis Maglaras. 2002. Optimal dynamic auctions for revenue management. *Management Science* **48**(11) 1388–1407.