Title
Online, Data Driven Learning Approaches in Operations Management Problems

Author
Ramamurthy, Vivek

Publication Date
2012

Peer reviewed|Thesis/dissertation

Online, Data Driven Learning Approaches in Operations Management Problems

by

Vivek Ramamurthy

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering : Industrial Engineering and Operations Research

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Zuo-Jun Max Shen, Chair
Professor J. George Shanthikumar
Associate Professor Andrew E. B. Lim
Professor Peter J. Bickel

Spring 2012

Online, Data Driven Learning Approaches in Operations Management Problems

Abstract

Online, Data Driven Learning Approaches in Operations Management Problems

by

Vivek Ramamurthy

Doctor of Philosophy in Engineering : Industrial Engineering and Operations Research

University of California, Berkeley

Professor Zuo-Jun Max Shen, Chair

Traditionally, stochastic models in operations research use specific probabilistic assumptions to model random phenomena, and determine optimal policies or decisions on this basis. Often, these probabilistic assumptions are parametric, and entail estimation of parameters using very small samples of data. Many a times, the available information is not sufficient to postulate a model with any degree of certainty. Consequently, policies based on parametric assumptions in this case, are very sensitive to the particular assumptions made. One of the goals of this thesis is therefore the development of objective, adaptive, data-driven, learning approaches to objective functions, that make as few parametric assumptions as possible, and give rise to optimal policies that perform well for small samples, without compromising large sample performance. While this clearly seems a very difficult problem, it is one that is observed in nearly every operations management problem and is certainly the right problem to pursue. In this thesis, we develop novel learning approaches to specific problems in inventory control, call center staffing and dynamic assortment optimization. We test these approaches computationally, and provide strong evidence for the adoption of our general approach in tackling model uncertainty in operations management problems.

To R. Muthuswamy, my late grandfather

For the Ph.D. thesis that he ghost-wrote.

Contents

List of Figures

List of Tables

Acknowledgments

This long journey toward obtaining a doctoral degree from the prestigious UC Berkeley would not have been possible without the help and support of a lot of different people in my life. I am grateful to be able to acknowledge them and their contributions here.

First of all, I have been fortunate enough to have enjoyed the support and encouragement of not one but two thesis supervisors: J. George Shanthikumar and Max Shen. I have yet to come across in all these years, someone with the depth and breadth of knowledge that George has, along with the humility, patience, and ever-helpful spirit that he possesses. I have indeed been extremely fortunate to have had George as one of my thesis supervisors. George was the primary driving force in motivating the broad orientation of the approaches I developed in this thesis. I am grateful to him for being so patient with all of my questions during our discussions over these years. He is certainly someone that I hope I can emulate in terms of knowledge and personality. Thank you George! I would also like to express my deep gratitude to Max. He gave me complete freedom over the course of my Ph.D. to pursue all of the classes that I was interested in, and occasionally nudged me toward keeping the focus required to complete my thesis. He was always available to help with any issues I faced related to research or otherwise, and was very prompt with feedback on the many drafts that I sent him. I also benefited greatly in my research from Max's ability to focus on the big picture in all of the problems we tackled. I would also like to thank Professors Andrew Lim and Peter Bickel for being on my qualifying exam and dissertation committees. I am grateful to them for all of the inputs I received from them during my Ph.D., and also for signing off on my thesis in a timely fashion.

Of course, my very entry into the Ph.D. program might never have happened, but for the generosity of some key people toward the end of my undergraduate years at IIT Bombay. I would first like to thank Prof. P. G. Awate, who advised me on my B.Tech. Seminar and B.Tech. Project. He was kind enough to take me on as his student, and during the course of our association, helped inculcate in me the qualities of diligence, discipline, and independent thinking. These qualities stood me in great stead over the course of my Ph.D. I would also like to thank Prof. N. Hemachandra for being kind enough to write me, what was undoubtedly a great letter of recommendation for grad school, inspite of our relatively low level of prior interaction. Finally, I would like to thank Prof. N. R. Srinivasa Raghavan at IISc Bangalore, for taking me on as a Research Assistant following my B.Tech. and providing me a crucial third letter of recommendation that surely helped me a great deal, in getting admitted to Berkeley with a fellowship.

My first year at Berkeley involved a very steep learning curve, and the daunting prospect of negotiating the preliminary exam. But all of this was made a lot of fun by the students in my cohort. I would like to thank Will, Anand, Ephrat, Kory, Anthony, Mauricio, Ming, Selina, Dashi, Daphne, Xingwei, and Yao for being wonderful and interesting colleagues. I would like to thank Vishnu Narayanan for being a great mentor. He was very generous with his time in answering all of the questions I had while I was considering grad school. He was also a great resource for advice once I joined Berkeley. I am also grateful to Ankit Jain, with whom I shared not only the same undergraduate background, but also a thesis adviser. Ankit was a great resource in helping shape my research approach, and in understanding job prospects after a Ph.D. Amongst the administrative staff, I would like to especially thank Mike Campbell for being such a great facilitator and all round problem solver. I have trouble understanding how the department functioned, before his cheerful presence showed up.

Outside of academics, I have been blessed with a great circle of friends, without whom this journey would have been so much harder. In terms of social life, my two years at International House have probably been my best times at Berkeley. First off, I have been privileged to be part of "Jeff's Table", where the greatest dinner table conversations took place. In no particular order, I would like to thank Daniel Hogan, John Wywras, Kaushik Krishnan, Norma Altshuler, Jeff Crosby, Sebastian Benthall, Rondu Gantt, Andrew Critch, Bryan Cockrell, and Benjamin Baker for sharing that table with me and being really cool people. I would like to thank Xiaoqi Feng, my occasional tennis buddy, for showing up this last year, and always bringing a smile to my face. Thanks are also due to Abhinaya Chandramohan, concert buddy and frequent dinner companion. Finally, special thanks are due to the following friends. Gurkaran Buxi, for always being around when there was a genuine need, and for fun outings in his Jaguar, be it on the Pacific Coast Highway, or to a swanky whiskey bar in San Francisco. Mahendra Prasad, my fellow wannabe Renaissance man, for his wacky, and often awkward sense of humour, and also for being one of the nicest people I know. Siva Darbha, my fellow 'Wire' and classical music buff, and frequent tennis buddy. Kartikeya Date, my fellow cricket tragic, and comprehensive resource on all things cricket and Marxist. And last, but not the least, Jeff Schauer, the "proprietor" of "Jeff's Table", for being an exceptional and inspiring person. Outside of International House, I would like to thank the broader Indian graduate student community in Berkeley, with whom I shared several road trips, camping trips, poker nights, movie nights, cricket games on weekends, and enjoyable times organizing concerts for SPICMACAY. I would like to thank CHAOS, the hiking club at Berkeley, for one amazing skiing trip to Tahoe and several hiking and camping trips around the Bay Area. I would like to thank NPR and Football Weekly for being the soundtrack of my life at Berkeley. Finally, I would like to thank Berkeley for its fantastic ambience, and all the amazing people I have rubbed shoulders with here.

Of course, these acknowledgements would not be complete without mention of my family. First, I would like to thank my cousin Gayatri Gururangan, her husband Raghu Gururangan, and their sons Kapil and Karthik, my family in the Bay Area. While I didn't nearly visit them enough during my time at Berkeley, their mere presence so close to me, gave me a great sense of security and comfort. I would like to thank them for always being around when I needed their advice or help, and for always inviting me to important functions at their wonderful house. And now, most importantly, to my immediate family: my father, M. Ramamurthy, my mother, Radha Ramamurthy, and my sister, Ramya. Words cannot describe how indebted I am to them. Through several upheavals in the extended family and in their personal lives over the last six years, they have given me their full emotional and moral support in my pursuit of this Ph.D. Their joy in my accomplishment does make the pursuit worth it after all. Thank you Appa, Amma, and Ramya!

Chapter 1

Introduction

Traditionally, stochastic models in operations research use specific probabilistic assumptions to model random phenomena, and determine optimal policies or decisions on this basis. Often, these probabilistic assumptions are parametric, and entail estimation of parameters using very small samples of data. Many a times, the available information is not sufficient to postulate a model with any degree of certainty. Consequently, policies based on parametric assumptions in this case, are very sensitive to the particular assumptions made. One of the goals of this thesis is therefore the development of objective, adaptive, data-driven, learning approaches to objective functions, that make as few parametric assumptions as possible, and give rise to optimal policies that perform well for small samples, without compromising large sample performance. While this clearly seems a very difficult problem, it is one that is observed in nearly every operations management problem and is certainly the right problem to pursue. In this thesis, we develop novel learning approaches to specific problems in inventory control, call center staffing and dynamic assortment optimization. We test these approaches computationally, and provide strong evidence for the adoption of our general approach in tackling model uncertainty in operations management problems.

The inventory problem is defined as the general problem of what quantities of goods to stock in anticipation of future demand. This fundamental problem of management science has received a great deal of interest, both in theory and practice, over the last 60 years. In spite of this, firms continue to face severe inventory problems even today. Highly popular products often face chronic shortages, while huge inventory surpluses of other products often cost firms a lot of money.

In Chapter 2, we consider the single period newsvendor problem with a parametric demand distribution. The parameters of the demand distribution are unknown; moreover, there is considerable uncertainty about the distributions of the unknown parameters. Our goal is to maximize a priori expected profit in this problem, given a finite sample of past demand data, assumed to be i.i.d. We attempt to tackle the problem posed by an unknown shape parameter for the demand distribution. We suggest a heuristic approach based on operational statistics to obtain improved ordering policies when the shape parameter was unknown. In the more general cases where the heuristic based on operational statistics was not applicable, we set out to find optimal order quantities as functions of parameter estimates, which are "optimized" to perform well for small sample sizes of data. The first kind of functions we consider are linear corrections to estimates, which are essentially functions of parameter estimates. Recognizing that a "linear" correction of estimates may not always be satisfactory, we seek a richer class of functions of parameter estimates. The particular approach we use to estimate the functions from a richer class is support vector regression. In certain cases, our proposed approaches are found to yield significant improvements. While this does not establish the universal effectiveness of our approach, we view our work as a first step in considering the most general classes of parametric demand distributions and estimating inventory policies in an optimal, data-driven, fashion for small data samples.

The goal of Chapter 3 is the development of an objective operational learning approach to call center staffing. Telephone call centers are an integral part of several businesses and their economic role has grown significantly over the last decade or so. In most call centers, capacity costs in general account for $60\% - 70\%$ of operating expenses, which makes capacity management critical from a cost perspective. Recent work in statistics and operations research has begun to address the problem of how call centers and other high volume service businesses can better manage the capacity-demand mismatch that results from arrival-rate uncertainty. While each of these streams of research has made important progress in addressing elements of the problems caused by arrival-rate uncertainty, none addresses the whole problem. Statistical papers dedicated to forecasting have used standard statistical measures of fit to assess performance. They have not, however, considered the downstream cost and quality of service implications of arrival rate forecast errors. In contrast, operations management papers have looked carefully at the cost implications of stochastic scheduling methods, but they have not used the sophisticated statistical forecasting methods that best capture the nature and dynamics of arrival-rate uncertainty. Hence, there is clearly a need for an integrated, data-driven approach that marries the best aspects of the two streams of research.

In Chapter 3, we propose an objective operational learning approach to optimal staffing in a call center. Our primary goal in this approach is making minimal assumptions about the distributions of call arrivals, customer waiting times and service times, and using empirical estimates wherever possible. In this sense, our approach is 'objective'. We focus on an 'operational' quantity of interest, the cost function, and try to estimate it for various staffing levels. In the long run, as more data are available, we aim to eliminate any errors introduced by using empirical estimates of parameters, and 'learn' the true cost. Hence, we call our approach 'objective operational learning'. Our novel approach uses available cost data from a call center, probabilistically 'extends' it, and then combines this data using kernel smoothing to construct an objective operational estimate of the cost, as a function of the call profile forecast and staffing level. We test this approach on real data from a call center, and comparisons with a recent approach in the literature are seen to be very promising. We also show the pointwise asymptotic convergence of our objective operational cost estimate to the true cost function.

In Chapter 4, we develop an adaptive, non-parametric, learning approach to dynamic assortment planning (optimization). Most models for dynamic assortment optimization incorporate various models for consumer choice. A popular model is the Multinomial Logit Choice model. While this model is widely used, it does possess the somewhat restrictive property known as the independence from irrelavant alternatives (IIA) property. More generally, using any parametric demand model for consumer choice requires estimating parameters of the model first, and then incorporating these estimates into an optimization algorithm. Considering that, in practice, prior data is often insufficient to justify any particular parametric model, using estimates of such parametric models for optimizing assortments is likely to reduce the reliability, or increase the variance, of solutions found. Hence, we would like to eschew a particular parametric demand model if possible, and make minimal assumptions with respect to demand. Toward this end, we seek to develop an adaptive, non-parametric approach to dynamic assortment optimization that directly maps assortments to revenues without estimating any parameters in the process. We fully explore the applicability of the multiple play multi-armed bandit approach in the dynamic assortment planning problem. We then develop an adaptive, non-parametric approach to dynamic assortment planning that incorporates ideas from multiple play multi-armed bandit problems, and show the asymptotic optimality of our approach. We also test our approach on sales data from Amazon, and preliminary results look quite promising.

Chapter 2

Inventory Control

2.1 Introduction

The inventory problem is defined as the general problem of what quantities of goods to stock in anticipation of future demand. Loss is caused by inability to meet demand or by stocking goods for which there is no demand. Hence, an optimal policy is a tradeoff between overstocking and understocking. This fundamental problem of management science has received a great deal of interest, both in theory and practice, over the last 60 years. Inspite of this, firms continue to face severe inventory problems even today. Highly popular products often face chronic shortages, while huge inventory surpluses of other products often cost firms a lot of money. Traditionally, firms use probabilistic distributions to model stochastic demand. Parameters of such distributions are estimated using past demand data, often with additional assumptions such as independence and stationarity. The estimated parameters are then used to compute optimal inventory policies. Clearly, the success of such inventory policies is very sensitive to the choice of probability distributions. The problem is further exacerbated by the fact that typically, samples of past demand data are quite small. Hence, the high variance of the parameter estimates often results in inventory policies of very poor quality. Thus, it would be highly desirable to have adaptive inventory policies, whose regret relative to optimal policies based on full knowledge of the demand distribution is small, and which tend to the same optimal policies as the demand data increase. This is the ultimate goal of our endeavor in this chapter. Before we elaborate on our specific contributions, let us review the literature in the area of stochastic inventory control, and chart the evolution of approaches to this problem in the way they used demand data to devise optimal control policies.

Literature Review

In their seminal work, Arrow et al. [7] were among the first to formulate a theory of optimal inventory policy. The general solution of the inventory problem was further developed to a substantial extent by Dvoretzky et al. in [28] and [29] using the framework of a fairly general class of stochastic processes. Scarf [73] considered a single period problem and found a stockage policy y to maximize the minimum profit that would occur, considering all distributions with a given mean and standard deviation. Scarf [74] later revisited the problem studied by Dvoretzky et al. in [29] by considering linear cost functions in order to obtain some detailed results about the optimal stockage policies. Karlin [50] formulated an extended version of the classical Arrow-Harris-Marschak dynamic inventory model, in which the demand distributions could change from period to period. Iglehart [46] extended the results of Scarf [74] to include more general distributions and costs. Veinott [82] was concerned with a multi-product dynamic nonstationary inventory problem in which the system was reviewed at the beginning of each of a sequence of periods of equal length. Hayes [43] defined and illustrated the use of the concept of "Expected Total Operating Cost" (ETOC) in dealing with inventory policy estimation when the cost structure was piecewise linear. He showed through examples, that estimates based on classical procedures were often unsatisfactory when viewed in terms of ETOC, and derived improved estimates. Then, he considered the problem from a Bayesian point of view, and derived the prior distributions (within a particular class) that were implied by the adoption of the aforementioned superior procedures.

In the last 30 to 40 years, several extensions and generalizations to the fundamental inventory control problem have been studied. Broadly speaking, the literature may be classified as follows. The first class of approaches assumes that the demand distribution belongs to a parametric family of distributions. Under these approaches, one may choose to estimate the unknown parameters, or choose a prior distribution for the unknown parameters and apply a Bayesian approach to incorporate the demand information available. One of the first papers to explicitly distinguish sales data from true demand was Conrad [24]. Braden and Freimer [18] characterized a family of distributions for which there exist fixed-dimensional sufficient statistics of purely censored observations. Recently, several qualitative insights into the case of censored demand data have been derived by Lariviere and Porteus [57], and Ding et al. [27]. Azoury [11] considered the periodic review inventory problem, for which one or more parameters of the demand distribution were unknown with a known prior distribution chosen from the natural conjugate family. The Bayesian formulation of this problem resulted in a dynamic program with a multi-dimensional state space, which was then reduced to one dimension for some specific demand distributions. Recently, Janssen et al. [48] studied an inventory model where demand was assumed to follow a particular distribution with unknown parameters. The formulae for the order-up-to levels were corrected analytically where possible and otherwise by use of simulation and linear regression.

The second class of approaches makes no assumptions regarding the parametric form of the unknown demand distributions. For instance, Godfrey and Powell [36] considered the problem of optimizing inventories for problems where the demand distribution was unknown, and they directly estimated the value function using a technique called the Concave Adaptive Value Estimation (CAVE) algorithm. Bertsimas and Thiele [15] proposed a general methodology based on robust optimization to address the problem of optimally controlling a supply chain subject to stochastic demand in discrete time. Huh and Rusmevichientong [45] studied stochastic inventory planning with lost sales and instantaneous replenishment, where the knowledge of the demand distribution was not available. In this setting, they proposed non-parametric adaptive policies to generate ordering decisions over time, and showed that the T-period average expected cost of their policy differed from the benchmark newsvendor cost by at most $O(1/\sqrt{T})$. Levi et al. [58] considered two fundamental inventory models, the single-period newsvendor problem and its multiperiod extension, but under the assumption that the explicit demand distributions were not known and that the only information available was a set of independent samples drawn from the true distributions. They described how to compute policies based on a sampling-driven algorithmic framework, and established bounds on the expected cost of the sampling-based policies compared to those of the optimal policies, which have full access to the demand distributions. Wagner [84] considered a generalization of the Wagner-Whitin model where demands were not known and there was no information to characterize the uncertainty. Using competitive analysis, he approached the inventory management problem from a worst-case perspective and designed flexible inventory procurement strategies that had provable performance guarantees that were best possible in certain cases.

Finally, a third class of approaches assumes that the demand is partially characterized by some of the moments of its distribution. Within this class, Gallego and Moon [31] presented a new, compact proof of the optimality of Scarf's ordering rule [73] for the newsboy problem where only the mean and variance of the demand were known, and extended the analysis to several cases. Roels and Perakis [69] studied the newsvendor problem with partial information about the demand distribution (e.g., mean, variance, symmetry, unimodality). In particular, they derived the order quantities that minimized the newsvendor's maximum regret of not acting optimally. Yue et al. [87] extended previous work on the distribution-free newsvendor problem, by considering a class $\mathcal{F}$ of demand distribution functions with mean μ and standard deviation σ. They computed the maximum 'expected value of distribution information' (EVDI) over all $f \in \mathcal{F}$ for any order quantity, and provided an optimization procedure to calculate the order quantity that minimizes the maximum EVDI. Furthermore, Yue et al. [88] discussed several mean-range based distribution-free decision procedures to minimize several types of "overage" and "underage" cost functions.

Motivation

In this chapter, we consider the single period newsvendor problem with a parametric demand distribution. The parameters of the demand distribution are unknown; moreover, there is considerable uncertainty about the distributions of the unknown parameters. For example, a practical situation where this might occur is in the following newsvendor problem. Suppose it has been established from large quantities of sales data that the I-Pod Touch 1 had a Gamma demand distribution. In this case, it is reasonable to assume that the I-Pod Touch 2 would also have a Gamma demand distribution. However, due to limited sales data for the I-Pod Touch 2, it is likely that there will be a great deal of uncertainty about the parameters of its demand distribution. Our goal is to maximize a priori expected profit in this problem, given a finite sample of past demand data, assumed to be i.i.d. Toward this end, we take the approach of operational statistics, introduced in Liyanage and Shanthikumar [60]. We also review the results reported in Chu et al. [22], which enable us to maximize a priori expected profit uniformly over all parameter values, when the demand distribution is known up to the location and scale parameters. Thereafter, we attempt to tackle the problem posed by an unknown shape parameter. We suggest a heuristic approach based on operational statistics to obtain improved ordering policies when the shape parameter is unknown. In the case of a Pareto demand distribution, this approach offers improvement over a traditional ordering policy that plugs parameter estimates into an optimal policy which assumes full distributional information. However, the heuristic based on operational statistics is only applicable when the optimal order quantity is a 'separable' function of the scale and shape parameters. This is a rather special case; which prompts us to consider more general scenarios. In the more general cases where the heuristic based on operational statistics is not applicable, we set out to find optimal order quantities as functions of parameter estimates, which are 'optimized' to perform well for small sample sizes of data. The first kind of functions we consider are linear corrections to estimates. Here, estimates refer to either distribution parameter estimates, or order quantities which are computed using estimated distribution parameters. Hence, linear corrections to either of these types of estimates are essentially functions of parameter estimates. The linear correction factor is computed by minimizing the average relative regret over the possible range of values of the unknown distribution parameter under consideration. Recognizing that a 'linear' correction of estimates may not always be satisfactory, we seek a richer class of functions of parameter estimates. The particular approach we use to estimate the functions from a richer class is support vector regression. We illustrate our proposed approaches using a Gamma demand distribution. The Gamma distribution is one of the most common distributions with a shape parameter, and is ideal to model demand that is left-skewed. Furthermore, as shown in Keaton [52], the continuous Gamma distribution is ideal for modeling slow moving items, and, with appropriate scaling of the units of measure, can easily be adapted for fast-moving items as well. The Gamma distribution is highly flexible, and can assume almost any shape that one would expect of daily demand, provided the distribution is unimodal. For each type of function (linear correction or support vector regression), we consider the cases where either one of the mean or shape parameters are unknown, or both of the parameters are unknown. In certain cases, our proposed approaches are found to yield significant improvements. While this does not establish the universal effectiveness of our approach, we view our work as a first step in considering the most general classes of parametric demand distributions and estimating inventory policies in an optimal fashion for small data samples.

Outline

The rest of this chapter is organized as follows. In Section 2, we define and illustrate the approach of operational statistics, and review two main results pertaining to the case where the scale and location parameters of a distribution are unknown. In Section 3, we consider the case where the shape parameter of a distribution is also unknown. We first suggest heuristics based on operational statistics. Following that, we consider linear correction of estimates and support vector regression, and examine both approaches with a comprehensive computational study. Finally, in Section 4, we discuss the results obtained and their potential managerial implications.

2.2 Preliminaries

In this section, we consider the operational statistics approach to inventory control problems, introduced in Liyanage and Shanthikumar [60]. We first define the concept, motivate it with an example, and then review some general results concerning computation of the optimal operational statistic.

Operational Statistics : Concept and Illustration

The aim of operational statistics is to improve the current practice in the implementation of inventory policies derived assuming full knowledge of distributional parameters. In current practice, the learning needed to implement policies is done using either *classical statistical estimation procedures* or *subjective Bayesian priors*. When the data available for learning is limited, the error induced by these approaches can be substantial, as discussed in Lim et al. [59]. Operational statistics integrates the estimation and optimization tasks so as to compute the optimal statistics of interest directly. In this sense, it is an implementation of the Main Principle of Inference due to Vapnik [81] stated as:

"If you possess a restricted amount of information for solving some problem, try to solve the problem directly and never solve a more general problem as an intermediate step."

Hence, the idea is to solve the problem of optimizing the order quantity in the newsvendor problem directly, without trying to solve the more general problem of estimating the parameters of the demand distribution.

By definition, a class of operational statistics $\mathcal{S}$ is a class of functions $S(\mathbf{x}, \mathbf{z})$ of the data $\mathbf{x}$, parametrized by a set of variables $\mathbf{z} \in \mathcal{L}$. In inventory control problems, the "operational" quantity of interest is the order quantity. Hence, finding the optimal order quantity in an inventory control problem is equivalent to finding the optimal operational statistic within an appropriate class of functions, by optimizing over $\mathbf{z}$.

We now illustrate this principle using the single period newsvendor inventory control problem considered in Liyanage and Shanthikumar [60]. Items are purchased at c per unit and sold at s ($s > c$) per unit, with no salvage value. $\{X_k, k = 1, 2, \cdots, n\}$ is the sequence of i.i.d. demand data with unknown distribution function F_D, assumed here to be exponential with unknown mean θ. We wish to find the optimal order quantity for the $(n+1)$th period, so as to maximize our a priori expected profit.

A statistic S of the data $\{X_1, X_2, \cdots, X_n\}$ is defined, parametrized by some optimization variables, say $\mathbf{z}$, within an acceptable range $\mathcal{L}$. The a priori expected profit is then maximized with respect to $\mathbf{z}$. Letting

$$\hat{X}(\mathbf{z}) = S(X_1, X_2, \cdots, X_n, \mathbf{z}) \tag{2.1}$$

be the order quantity estimated from the data $\{X_1, X_2, \cdots, X_n\}$ with the optimization parameters $\mathbf{z}$, the a priori expected profit for the order quantity $\hat{X}(\mathbf{z})$ is given by

$$\eta(\mathbf{z}) = \mathbb{E}[\phi(\hat{X}(\mathbf{z}), \theta)] \tag{2.2}$$

where $\phi(x, \theta)$ is the expected newsvendor profit when demand is distributed with parameter θ and the order quantity is x. If

$$\mathbf{z}^* = \arg\max\{\eta(\mathbf{z}) : \mathbf{z} \in \mathcal{L}\} \tag{2.3}$$

then the optimal order quantity for the class of ordering policies $\hat{X}(\mathbf{z})$ is $\hat{X}(\mathbf{z}^*)$. Motivated by the ordering policies estimated with the sample mean $\overline{X}$ of the data $\{X_1, X_2, \cdots, X_n\}$, given by

$$\hat{X}_{sm} = \overline{X} \ln\left(\frac{s}{c}\right), \tag{2.4}$$

the following class of order policies (or operational statistics) is considered:

$$\hat{X}(\mathbf{a}) = \sum_{k=1}^{n} a_k Z_k, \;\; \mathbf{a} \in \mathbb{R}^n \tag{2.5}$$

where $\{Z_1, \cdots, Z_n\}$ are the spacings between the order statistics of the data $\{X_1, \cdots, X_n\}$, i.e.,

$$Z_1 = X_{[1]}$$

and

$$Z_i = X_{[i]} - X_{[i-1]} \text{ for } i = 2, \cdots, n$$

For details on the chosen form of $\hat{X}_{sm}$ and the above class of order policies, the reader is referred to Liyanage and Shanthikumar [60]. It is shown that $\hat{X}_{sm}$ belongs to the above class, and that the optimal value of $\mathbf{a}$ (which maximizes $\eta(\mathbf{a})$) is given by

$$a_k = \frac{1}{n}(n-k+1)\left(\left(\frac{s}{c}\right)^{1/(n+1)} - 1\right),\; k = 1, \cdots, n \tag{2.6}$$

The corresponding optimal order quantity (or operational statistic) is then given by

$$\hat{X}(\mathbf{a}^*) = n\left(\left(\frac{s}{c}\right)^{1/(n+1)} - 1\right)\overline{X} \tag{2.7}$$

It may be noted that this order quantity differs from $\hat{X}_{sm}$, but is a consistent estimator of the theoretical optimal order quantity (when the parameter θ is known).

Computing Optimal Operational Statistics

Chu et al. [22] propose a Bayesian analysis to find the optimal operational statistic. By specifying a non-informative prior, they show the relationship between the objective values in operational statistics and in the Bayesian analysis, and show that for all possible sample data, the Bayesian analysis derives the value of the decision variable that is an optimal operational statistic.

The single period newsvendor inventory control problem from above is considered again. First, it is assumed that $Z = D/\theta$ has a known density function f_Z, with no prior information about an unknown scale parameter θ. In this case, Chu et al. restrict the operational statistics to the class of degree one homogeneous functions $\mathcal{H}_1^+$ defined by

$$\mathcal{H}_1^+ = \{g : \mathbb{R}_+^n \to \mathbb{R}_+; g(\alpha\mathbf{x}) = \alpha g(\mathbf{x}), \alpha \geq 0, \mathbf{x} \geq \mathbf{0}\}. \tag{2.8}$$

and show that the optimal operational statistic in this class maximizes the expected a priori profit uniformly for all θ. Further, Chu et al. show that this operational statistic may also be found using a Bayesian analysis of the problem, by specifying an objective function and a prior likelihood of the unknown parameter Θ. For the objective function, they choose $\Phi(y, \Theta) = \phi(y, \Theta)/\Theta$, which is normalized for the scale effect of the demand. For the prior, they choose Jeffrey's non-informative prior $\pi(\theta) = 1/\theta$ (e.g., see Kass and Wasserman [51]).

Furthermore, they extend the above analysis to demand distributions with unknown location and scale parameters. In this case, it is assumed that $Z = (D - \tau)/\theta$ has a known density function f_Z, with no prior information on unknown parameters τ (location) and θ (scale, as before). In this case, Chu et al. restrict the operational statistics to the class of functions $\mathcal{H}_1^e$ defined by

$$\mathcal{H}_1^e = \{g : \mathbb{R}^n \to \mathbb{R}; g(\alpha(\mathbf{x} - \delta\mathbf{e})) = \alpha(g(\mathbf{x}) - \delta), \alpha \geq 0, -\infty < \delta < \infty, \mathbf{x} \in \mathbb{R}^n\}. \tag{2.9}$$

and show the optimal operational statistic in this class maximizes the expected a priori profit uniformly for all τ and θ. Further, Chu et al. show that the above operational statistic may also be found using a Bayesian analysis of the problem, by specifying an objective function and a prior likelihood for the unknown parameters Γ and Θ. For the objective function, they choose $\Phi(y, \Gamma, \Theta) = (\phi(y, \Theta) - (s - c)\Gamma)/\Theta$, which is normalized for the location and scale effect of the demand. For the prior, they choose Jeffrey's non-informative prior $\pi(\tau, \theta) = 1/\theta, -\infty < \tau < \infty; \theta > 0$ (e.g., see Kass and Wasserman [51]).

2.3 Shape Parameter Considerations

In the last section, we considered distributions for demand given by

$$D = \tau + \theta Z \tag{2.10}$$

where Z is assumed to have a known density function f_Z, though we do not assume any prior distribution for the location parameter τ and the scale parameter θ. We may also assume here, without loss of generality, that Z is a mean 1 random variable. Chu et al. [22] found an operational statistic $g(\mathbf{X})$ of the data $\mathbf{X}$, such that if they set the order quantity to $y = g(\mathbf{X})$, the expected a priori profit $\mathbb{E}_{\tau,\theta}[\phi(g(\mathbf{X}), \tau, \theta)]$ is maximized uniformly for all τ and θ.

Now, it is desired to extend the above analysis to demand distributions with three unknown parameters. These include, the location and scale parameters as considered before, and additionally a shape parameter. Here we note that almost all known probability distributions may be fully characterized by the knowledge of at most three parameters: location, scale, and shape. While distributions such as the Normal and Exponential distribution do not possess a shape parameter, other important distributions such as Beta, Gamma, Generalized Extreme Value, Log-logistic, Pareto, and Weibull possess shape parameters. Hence, if it is possible for us to find a class of operational statistics such that expected profit is maximized uniformly for all values of the location, scale and shape parameters, we would have a unified analytical approach for dealing with any parametric demand distribution with unknown parameters, in the presence of a small sample of demand data.

Hence, we now assume that the form of f_D is known up to a location parameter τ, a scale parameter θ, and a shape parameter k. The goal, as usual, is to find an operational statistic $g(\mathbf{X})$ of the data $\mathbf{X}$ so that if $y = g(\mathbf{X})$, then the expected profit $E_{\tau,\theta,k}[\phi(g(\mathbf{X}), \tau, \theta, k)]$ is maximized for all τ, θ and k.

However, herein lies the problem. A shape parameter is any parameter of a probability distribution that is neither a location parameter nor a scale parameter (nor a function of either or both of these only, such as a rate parameter). Such a parameter must affect the *shape* of a distribution rather than simply shifting it (as a location parameter does) or stretching or shrinking it (as a scale parameter does). Thus, it is possible to restrict operational statistics to specific classes of functions in the case of unknown location and scale parameters, precisely because these parameters have a fixed type of effect on the distribution. The same cannot however be said for the shape parameter, and hence it is not possible to define universal classes of functions for the shape parameter in the manner of the location and scale parameters. This requires that we modify our approach to deal with distributions having shape parameters.

Methodology: Separable case

In the following, we describe one possible approach when all three parameters of the demand distribution are unknown. Denote the location, scale and shape parameters respectively by τ, θ, and k. The following approach is only valid when the theoretical optimal order quantity (all three parameters are known) is separable into the form $g(\tau, \theta) \cdot h(k)$. When it is not separable, other approaches are in order; see, for example, Section 3.3.

Step 1: Consider the case when all parameters of the demand distribution are known. Determine the optimal order quantity.

Step 2: Verify that the optimal order quantity determined in Step 1 is separable into the form $g(\tau,\theta)\cdot h(k)$.

Step 3: Assume that the only unknown parameter is the shape parameter k. Consider the class of operational statistics suggested by the function $h(k)$. Determine the optimal operational statistic in this class. If the optimal operational statistic is a function of k, which we do not know a priori, replace occurrences of k in the formula by an appropriate statistical estimator of k.

Step 4: Next, assume that the only known parameter is the shape parameter. This reduces to the unknown scale and location parameter case studied in Chu et al. [22]. Determine the optimal operational statistic as described in Chu et al. [22].

Step 5: Consider the ordering policies hence suggested, when all the parameters are unknown.

Example : Pareto distribution

We illustrate the above approach by considering demand having the Pareto distribution, with unknown scale and shape parameters θ and k respectively. The Pareto distribution, named after the Italian economist Vilfredo Pareto, is a power law probability distribution that coincides with many social, scientific, geophysical, and actuarial phenomena. The Pareto distribution has support on $[\theta,\infty)$ for $\theta>0$. Its probability density function is given by:

$$f(x;\theta,k)=\frac{k\theta^k}{x^{k+1}} \tag{2.11}$$

and its cumulative distribution function is given by

$$F(x;\theta,k)=1-\left(\frac{\theta}{x}\right)^k \tag{2.12}$$

In this case, the expected profit function, for $x\geq 0$, is given by

$$\phi(x;\theta,k)=s\int_{y=0}^{x}\overline{F}(y;\theta,k)dy-cx=\frac{s\theta}{k-1}-\frac{sx}{k-1}\left(\frac{x}{\theta}\right)^{-k}-cx \tag{2.13}$$

Step 1: By straightforward calculus, the optimal order quantity, when both the scale and shape parameter are known, is given by

$$x^*(\theta,k)=\theta\left(\frac{s}{c}\right)^{\frac{1}{k}}=g(\theta)h(k) \tag{2.14}$$

for some functions g and h.

Step 2: Clearly, the optimal order quantity is separable into the form $g(\tau,\theta)\cdot h(k)$. We assume that we are given demand data from the past n periods $X_1,\cdots,X_n$, such that $X_i\sim$ Pareto(θ,k) are i.i.d. for all i.

Step 3: We first consider the case where θ is known and k is unknown. In this case, the Maximum Likelihood Estimator (MLE) of $1/k$ is given by

$$\hat{T} := \frac{1}{n}\sum_{i=1}^{n} \ln\left(\frac{X_i}{\theta}\right) \tag{2.15}$$

We are now required to consider the class of operational statistics suggested by $h(k)$. Since, for the case of the Pareto distribution, $h(k) = \left(\frac{s}{c}\right)^{\frac{1}{k}}$, an intuitive class of operational statistics suggested by $h(k)$ is $z^{\hat{T}} = \exp\{\hat{T}\ln z\}$, where a variable of optimization z replaces s/c and $1/k$ is replaced by its estimator $\hat{T}$, in the formula for $h(k)$. This is just a specific instance of a general approach that we propose for choosing a class of operational statistics based on $h(k)$. In general, we replace occurrences of the unknown parameter k by an appropriate statistical estimator, and replace occurrences of exogeneous parameters such as s/c, by variables to be optimized. Hence, in the above case, we consider the class of ordering policies of the form $\hat{X}(z) = \theta\exp\{\hat{T}\ln z\}$. Following the principle of operational statistics, we find the optimal value of z, which maximizes the a-priori expected profit function

$$\eta(z) = \mathbb{E}[\phi(\hat{X}(z); k)] \tag{2.16}$$

Taking the derivative w.r.t. z, and setting it to 0, we get,

$$\ln z^* = \frac{nlk}{l+k} = \frac{nl}{\frac{1}{k}\cdot l + 1} \tag{2.17}$$

where $l = \left(\frac{s}{c}\right)^{\frac{1}{n+1}} - 1$. Since by assumption, we do not know the value of $\frac{1}{k}$ a priori, we replace it with an appropriate statistical estimator, namely the MLE $\hat{T}$. Thus, we consider an order quantity based on the above operational statistic, given by

$$\hat{X}_\theta = \theta\exp\left\{\frac{\hat{T}nl}{\hat{T}l+1}\right\} \tag{2.18}$$

where the subscript θ denotes the known parameter in this case.

Step 4: Now, we consider the case where k is known and θ is unknown. Following Chu et al. [22], for the case where only the scale parameter is unknown, the optimal order quantity is given by the optimal operational statistic, computed as

$$y^* = \arg\max_{y\geq 0} \int_{\theta=0}^{\infty} \phi(y,\theta,k)\frac{1}{\theta^{n+2}}\prod_{i=1}^{n} f_Z\left(\frac{X_i}{\theta}\right)d\theta \tag{2.19}$$

The complete and sufficient statistic (and maximum value) for θ is given by

$$\hat{M} = \min\{X_1, \cdots, X_n\} \tag{2.20}$$

Hence, for the case of the Pareto distribution, y^* is given by

$$y^* = \arg\max_{y\geq 0} \int_{\theta=0}^{\hat{M}} \left[s \int_{z=\theta}^{y} \left(\frac{\theta}{z}\right)^k dz - cy \right] \frac{1}{\theta^{n+2}} \prod_{i=1}^{n} \frac{k\theta^{k+1}}{X_i^{k+1}} d\theta \tag{2.21}$$

Solving the above optimization problem, we get

$$\hat{X}_k = y^* = \left(\frac{n - \frac{1}{k}}{n + 1 - \frac{1}{k}}\right)^{\frac{1}{k}} \cdot \hat{M} \cdot \left(\frac{s}{c}\right)^{\frac{1}{k}} \tag{2.22}$$

where the subscript k denotes the known parameter in this case.

Step 5: Several ordering policies may now be considered, when both the shape and scale parameters are unknown. Our approach is to replace the occurrences of parameters in the formulae by appropriate statistical estimators, such as MLEs. To begin with, we consider the traditional ordering policy

$$\hat{X}_n = \hat{M} \left(\frac{s}{c}\right)^{\hat{R}} \tag{2.23}$$

where $\hat{M}$ is the MLE of θ and $\hat{R}$ is the MLE of $1/k$, given by

$$\hat{R} = \frac{1}{n} \sum_{i=1}^{n} \ln\left(\frac{X_i}{\hat{M}}\right) \tag{2.24}$$

The above policy is obtained by replacing the parameters in the ordering policy $x^*(\theta, k)$ (Eq. (2.14)) by their respective MLEs. Furthermore, we consider 3 other ordering policies as follows:

$$\hat{X}_1 = \left(\frac{n - \hat{R}}{n + 1 - \hat{R}}\right)^{\hat{R}} \hat{M} \left(\frac{s}{c}\right)^{\hat{R}} \tag{2.25}$$

$$\hat{X}_2 = \left(\frac{n - \hat{R}}{n + 1 - \hat{R}}\right)^{\hat{R}} \hat{M} \exp\left\{\frac{\hat{R}nl}{\hat{R}l + 1}\right\} \tag{2.26}$$

$$\hat{X}_3 = \hat{M} \exp\left\{\frac{\hat{R}nl}{\hat{R}l + 1}\right\} \tag{2.27}$$

The intuition for the above policies is as follows. Policy $\hat{X}_1$ is obtained by replacing the parameter $1/k$ in the ordering policy $\hat{X}_k$ (Eq. (2.22)), by its MLE. Policy $\hat{X}_3$ is obtained as follows. First, replace the parameter θ in the ordering policy $\hat{X}_\theta$ (Eq. (2.18)), by its MLE. Then, in the same policy, replace $\hat{T}$ by $\hat{R}$, which is the actual MLE of $1/k$, when both k and θ are unknown. In terms of the traditional policy $\hat{X}_n$, $\hat{X}_1$ is obtained by replacing $\hat{M}$

in $\hat{X}_n$ by a corrected term $\left(\frac{n-\hat{R}}{n+1-\hat{R}}\right)^{\hat{R}}\hat{M}$, while $\hat{X}_3$ is obtained by replacing $\left(\frac{s}{c}\right)^{\hat{R}}$ in $\hat{X}_n$ by a corrected term $\exp\left\{\frac{\hat{R}nl}{\hat{R}l+1}\right\}$. Applying both these corrections simultaneously in $\hat{X}_n$ then gives us policy $\hat{X}_2$, as may be easily verified. The expected a-priori profit is then given by

$$\psi_i(\theta, k) = \mathbb{E}[\phi(\hat{X}_i; \theta, k)] = \frac{s\theta}{k-1} - \frac{s\theta^k}{k-1}\mathbb{E}\left[\hat{X}_i^{1-k}\right] - c\mathbb{E}\left[\hat{X}_i\right] \tag{2.28}$$

for each i.

Analysis of Ordering Policies

The above expectations are difficult to evaluate analytically, so it is proposed that the ordering policies considered above be compared numerically. This is done by making use of the law of large numbers to appproximate the expectations by taking averages over a large number of repetitions. The number of repetitions used in our experiments is 2000. In our numerical study, we consider samples of size $n = 20$, $\theta = 1$ and cost price = 100 for a range of sale-price to cost-price ratios. The percent improvement in expected profit over the traditional policy for the 3 policies considered, is plotted against a range of k values. This is illustrated in Figure 2.1. We observe the following:

- For any value of s/c, at least one of the three policies considered, offers improvement over $\hat{X}_n$ for all the values of k considered for $n = 20$.
- This improvement decreases as k increases, as the Pareto distribution rapidly tends to a degenerate distribution for large values of k.
- The improvement is comparatively negligible as the sample size grows large, since in this case, all policies tend to the optimal policy when the distribution parameters are known.
- The magnitude of percent improvement increases with the value of s/c.
- In general, for values of $s/c < 2$, policy 1 is found to offer the most improvement. For values of $s/c > 2$, policy 3 is found to offer the most improvement.

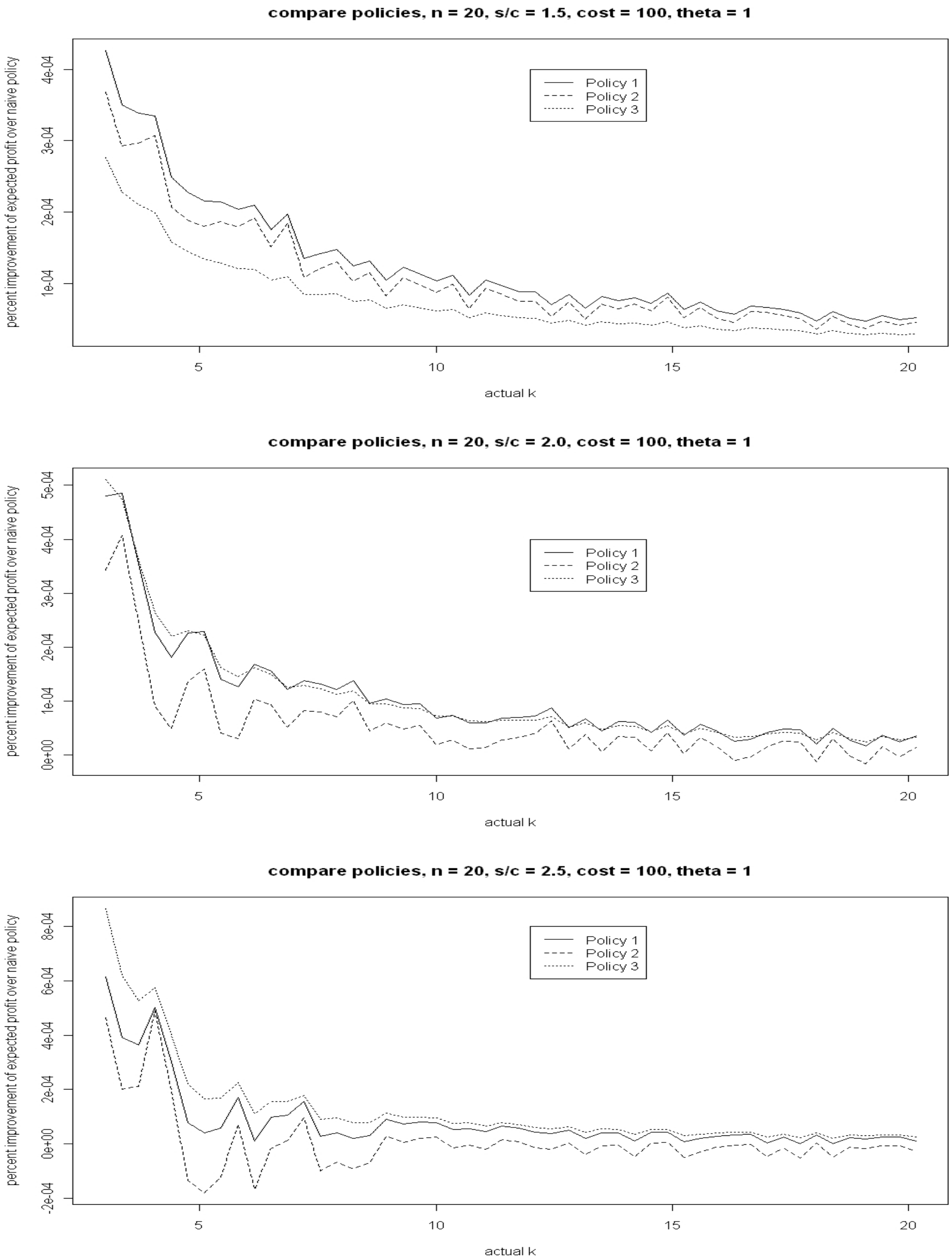

Figure 2.1: Pareto demand distribution : Percent improvement in expected profit

Methodology: non-separable case

In section 2.3, we considered the Pareto distribution and found closed-form approximations to optimal operational statistics, whose performances we then studied numerically. We used the special form of the Pareto distribution and its relation to the exponential distribution to derive these closed-form approximations. This is not always possible. For instance, to take the example of the Gamma distribution, one of the most well-known distributions with a shape parameter, there isn't even an explicit formula for the optimal order quantity in the newsvendor problem considered. Hence, at the very outset, we need to make use of numerical procedures. In the following, we consider improved inventory policy using corrections to parameter estimates, as well as corrections to optimal order quantities based on estimated parameters. We generate i.i.d. demand data over the range of values of a parameter, estimate parameter values using the data, and choose correction factors that minimize the regret relative to the case where true values of the parameter are used. For simplicity, we consider linear correction factors. The performance of these correction factors is then studied numerically, for the cases when one or more distribution parameters are unknown. Following this, to account for cases where a linear correction alone may not be justified, we take a regression approach towards determining the optimal order quantities when there is parameter uncertainty. Toward this end, we describe SV regression methods, due to Schölkopf et al. [76], which use the theory of Support Vector Machines, to determine an optimal regression fit. This approach is then numerically compared with the linear correction approaches considered.

Linear Correction of Estimates

We propose the following general methodology for linear correction of estimates. To begin with, we define an interval Ω, in which our unknown shape parameter k lies. Moreover, we assume our demand distribution has mean 1. This defines a corresponding interval for the scale parameter θ. We have no prior information about the distribution of k over the interval Ω. Hence, we pick r points uniformly from Ω and form a vector k_a. This defines a corresponding vector θ_a, of length r, as required.

The above assertion requires that, given the mean, for every shape parameter value, there exists a unique scale parameter value. The satisfaction of this requirement may be shown as follows. Without loss of generality, we suppose that the location parameter is 0. If the probability density exists for all values of the complete parameter set, then the density (as a function of the scale parameter only) satisfies

$$f_\theta(x) = f(x/\theta)/\theta \tag{2.29}$$

where f is a standardized version of the density. It then follows that

$$\mathbb{E}_{k,\theta}[X] = \int x f_\theta(x) dx = \int x \frac{f(x/\theta)}{\theta} dx = \theta \int \frac{x}{\theta} f\left(\frac{x}{\theta}\right) \frac{dx}{\theta} = \theta \int z f(z) dz = \theta \mathbb{E}_{k,1}[X] \tag{2.30}$$

where X is a random variable following the distribution under consideration. Here, $\mathbb{E}_{k,1}[X]$ is only a function of k, and not θ. It follows that when the mean is assumed to be 1, we have $\theta = (\mathbb{E}_{k,1}[X])^{-1}$, which is unique for any given k. The implicit assumption here is that $\mathbb{E}_{k,1}[X] > 0$, which may reasonably be assumed to be true for newsvendor demand distributions.

Now, we determine the corresponding optimal order quantity vector x_a; this becomes our benchmark. Next, for each value in k_a and/or θ_a, we generate a sample of size n and determine the corresponding vector of estimates (k_e and/or θ_e). We repeat the above step several times to obtain average values of k_e and/or θ_e. Finally, we compute the required linear correction factors using the vectors computed above. We illustrate this methodology in the following for a Gamma distributed demand.

Example : Gamma distribution

To be concrete, we consider the single period newsvendor problem with demand distributed according to a Gamma distribution with shape parameter k and mean 1. This in turn implies that the scale parameter is $1/k$. For such a Gamma distribution, we have,

$$F_D(x) = \frac{\gamma(k, kx)}{\Gamma(k)} \Longrightarrow \overline{F_D}(x) = \frac{\Gamma(k, kx)}{\Gamma(k)} \tag{2.31}$$

where $\gamma(k, x)$ and $\Gamma(k, x)$ are the lower incomplete and the upper incomplete gamma functions respectively, corresponding to the gamma function $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}dt$, given by

$$\gamma(k, x) := \int_0^x t^{k-1}e^{-t}dt \text{ and } \Gamma(k, x) := \int_x^\infty t^{k-1}e^{-t}dt \tag{2.32}$$

For i.i.d. data $X_1, \cdots, X_n$, the Method of Moments (MOM) estimators of the shape parameter k and the scale parameter θ, are respectively given by

$$\hat{k} = \frac{m_1^2}{m_2 - m_1^2} \tag{2.33}$$

and

$$\hat{\theta} = \frac{m_2 - m_1^2}{m_1} \tag{2.34}$$

where m_1 and m_2 are the first and second sample moments respectively, of the data. When all parameters are known, the optimal order quantity is given by

$$kx^*(k) = \overline{F_D}^{inv}\left(\frac{c}{s}\right) \Longrightarrow \frac{\Gamma(k, kx^*(k))}{\Gamma(k)} = \frac{c}{s} \tag{2.35}$$

and the expected profit function is given by

$$\phi(x, k) = s\int_0^x \frac{\Gamma(k, ky)}{\Gamma(k)}dy - cx \tag{2.36}$$

where we substitute $x^*(k)$ for x to get the expected profit function for the optimal order quantity. This is the ideal case. In practice, the values of distribution parameters are uncertain, and determined by estimators based on finitely many data. We will now numerically study the effect of correcting these estimates using linear factors, on the expected profit.

In the following, we consider a set of values for the shape parameter k ranging from $k = 1$ to $k = 16$. For $k = 1$, the Gamma distribution reduces to the Exponential distribution; moreover, for $k > 15$, the Gamma distribution rapidly tends to the Normal distribution. Hence, the most plausible range of values for the shape parameter for the Gamma distribution is considered. Moreover, the true mean of the Gamma distribution is assumed to be 1. We consider three cases:

- Known mean, unknown shape parameter
- Unknown mean, known shape parameter
- Unknown mean, unknown shape parameter

In each of the above cases, we consider the MOM estimate of either the unknown shape parameter or the unknown mean or both, an optimal linear correction to the estimate w.r.t. relative regret, and an optimal linear correction to the order quantity based on the MOM estimate, again w.r.t. relative regret. We plot the expected relative regret versus the range of k values in each case.

Known mean, unknown shape parameter We pick $r = 30$ points uniformly from the likely range of k values and form a vector k_a of length $r = 30$. Corresponding to each element k_{ai} of k_a, we determine the optimal order quantity x_{ai} and hence generate the optimal order quantity vector x_a. This is our benchmark. Next, for each value in k_a, we generate a sample of size $n = 20$ and find the corresponding MOM estimate. This is done $i = 1000$ times and the average over these repetitions is used to form a vector k_e, again of length $r = 30$ of expected MOM estimates. Now, we consider two types of linear correction factors. The first is a correction to the MOM estimate of k. This factor is computed as follows:

$$c_1 = \arg\min_{\beta} \sum_{i=1}^{r} \left(\frac{\phi(x_{ai}, k_{ai}) - \phi(x_{\beta k_{ei}}, \beta k_{ei})}{\phi(x_{ai}, k_{ai})} \right)^2 \tag{2.37}$$

Here k_a and k_e are vectors representing respectively, the range of actual k values, and the corresponding range of estimated k values, and $\phi(\cdot, \cdot)$ is the expected profit function. Moreover, for each value of β, $x_{\beta k_e}$ is the optimal order quantity vector corresponding to the parameter vector βk_e. The above problem is a nonlinear minimization problem in one variable, and is solved using the 'optimize' routine in R.

The second type of correction is a correction to the order quantity based on the MOM estimate of k. This factor is computed as follows:

$$c_2 = \arg\min_{\beta} \sum_{i=1}^{r} \left(\frac{\phi(x_{ai}, k_{ai}) - \phi(\beta x_{k_{ei}}, k_{ei})}{\phi(x_{ai}, k_{ai})} \right)^2 \tag{2.38}$$

Here k_a, k_e and $\phi(\cdot,\cdot)$ are as above, and x_{k_e} is the optimal order quantity vector corresponding to the parameter vector k_e. The above problem is also a nonlinear minimization problem in one variable, and is again solved using the 'optimize' routine in R.

We now consider a sale price to cost price ratio of $s/c = 1.2$, with a scaling factor of 1. Having computed the correction factors c_1 and c_2, the expected relative regret functions corresponding to the various cases are approximately computed as follows. A new sample of size $n = 20$ is generated, k_e is estimated, and the correction factors are applied to determine the relative regret functions. Repeating this procedure $j = 1000$ times, we get approximations to the expected relative regret. These functions correspond to the MOM estimate, the corrected MOM estimate, and the corrected order quantity over the whole range of values of k. We notice that there is no one method that dominates the other over the whole range of k values. Hence, we break the interval of k values into sub-intervals, and estimate correction factors over these sub-intervals. On plotting the expected relative regret functions, we see that the corrected order quantity dominates the corrected MOM estimate and the plain MOM estimate. We also note that over the full range of k values, the improvement due to the correction factors is large for the case of low k values. The difference between the various procedures becomes almost negligible for larger values of k, which may be attributed to the negligible variance, and consequent degeneracy, of the mean 1 Gamma distribution for larger values of k. See Figure 2.2 for details.

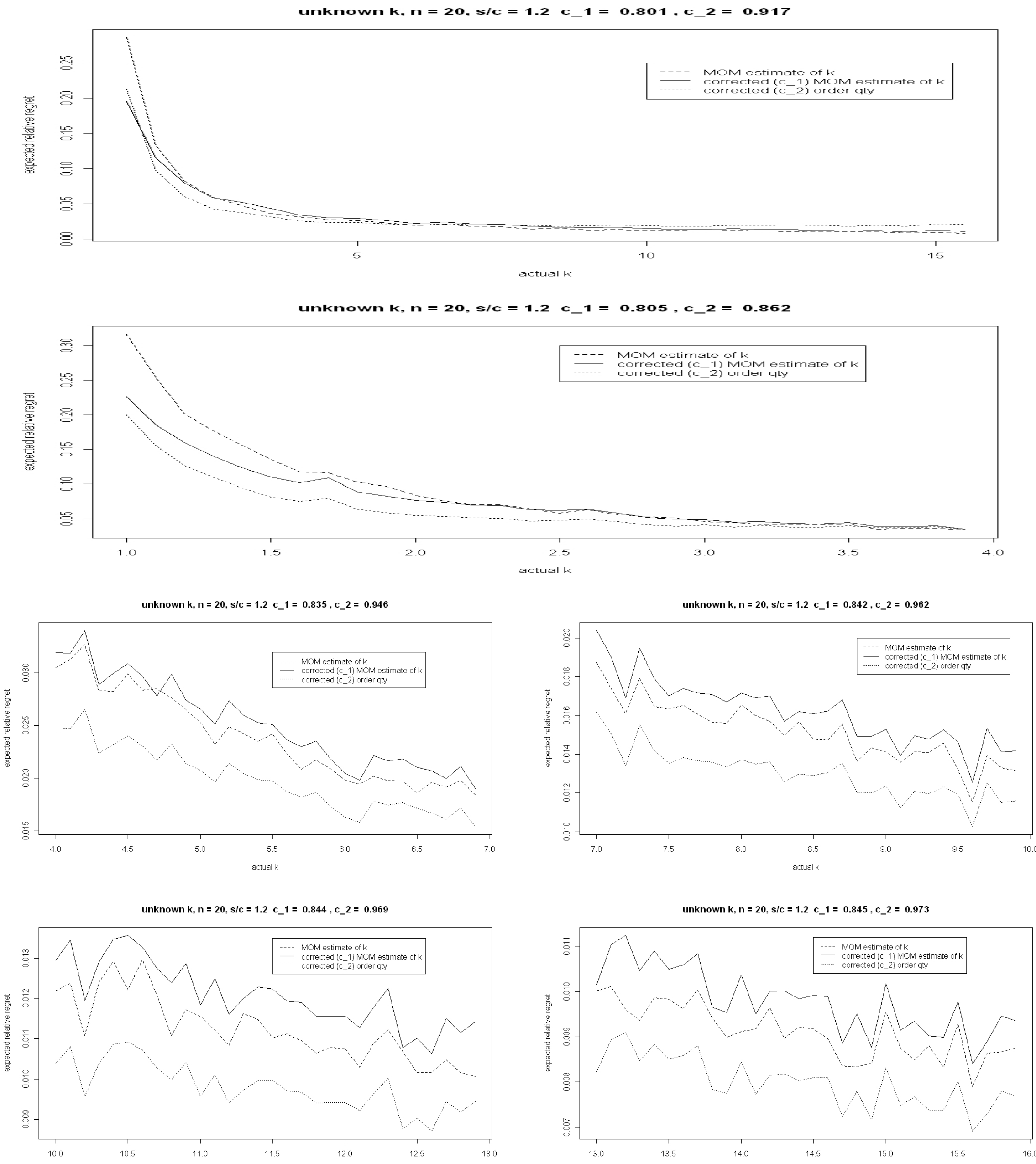

Figure 2.2: Known mean, unknown shape parameter

Furthermore, we proceed to study the effect of varying s/c on the expected relative regret functions. We observe that expected relative regret increases as s/c decreases towards 1. It is also seen that the linear corrections offer significant improvements, for low values of k, as the value of s/c is decreased towards 1. Shown in Figure 2.3 is a sample graph for the case of $n = 20$ and $s/c = 1.01$.

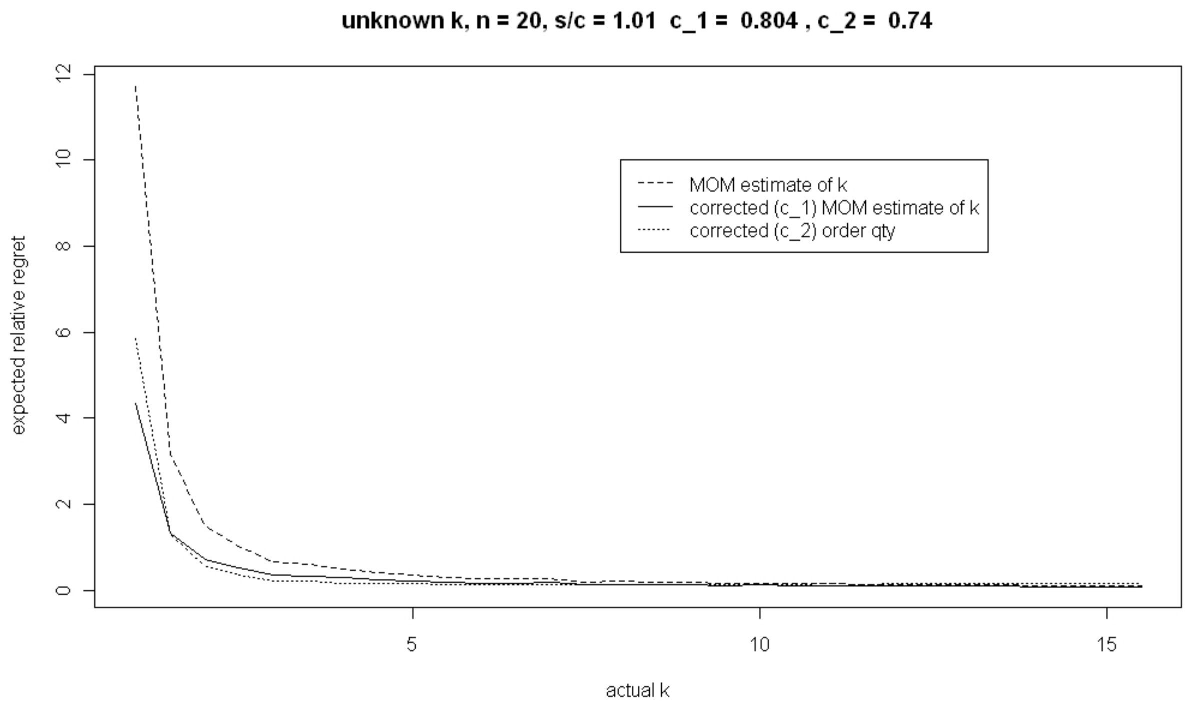

Figure 2.3: Low s/c

Finally, the effect of increasing sample size is considered. As expected, the magnitude of the relative regret is greatly reduced, as is the quantitative difference between the various procedures.

Unknown mean, known shape parameter In this case, the implication is that θ is unknown and needs to estimated. Again, we pick $r = 30$ points uniformly from the likely range of k values and form a vector k_a of length $r = 30$. This simultaneously defines a vector of corresponding scale parameter values θ_a of length $r = 30$, where $\theta_{ai} = 1/k_{ai}$ for all i, using the fact that the true mean is 1. Corresponding to k_a, we determine the optimal order quantity vector x_a. This is our benchmark. Next, for each value in θ_a, we generate a sample of size $n = 20$ and find the corresponding MOM estimate, using the known values of k. This is done $i = 1000$ times and the average over these repetitions is used to form a vector θ_e, again of length $r = 30$ of expected MOM estimates. Now, we consider a correction to the

MOM estimate of θ. This factor is computed as follows:

$$c_1 = \arg\min_{\beta} \sum_{i=1}^{r} \left(\frac{\phi(x_{ai}, \theta_{ai}) - \phi(x_{\beta\theta_{ei}}, \beta\theta_{ei})}{\phi(x_{ai}, \theta_{ai})} \right)^2 \tag{2.39}$$

where θ_a, θ_e and $\phi(\cdot, \cdot)$ are as defined previously. Moreover, for each value of β, $x_{\beta\theta_e}$ is the optimal order quantity vector corresponding to the parameter vector $\beta\theta_e$. The above problem is a nonlinear minimization problem in one variable, and is solved using the 'optimize' routine in R.

We do not separately consider a linear correction factor for the order quantity based on the MOM estimate, since the set of all such order quantities is part of the feasible set of the optimization problem considered above.

Again, we consider a sale price to cost price ratio of $s/c = 1.2$, with a scaling factor of 1. Having computed the correction factor c_1, the expected relative regret functions corresponding to the various cases are approximately computed as follows. A new sample of size $n = 20$ is generated, θ_e is estimated, and the correction factor is applied to determine the relative regret functions. Repeating this procedure $j = 1000$ times, we get approximations to the expected relative regret. These functions correspond to the MOM estimate and the corrected MOM estimate. It is seen that both procedures have nearly the same expected relative regret over the entire k range, with negligible difference. So, we break the interval of k values into sub-intervals, and estimate correction factors over these sub-intervals. In this case too, it is seen that the values follow those in the case of the full interval.

Furthermore, we proceed to study the effect of varying s/c on the expected relative regret functions. It is seen that the performance of the linear correction is largely insensitive to the value of s/c, with maximum expected relative regret decreasing slightly as the value of s/c increases. Shown in Figure 2.4 is a sample graph for the case of $n = 20$ and $s/c = 1.01$. This tells us that the linear correction approach is rather unnecessary when the shape parameter is already known. Traditional approaches would serve us well when the mean is the only unknown parameter.

Finally, we study the effect of increasing sample size. As expected, the magnitude of the maximum expected relative regret is greatly reduced.

Unknown mean, unknown shape parameter As before, we pick $r = 30$ points uniformly from the likely range of k values and form a vector k_a of length $r = 30$. This simultaneously defines a vector of corresponding scale parameter values θ_a of length $r = 30$, where $\theta_{ai} = 1/k_{ai}$ for all i, using the fact that the mean is 1. Corresponding to k_a, we determine the optimal order quantity vector x_a. This is our benchmark. Next, for each value in k_a, we generate a sample of size $n = 20$ and find the corresponding MOM estimates of k and θ. This is done $i = 1000$ times and the average over these repetitions is used to form respective vectors k_e and θ_e, again of length $r = 30$, of expected MOM estimates. Now, we

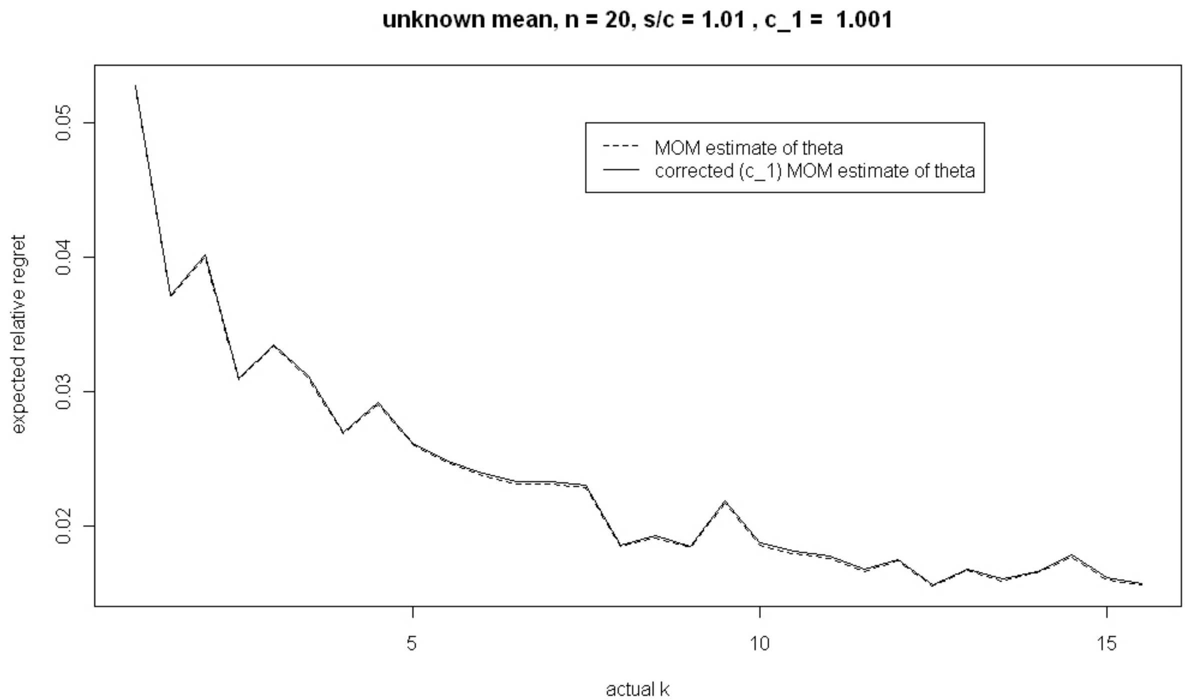

Figure 2.4: Unknown mean, known shape parameter

correct both k_e and θ_e by linear factors, which are simultaneously determined as follows:

$$(c_1, c_2) = \arg\min_{\beta_1,\beta_2} \sum_{i=1}^{r} \left(\frac{\phi(x_{ai}, k_{ai}, \theta_{ai}) - \phi(x_{\beta_1 k_{ei}, \beta_2 \theta_{ei}}, \beta_1 k_{ei}, \beta_2 \theta_{ei})}{\phi(x_{ai}, k_{ai})} \right)^2 \tag{2.40}$$

where k_a, k_e, θ_e are as defined $\phi(\cdot,\cdot,\cdot)$ is the expected profit function, defined now, using two parameters k and θ, as follows:

$$\phi(x, k, \theta) = s \int_0^x \frac{\Gamma(k, y/\theta)}{\Gamma(k)} dy - cx \tag{2.41}$$

Moreover, for each value of β_1 and β_2, and for each $i = 1, \cdots, r$, $x_{\beta_1 k_{ei}, \beta_2 \theta_{ei}} = x^*(\beta_1 k_{ei}, \beta_2 \theta_{ei})$, where $x^*(k, \theta)$ is given by

$$\frac{\Gamma(k, x^*(k,\theta)/\theta)}{\Gamma(k)} = \frac{c}{s} \tag{2.42}$$

and is defined as the optimal order quantity corresponding to known parameters k (shape) and θ (scale) of a Gamma distribution. The above problem is a nonlinear minimization problem in two variables, and is solved using the 'optim' routine in R.

As in the unknown mean, known shape parameter case, we do not separately consider a linear correction factor for the order quantity based on the MOM estimates. The set of all such order quantities is already part of the feasible set of the optimization problem considered above.

Again, we consider a sale price to cost price ratio of $s/c = 1.2$, with a scaling factor of 1. Having computed the correction factors c_1 and c_2, the expected relative regret functions corresponding to the cases of plain estimation and corrected estimation are approximately computed as follows. A new sample of size $n = 20$ is generated, k_e and θ_e are estimated, and the correction factors are applied to determine the relative regret function. Repeating this procedure $j = 1000$ times, we get approximations to the expected relative regret. These functions correspond to the MOM estimate, and the corrected MOM estimates over the whole range of values of k. We notice that for low values of k, the correction helps, while it does not for high values of k. So, as before, we break the interval of k values into sub-intervals, and estimate correction factors over these sub-intervals. On plotting the expected relative regret functions, we see that the performance of the linear correction remains about the same, as when the full interval is considered. We also note that over the full range of k values, the improvement due to the correction factors is relatively large for the case of low k values, while the difference between the two procedures is within about 5 percent for larger values of k. See Figure 2.5 for details.

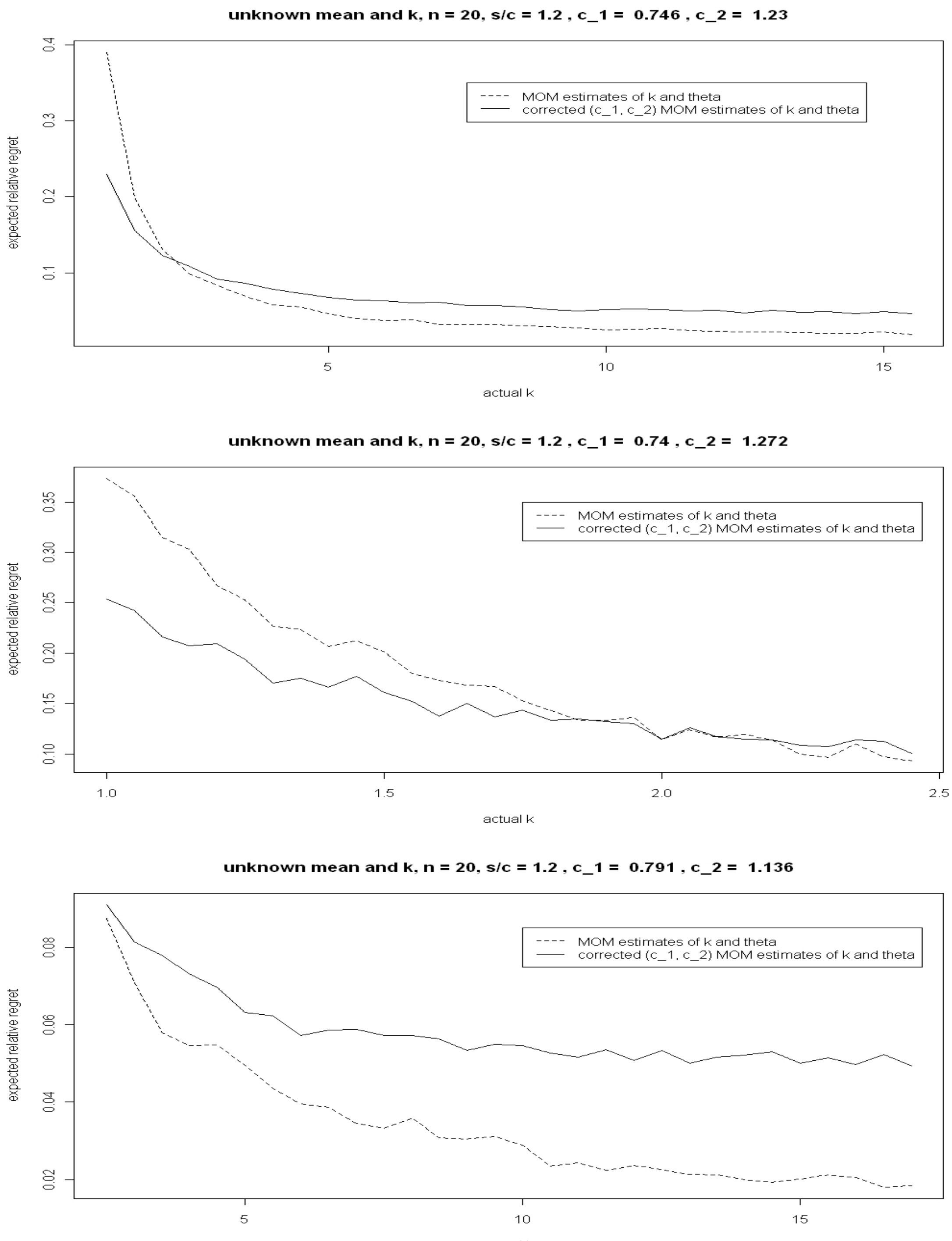

Figure 2.5: Unknown mean, unknown shape parameter

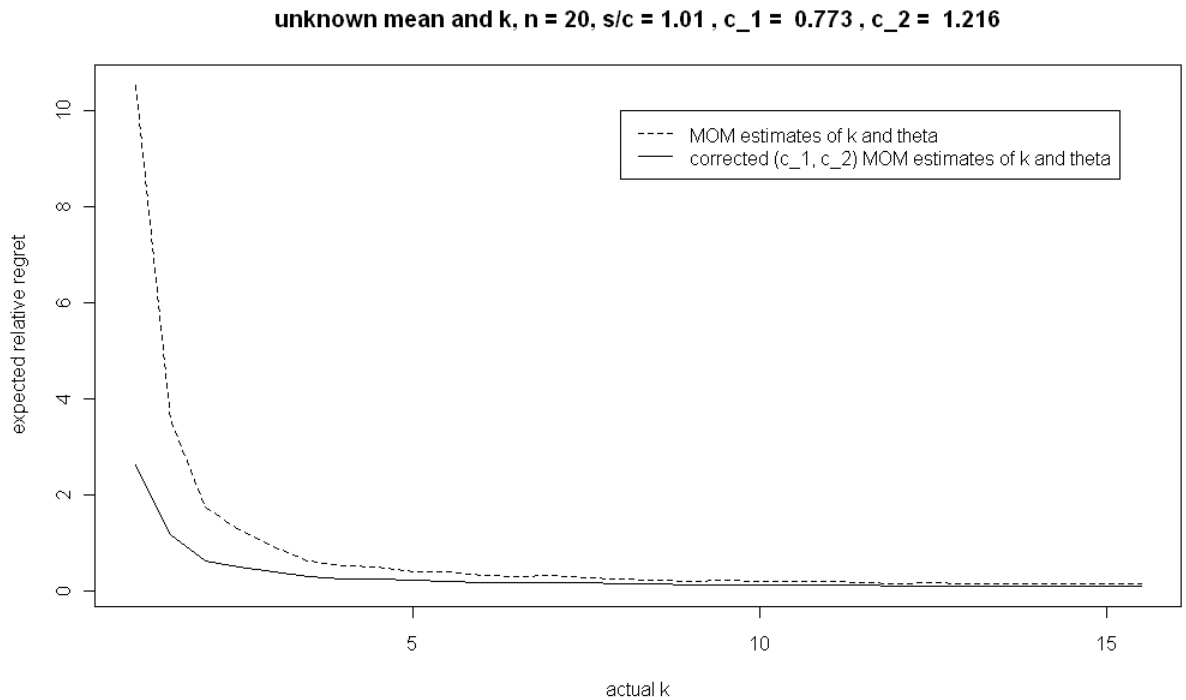

Figure 2.6: Low s/c

Furthermore, we proceed to study the effect of varying s/c on the expected relative regret functions. We observe that expected relative regret increases as s/c decreases towards 1. It is also seen that the linear corrections offer significant improvements, for low values of k, as the value of s/c is decreased towards 1. Figure 2.6 shows a sample graph for the case of $n = 20$ and $s/c = 1.01$.

Finally, we study the effect of increasing sample size. As expected, the magnitude of the relative regret is greatly reduced, as is the quantitative difference between the two procedures.

Support Vector Regression

In the previous section, we considered linear corrections of estimates. However, it is conceivable that a linear correction may not be satisfactory, and a higher order correction may be desired. We may wish to express the relationship between parameter estimates and optimal inventory policy using a richer class of functions. Toward this end, we consider the application of Support Vector Regression to minimize the expected relative regret in the newsvendor model. Support Vector Machines (SVMs) were first developed for pattern recognition. They represent the decision boundary in terms of a typically small subset of all training samples - the support vectors. This particular sparseness property needed to be retained to construct loss functions when the SV algorithm was generalized to the case of regression estimation (that is, to the estimation of *real-valued* functions, rather than just $\{\pm 1\}$-valued ones, as in the case in pattern recognition).

Theoretical Background

To construct an SVM for real-valued functions, Vapnik [81] devised a new type of loss functions, the so-called ϵ-insensitive loss functions, given by

$$L(y, f(x,\alpha)) = L(|y - f(x,\alpha)|_\epsilon) \tag{2.43}$$

where he set

$$|y - f(x,\alpha)|_\epsilon = \begin{cases} 0 & \text{if } |y - f(x,\alpha)| \le \epsilon \\ |y - f(x,\alpha)| - \epsilon, & \text{otherwise} \end{cases} \tag{2.44}$$

These loss functions describe the ϵ-insensitive model: The loss is equal to 0 if the discrepancy between the predicted and the observed values is less than ϵ.

The basic SV regression algorithm, henceforth called ϵ-SVR, seeks to estimate affine functions of the form,

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b, \text{ where } \mathbf{w}, \mathbf{x} \in \mathcal{H}, b \in \mathbb{R} \tag{2.45}$$

based on i.i.d. data,

$$(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_n, y_n) \in \mathcal{H} \times \mathbb{R} \tag{2.46}$$

where $\mathcal{H}$ is an inner-product space in which the (mapped) input patterns live (i.e., the feature space induced by a kernel). The goal of the learning process is to find a function f with a small risk (or test error) [75]

$$R[f] = \int c(f, \mathbf{x}, y) dP(\mathbf{x}, y), \tag{2.47}$$

where P is the probability measure which is assumed to be responsible for the generation of the observations (2.46) and c is a loss function, such as $c(f, \mathbf{x}, y) = (f(\mathbf{x}) - y)^2$, or one of many other possible choices. In practice, we are given the sample (2.46), and we try to obtain a small risk by minimizing the regularized risk functional,

$$\frac{1}{2}\|\mathbf{w}\|^2 + C \cdot R^\epsilon_{emp}[f] \tag{2.48}$$

where

$$R^\epsilon_{emp}[f] := \frac{1}{n}\sum_{i=1}^{n} |y_i - f(\mathbf{x}_i)|_\epsilon \tag{2.49}$$

measures the ϵ-insensitive training error, i.e., the average ϵ-insensitive loss of the estimates $f(\mathbf{x}_1), \cdots, f(\mathbf{x}_n)$ with respect to the training data $y_1, \cdots, y_n$. Moreover, C is a constant determining the trade-off of the training error with a penalty function $\|\mathbf{w}\|^2$. In short, minimizing (2.48) captures the main insight of statistical learning theory, stating that in order to obtain a small risk, we need to control both training error and model complexity, by explaining the data with a simple model (Schölkopf and Smola [75]). The minimization of (2.48) is equivalent to solving the following constrained optimization problem (Schölkopf and Smola [75]):

$$\min_{\mathbf{w}\in\mathcal{H},\xi^{(*)}\in\mathbb{R}^n,b\in\mathbb{R}} \tau(\mathbf{w},\xi^{(*)}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\cdot\frac{1}{n}\sum_{i=1}^{n}(\xi_i+\xi_i^*), \tag{2.50}$$

$$\text{subject to } (\langle\mathbf{w},\mathbf{x}_i\rangle+b)-y_i \le \epsilon+\xi_i, \text{ for } i=1,\cdots,n \tag{2.51}$$

$$y_i-(\langle\mathbf{w},\mathbf{x}_i\rangle+b) \le \epsilon+\xi_i^*, \text{ for } i=1,\cdots,n \tag{2.52}$$

$$\xi_i^{(*)} \ge 0. \text{ for } i=1,\cdots,n \tag{2.53}$$

Here it is understood that bold face Greek letters denote n-dimensional vectors of the corresponding variables; $^{(*)}$ is a shorthand implying both the variables with and without asterisks. The regression function is obtained from the Lagrange multiplier conditions of the above optimization problem (Schölkopf and Smola [75]).

The parameter ϵ of the ϵ-insensitive loss is useful if the desired accuracy of the approximation can be specified beforehand. Sometimes, however, we just want the estimate to be as accurate as possible, without having to specify a level of accuracy a priori. In the following, we describe a modification of the ϵ-SVR algorithm, called ν-SVR, which automatically computes ϵ (Schölkopf et al. [76]).

To estimate functions (2.45) from empirical data (2.46) we proceed as follows. At each point $\mathbf{x}_i$, we allow an error ϵ. Everything above ϵ is captured in slack variables $\xi_i^{(*)}$, which are penalized in the objective function via a regularization constant C, chosen a priori. The size of ϵ is traded off against model complexity and slack variables via a constant $\nu \ge 0$, giving the primal problem:

$$\min_{\mathbf{w}\in\mathcal{H},\xi^{(*)}\in\mathbb{R}^n,b\in\mathbb{R}} \tau(\mathbf{w},\xi^{(*)},\epsilon) = \frac{1}{2}\|\mathbf{w}\|^2 + C\cdot\left(\nu\epsilon+\frac{1}{n}\sum_{i=1}^{n}(\xi_i+\xi_i^*)\right), \tag{2.54}$$

$$\text{subject to } (\langle\mathbf{w},\mathbf{x}_i\rangle+b)-y_i \le \epsilon+\xi_i, \text{ for } i=1,\cdots,n \tag{2.55}$$

$$y_i-(\langle\mathbf{w},\mathbf{x}_i\rangle+b) \le \epsilon+\xi_i^*, \text{ for } i=1,\cdots,n \tag{2.56}$$

$$\epsilon \ge 0, \xi_i^{(*)} \ge 0. \text{ for } i=1,\cdots,n \tag{2.57}$$

By standard Lagrange multiplier techniques we get the dual optimization problem, which we state in the kernelized form, using the kernel k given by

$$k(\mathbf{x},\mathbf{x}') = \langle\Phi(\mathbf{x}),\Phi(\mathbf{x}')\rangle = \langle\mathbf{x},\mathbf{x}'\rangle. \tag{2.58}$$

Hence, the ν-SVR dual program, for $\nu \ge 0$, $C \ge 0$, is given by

$$\max_{\alpha^{(*)}\in\mathbb{R}^n} W(\alpha^{(*)}) = \sum_{i=1}^{n}(\alpha_i^*-\alpha_i)y_i - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)k(\mathbf{x}_i,\mathbf{x}_j), \tag{2.59}$$

$$\text{subject to } \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) = 0, \tag{2.60}$$

$$\alpha_i^{(*)} \in \left[0, \frac{C}{n}\right], \tag{2.61}$$

$$\sum_{i=1}^{n} (\alpha_i + \alpha_i^*) \leq C \cdot \nu. \tag{2.62}$$

The regression estimate then takes the form (Schölkopf et al. [76]):

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) k(\mathbf{x}_i, \mathbf{x}) + b, \tag{2.63}$$

where b and ϵ are again computed using the complementary slackness conditions.

Now, if $\nu > 1$, then necessarily, $\epsilon = 0$, since it does not pay to increase ϵ. Hence, (2.62) is redundant, and all values of $\nu \geq 1$ are in fact, equivalent. So, we are only interested in the case when $0 \leq \nu \leq 1$ (Schölkopf et al. [76]).

Inventory Control using SV Regression

The general methodology for inventory control using SV Regression is as follows. As in the case of linear correction of estimates, we define an interval Ω, in which our unknown shape parameter k lies. Moreover, we assume our demand distribution has mean 1. This defines a corresponding interval for the scale parameter θ. We have no prior information about the distribution of k over the interval Ω. Hence, we pick r points uniformly from Ω and form a vector k_a. This defines a corresponding vector θ_a, of length r, as required. Then, we determine the corresponding optimal order quantity vector x_a; this becomes our benchmark. Next, for each value in k_a and/or θ_a, we generate a sample of size n and set corresponding parameter estimates as inputs. We train the SV model by fixing the response for each input equal to the optimal order quantity corresponding to the true value of k. The radial basis function ($\exp\{-\gamma\|\mathbf{x} - \mathbf{x}'\|^2\}$) is used as a kernel. We illustrate this methodology in the following for a Gamma distributed demand.

Example : Gamma distribution

ν-SV and ϵ-SV Regression were applied to each of the following three cases considered earlier:

- Known mean, unknown shape parameter
- Unknown mean, known shape parameter
- Unknown mean, unknown shape parameter

We picked $r = 30$ points uniformly from the likely range of k values and formed a vector k_a of length $r = 30$. For each point picked, for each case mentioned above, $j = 50$ samples were generated. For each such sample the parameter estimates were set as inputs. The SV model was trained by fixing the response for each input equal to the optimal order quantity corresponding to the true value of k. Hence, a total of 1500 training points were used. The radial basis function with parameter γ was used as a kernel.

Inventory Control using ν-SV Regression Three parameters, namely γ, the cost C and ν needed to be set. These needed to be tuned optimally for the given scenario, i.e. $n = 20$ and $s/c = 1.2$. A rough tuning was performed by performing grid search over a range of values for each parameter, around their default values. For instance, the default values for the three parameters were $\gamma = 1/\text{data dimension}$ (which was either 1 or 0.5 in our case), $C = 1$ and $\nu = 0.5$. Hence the range of consideration for each parameter was $\gamma = (0.125, 0.25, 0.5, 1, 2, 4)$, $C = 0.01, 0.1, 1, 10, 100, 1000$ and $\nu = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$. Hence, grid search was performed over all 324 combinations of the parameters to determine the optimal set of parameters for each case, using 10-fold cross validation. The optimal parameters determined a model which was then used to determine the regression function. This approach was then compared with the linear correction approaches considered earlier in terms of expected relative regret. As before, a new sample of size $n = 20$ was generated, the unknown parameters were estimated, and the correction factors or the regression function were applied to determine the relative regret functions. Repeating this procedure $rep = 1000$ times, the approximations to the expected relative regret were determined. The ν-SV regression was implemented in R using the e1071 package.

Known mean, unknown shape parameter In this case, the training inputs were the MOM estimates of k for each sample and the training responses were the optimal order quantities corresponding to the true values of k. The optimal parameters determined by tuning were $\gamma = 0.25$, $C = 1000$ and $\nu = 0.6$. The graph of the performance of ν-SV regression against the linear correction procedures is shown in Figure 2.7.

Unknown mean, known shape parameter In this case, the training inputs were the MOM estimates of θ for each sample and the training responses were the optimal order quantities corresponding to the true values of k and θ. The optimal parameters determined by tuning were $\gamma = 1$, $C = 10$ and $\nu = 0.8$. The graph of the performance of ν-SV regression against the linear correction procedures is shown in Figure 2.8.

Unknown mean, unknown shape parameter In this case, the training inputs were the MOM estimates of k and θ for each sample and the training responses were the optimal order quantities corresponding to the true values of k and θ. The optimal parameters determined by tuning were $\gamma = 0.125$, $C = 100$ and $\nu = 0.7$. The graph of the performance of ν-SV regression against the linear correction procedures is shown in Figure 2.9.

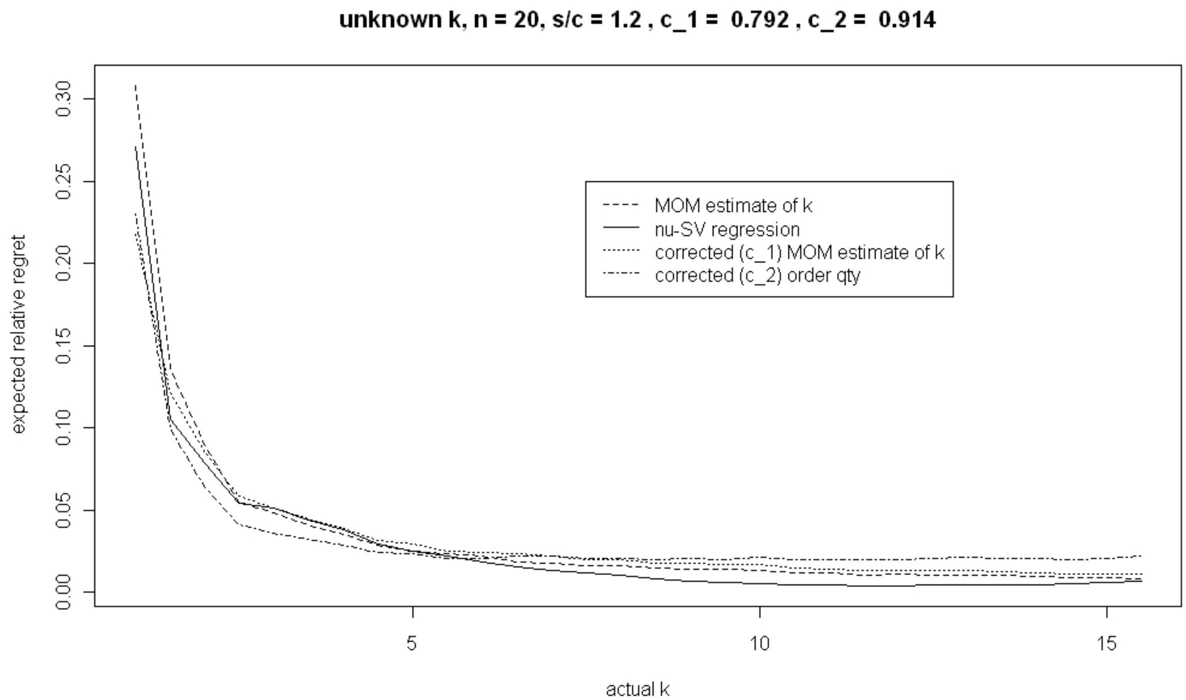

Figure 2.7: Known mean, unknown shape parameter

Inventory Control using ϵ-SV Regression In the case of ϵ-SV Regression, the value of the ϵ parameter was set to be 0.0001 (arbitrarily). Furthermore, two parameters, namely γ, and the cost C needed to be set. These needed to be tuned optimally for the given scenario, i.e. $n = 20$ and $s/c = 1.2$. A rough tuning was performed by performing grid search over a range of values for each parameter, around their default values. For instance, the default values for the three parameters were $\gamma = 1/\text{data dimension}$ (which was either 1 or 0.5 in our case), and $C = 1$. Hence the range of consideration for each parameter was $\gamma = (0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8)$, and $C = 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10^4, 10^5$. Hence, grid search was performed over all combinations of the parameters to determine the optimal set of parameters for each case, using 10-fold cross validation. The optimal parameters determined a model which was then used to determine the regression function. This approach was then compared with the linear correction approaches considered earlier in terms of expected relative regret. As before, a new sample of size $n = 20$ was generated, the unknown parameters were estimated, and the correction factors or the regression function were applied to determine the relative regret functions. Repeating this procedure $rep = 1000$ times, the approximations to the expected relative regret were determined. The ϵ-SV regression was implemented in R using the e1071 package.

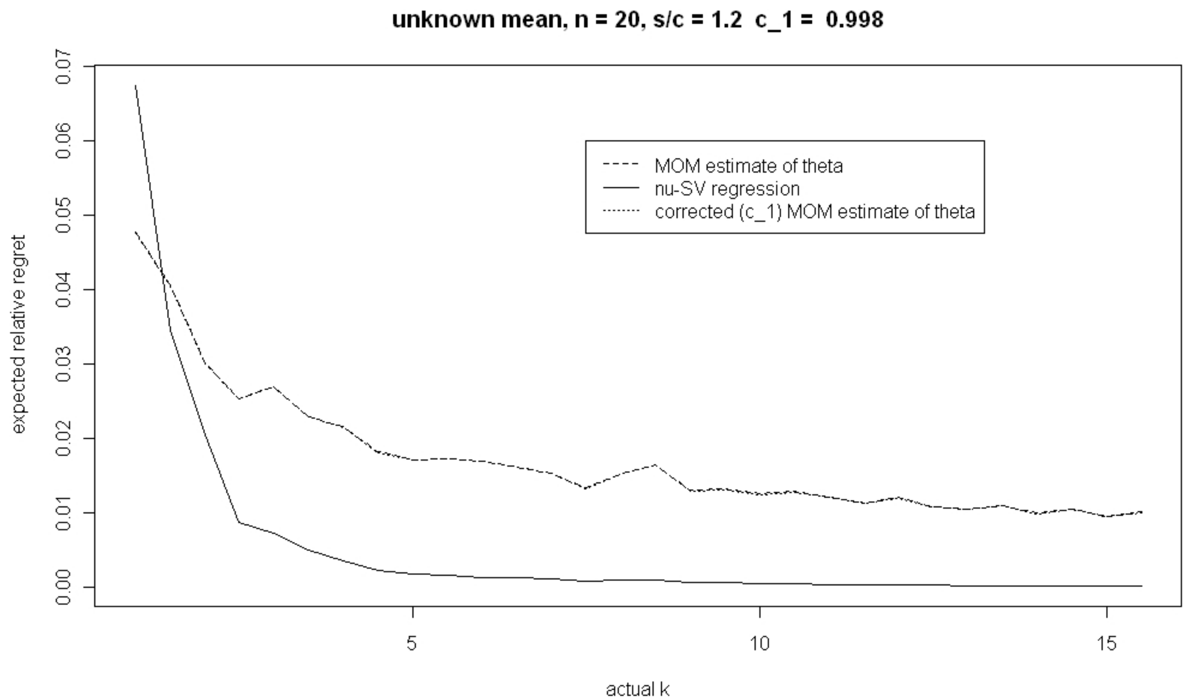

Figure 2.8: Unknown mean, known shape parameter

Known mean, unknown shape parameter In this case, the training inputs were the MOM estimates of k for each sample and the training responses were the optimal order quantities corresponding to the true values of k. The optimal parameters determined by tuning were $\gamma = 0.125$ and $C = 10000$. The performance of ϵ-SV regression against the linear correction procedures was found to be similar to that of ν-SV regression.

Unknown mean, known shape parameter In this case, the training inputs were the MOM estimates of θ for each sample and the training responses were the optimal order quantities corresponding to the true values of k and θ. The optimal parameters determined by tuning were $\gamma = 0.0625$ and $C = 100$. The performance of ϵ-SV regression against the linear correction procedures was again found to be similar to that of ν-SV regression.

Unknown mean, unknown shape parameter In this case, the training inputs were the MOM estimates of k and θ for each sample and the training responses were the optimal order quantities corresponding to the true values of k and θ. The optimal parameters determined by tuning were $\gamma = 0.0625$ and $C = 10^5$. The performance of ϵ-SV regression against the linear correction procedures, was again found to be similar to that of ν-SV regression.

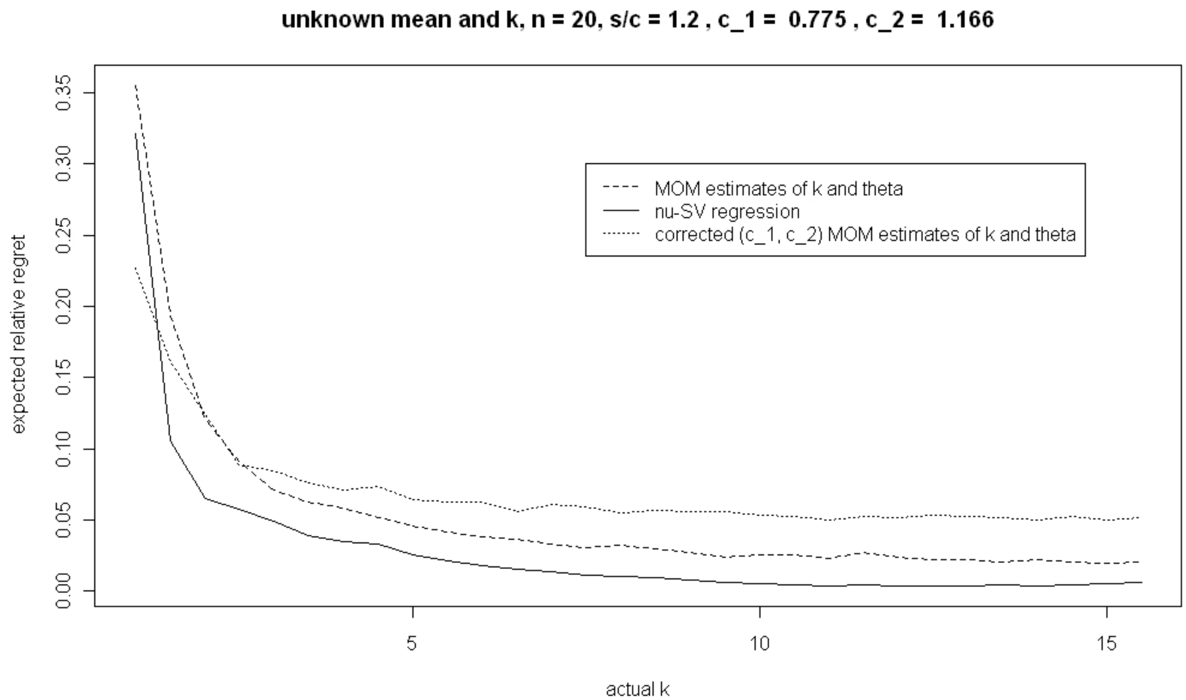

Figure 2.9: unknown mean, unknown shape parameter

2.4 Discussion of Results and Managerial Implications

In this chapter, we have considered several different approaches to improved inventory control when the parameters of the demand distribution, in particular the shape parameter, are uncertain or unknown. At the outset, we considered a heuristic based on operational statistics in cases where the optimal order quantity was 'separable'. In this case, while the benefits were admittedly negligible, we did manage to find policies which dominated traditional plug-in policies. The exercise also directed us towards approaches which would be more generally applicable, which we then considered in the 'non-separable' case.

As far as linear correction of estimates was concerned, we observed that the knowledge of the shape parameter rendered the approach quite redundant. In this case, the performance of this approach was no better than a traditional approach that estimated parameters and plugged them in. In the more interesting cases where at the very least, the shape parameter was unknown, we found that linear correction approach was more useful when there was uncertainty about, rather than complete lack of knowledge of the shape parameter. In particular, for the case of the Gamma distribution, the linear correction approach provided improvement over traditional approaches for low values of k. An interesting line of examination was the variation of performance as the profit margin s/c varied. In this case, we observed that the expected relative regret increased, as s/c decreased towards 1. Moreover, there was dramatic improvement in performance of linear correction over the traditional approach as s/c decreased towards 1. This was in contrast to the separable case, where the performance of the heuristic policies improved as s/c increased.

In the case of support vector regression, we observed that there wasn't a lot of improvement when only the shape parameter was unknown. However, in all other cases, when at the very least, the scale parameter was unknown, support vector regression offered significant improvement over almost the entire range of values considered for the shape parameter. This was an improvement over the linear correction approach, which really only offered improved performance for low values of k. Furthermore, as in the case of linear correction, expected relative regret increased, but there was dramatic improvement offered by support vector regression, as s/c decreased towards 1.

Thus, the immediate managerial implication of this work is as follows. When the shape parameter is known to have a low value (at least in the case of the Gamma distribution), but its value is uncertain, and the profit margin is very low, there is definite benefit in applying the linear correction approach, when past data is limited. In the cases, where at the very least the scale parameter is unknown (and all other parameters are also most probably unknown), support vector regression offers significantly improved performance when the past data is limited. The performance improvement is even more stark when the profit margin is low. While our approaches are not guaranteed to work for all classes of distributions, the illustrative examples do suggest potential utility of our approaches to a wider class of distrbutions, which is worthy of further investigation.

Finally, we note that the improvement due to support vector regression and linear correction, in particular, is quite sensitive to the values of k. While we do not have an explanation for this, we believe this has to do with certain aspects of the Gamma distribution which was considered. In particular, the negligible variance, and consequent degeneracy, of the mean 1 Gamma distribution for larger values of k is the likely cause for sharp drop in performance improvement for large values of k. This is yet another avenue that deserves further examination that could be potentially rewarding.

Chapter 3

Call Center Staffing

3.1 Introduction

Telephone call centers are an integral part of several businesses and their economic role has grown significantly over the last decade or so. In most call centers, capacity costs in general account for 60%−70% of operating expenses, which makes capacity management critical from a cost perspective (see Gans et al. [32]). In this chapter, we consider inbound telephone call centers which handle service requests that originate from customers calling in. As described in Gans et al. [32], such call centers use a hierarchical staffing and scheduling system. In the classical approach to staffing, (see, for example, Gans et al. [32], Mok and Shanthikumar [65]), the call center is modeled as a time varying, Poisson arrival, multiple server, queueing system, with abandonment. The process begins with forecasts of the arrival process over a planning horizon, which may range from a day to several weeks. The distributions of the service time and the time to abandonment would also be estimated. These forecasts and estimates would then be used to analyze the queueing model to obtain an approximate performance measure and cost under different staffing levels. These performance measures would then be optimized to determine staffing levels over the short time intervals, and in turn, constraints to be met, as the call center develops staff schedules. In this manner, the forecasted arrival process of calls drives employee schedules at the call center.

The underlying assumption in the above approach is that the arrival rate forecasts are correct. This is generally not a valid assumption, and leads the system performance to deviate from the expected behaviour. Recent work in statistics and operations research has begun to address the problem of how call centers and other high volume service businesses can better manage the capacity-demand mismatch that results from arrival-rate uncertainty. The statistical research has mainly sought to improve the forecasting of arrival rates. Avramidis et al. [10] developed stochastic models of time-dependent arrivals, with focus on the application to call centers. Brown et al. [20] performed statistical analysis of a unique record of call center operations from a queueing science perspective. Weinberg et al. [86] proposed a multiplicative model based on Markov Chain Monte Carlo for modeling and forecasting within-day arrival rates to a U.S. commercial bank's call center. Shen and Huang [77] used singular value decomposition and time series techniques to develop methods for interday and dynamic intraday forecasting of incoming call volumes.

Papers in operations research have tended to use the stochastic programming framework to account for arrival-rate uncertainty when making short-run staffing and scheduling decisions. Harrison and Zeevi [42] used stochastic fluid models to reduce the staffing problem to a multidimensional newsvendor problem, which can be solved numerically by a combination of linear programming and Monte Carlo simulation. Bassamboo and Zeevi [12] proposed a data-driven method for optimal staffing in large call centers, and performed some asymptotic analysis on it. Bertsimas and Doan [13] proposed both robust and data-driven approaches to a fluid model of call centers that incorporated random arrival rates with abandonment to determine staff level and dynamic routing policies. Mehrotra et al. [64] used mid-day recourse actions to adjust pre-scheduled staffing levels in reaction to realized deviations from arrival-rate forecasts.

While each of these streams of research has made important progress in addressing elements of the problems caused by arrival-rate uncertainty, none addresses the whole problem. Statistical papers dedicated to forecasting have used standard statistical measures of fit to assess performance. They have not, however, considered the downstream cost and quality of service implications of arrival-rate forecast errors. In contrast, operations management papers have looked carefully at the cost implications of stochastic scheduling methods, but they have not used the sophisticated statistical forecasting methods that best capture the nature and dynamics of arrival-rate uncertainty. In turn, their measures of cost improvements may not accurately reflect the gains that can be made when better forecasting and scheduling methods are used in concert. Hence, there is clearly a need for an integrated, data-driven approach that marries the best aspects of the two streams of research. One such approach is described by Gans et al. [33], who developed and tested an integrated forecasting and stochastic programming approach to workforce management in call centers.

In this chapter, we propose an objective operational learning approach to optimal staffing in a call center. Our primary goal in this approach is making minimal assumptions about the distributions of call arrivals, customer waiting times and service times, and using empirical estimates wherever possible. In this sense, our approach is 'objective'. We focus on an 'operational' quantity of interest, which is the cost function, and try to estimate it for various staffing levels. In the long run, as more data are available, we aim to eliminate any errors introduced by using empirical estimates of parameters, and 'learn' the true cost, as a function of the call profile feature forecast and staffing level. Hence, we call our approach 'objective operational learning'. We consider the following situation. A call center manager has data on the call arrivals for the past n days. It may be reasonably assumed that this data includes detailed information on every call. In particular, it includes the arrival times, waiting times, and service times for each call, as well as detailed staffing level information by shift. Suppose that the manager uses this data to come up with the best possible forecast for the call profile on the next day. He now wishes to determine the optimal staffing level for this call profile on the next day. Our approach to solving this problem involves using the past data to learn the total cost (staffing and abandonment) objective, as a function of the call profile forecast and the staffing level. Having estimated this objective function, we then seek to optimize it with respect to the staffing level.

In objective operational learning (which can be seen as an operational extension of non-parametric statistics) we do not exclusively use the cost estimates obtained with the forecast estimates and the analytic (approximate) queueing results. Rather we use the observed cost, and an operational extension of it, to find the optimal staffing level. Specifically we use the estimated distributions to complete the data on service times (for those callers who abandoned on a particular day) and the time to abandon (for those callers who got served on a particular day). This then provides us with a complete data set to compute the sample performance for a particular day. We call this the probabilistically "extended data set". Using the extended data set, for each day, we construct the sample path for (i.e., simulate) the multiple server queueing system and compute the cost for various staffing levels. Observe that if the dynamics of the call center conform to the "modeled", multiple server queueing system, then the cost computed this way for the actual staffing level will be the cost observed that day. This may not be the case in practice (for example, see Mok and Shanthikumar [65]). In practice, the staff may not be available to serve a waiting call immediately after completing the service to another call (depending on the number of calls waiting) and supervisors may take calls when excessive waiting is seen. This is usually done in an ad-hoc manner, making it difficult to model the decisions by a set of well defined rules. In such cases, an operationally "adjusted" queueing model can be used to match the cost computed using the model and the actual cost. In such a case we assume that the cost function for any day is computed using this "adjusted" queueing model. The objective function used to find the staffing level is then obtained by a kernel smoothing of the cost functions computed with the data available for the n days.

The remainder of the chapter is organized as follows. Section 2 reviews additional relevant literature. Section 3 provides a mathematical description of our objective operational learning model. Section 4 describes our numerical experiments and Section 5 provides a discussion of the results obtained.

3.2 Literature Review

The introduction cites papers that are closest to our work in spirit. In addition to Gans et al. [32], another comprehensive review article on call center operations is by Akşin et al. [2]. Most of the early literature on call center staffing focused on a single pool of identical servers. In that realm, the case where there is only a single class of customers leads to trivial control decisions, and if the system is Markovian, the Erlang-C formula provides the main mathematical tool for solving the staffing problem. An important rule-of-thumb that arises from the Erlang-C formula is the so-called *square-root staffing rule*; see Gans et al. [32] for further discussion. Borst et al. [17] refine the square-root safety staffing principle by trading off agents' costs with service quality. Garnett et al. [34] were among the first to account for abandonments, and analyzed the simplest abandonment model, in which customers' patience is exponentially distributed and the system's waiting capacity is unlimited. More recently, Mandelbaum and Zeltyn [62] studied asymptotically optimal staffing of many-server queues with abandonment. Their call center is characterized by Poisson arrivals, exponential service times, n servers, and generally distributed patience times of customers.

Staffing a single pool of servers when there are multiple customer classes involves a significant escalation in complexity. Research on this problem has started only recently, and the primary example of such work is that of Gurvich et al. [40], which exploits many server diffusion limits in the so-called quality- and efficiency-driven regime (QED) first introduced by Halfin and Whitt [41]. Work on the staffing problem in the context of a multiclass/multipool model is still in its infancy and relies mostly on simulation-based methods; for an example of the latter see Wallace and Whitt [85].

In other work on optimal staffing in call centers, Armony et al. [6] studied the sensitivity of optimal capacity to customer impatience in an unobservable M/M/S Queue. They employed sample path arguments to derive several convexity properties and comparative statics for an M/M/S queue with impatient customers. Gurvich et al. [38] considered the problem of staffing call-centers with multiple customer classes and agent types operating under quality-of-service (QoS) constraints and demand rate uncertainty. They introduced a formulation of the staffing problem that required that the QoS constraints are met with high probability with respect to the uncertainty in the demand rate. They then proposed a two-step solution for the staffing problem under chance constraints. Their formulation and solution approach had the important property that it translates the problem with uncertain demand rates to one with known arrival rates. Cross-selling is becoming an increasingly prevalent practice in call centers, due, in part, to its unique capability to allow firms to dynamically segment their callers and customize their product offerings accordingly. Two recent papers reflect the growing interest in this phenomenon. Gurvich et al. [39] considered a call center with cross-selling capability that served a pool of customers that were differentiated in terms of their revenue potential and delay sensitivity. They studied the operational decisions of staffing, call routing, and cross-selling under various forms of customer segmentation, and derived near-optimal controls in each of the settings analyzed. Armony and Gurvich [5] addressed the following two questions: How many customer-service representatives are required (staffing), and when should cross-selling opportunities be exercised (control) in a way that will maximize the expected profit of the center while maintaining a prespecified service level target? They tackled these questions by characterizing control and staffing schemes that are asymptotically optimal in the limit, as the system load grows large. Their main finding was that a threshold priority control, in which cross-selling was exercised only if the number of callers in the system was below a certain threshold, was asymptotically optimal in great generality. The asymptotic optimality of threshold priority reduced the staffing problem to a solution of a simple deterministic problem in one regime and to a simple search procedure in another. They showed that their joint staffing and control scheme was nearly optimal for large systems.

3.3 Objective Operational Learning Model

Model Setup

We suppose that the call center has J disjoint staffing intervals of interest during any day, over T time units. We assume each such staffing interval is of the same length (say one hour). As our variable of optimization, we define a staffing vector $\mathbf{z} = (z_1, \cdots, z_J)$ where $z_j \in \mathbb{Z}_+$ for all $j = 1, \cdots, J$. For any j, z_j denotes the number of staffing agents that work during the staffing interval j. We now assume we are given data for the previous n days in the following form (here, i is a day index, and k is a call index):
$\{a_{ik}\}$:- ordered sequence of call arrival times
$\{q_{ik}\}$:- corresponding sequence of call queue times
$\{s_{ik}\}$:- corresponding sequence of call service times
$\{r_{ik}\}$:- corresponding binary sequence of call outcomes; 1 if served, and 0 if hung up
A :- total number of agents available to the call center

Hence, a_{ik} is the arrival time (or epoch) of the kth call on day i. q_{ik} and s_{ik} are similarly defined. By default, if the kth call on the ith day is served immediately upon arrival, then $q_{ik} = 0$, and if it is hung up without service, then $s_{ik} = 0$. In addition, for each served call, we know which particular agent served the call. Also, we have the actual staffing level for any day, over every staffing interval of interest.

Now, both times to abandonment and times to service are censored data. We denote by
R :- the "patience" or "time willing to wait before abandonment" and
V :- the "virtual waiting time".

We equip R and V with steady-state distributions. The virtual waiting time amounts to the time that a (virtual) customer, equipped with infinite patience, would have waited until being served. One actually samples $Q = \min\{R, V\}$, as well as the indicator $1\{R < V\}$, for observing R or V. One considers all calls that reached an agent as censored observations for estimating the distribution of R, and vice versa for estimating the distribution of V. We make the assumption that (as random variables) R and V are independent given the covariates relevant to the individual customer. Under this assumption, the distributions of R and V (given the covariates) can be estimated using the standard Kaplan-Meier product-limit estimator. In computing the empirical distribution of R, we use the quantities $q_{ik}1(q_{ik} > 0)$ for all i and k. Moreover, we estimate the empirical distribution of the service time S using the quantities $s_{ik}1(s_{ik} > 0)$ for all i and k.

Cost Estimation

We would now like to construct estimates of the cost function for all n days for which we have data, for a set of feasible staffing vectors. In the following, we assume that we have detailed call information for each of the n days. We also assume that, the call arrival times, and the call profiles, are fixed and known. In subsequent sections, we will use call profile forecast information and integrate it with the costs estimated in this section.

For $i = 1, \cdots, n$, $j = 1, \cdots, J$ and $k = 1, \cdots, A$, we define a binary variable

$$\begin{aligned} t_{ijk} &= 1 \quad \text{if agent } k \text{ works during the } j\text{th staffing interval of day } i \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{3.1}$$

For $i = 1, \cdots, n$, we denote by
$\mathbf{y}_i = (y_{i1}, \cdots, y_{iJ})$:- the actual staffing vector observed on day i
$K_i = \{k : a_{ik}\}$:- the ordered sequence of call indices on day i
K_{ij} :- the ordered sequence of call indices on day i, during staffing interval j, for $j = 1, \cdots, J$

It may easily be seen that for all $i = 1, \cdots, n$ and $j = 1, \cdots, J$,

$$y_{ij} = \sum_{m=1}^{A} t_{ijm} \tag{3.2}$$

Furthermore, we denote by
c_S :- the staffing cost per agent per staffing interval
c_A :- the per head abandonment cost

We assume the calls are served on a first-come-first-served (FCFS) basis. Now, observe that, for the given ordered sequence of call arrival times $\{a_{ik}\}$ for every day i, the sequences $\{q_{ik}\}$, $\{s_{ik}\}$ and $\{r_{ik}\}$ are particular realizations that correspond to the actually observed staffing vector $\mathbf{y}_i$. However, for any staffing vector $\mathbf{z} \neq \mathbf{y}_i$, the realized sequences would be random and not known a priori. Hence, for the given ordered sequence of call arrival times $\{a_{ik}\}$ for every day i, we define the sequences of random variables
$\{\tilde{q}_{ik}\}(\mathbf{z})$:- the realized sequence of call queue times corresponding to staffing vector $\mathbf{z}$
$\{\tilde{s}_{ik}\}(\mathbf{z})$:- the realized sequence of call service times corresponding to staffing vector $\mathbf{z}$
$\{\tilde{r}_{ik}\}(\mathbf{z})$:- the realized sequence of call outcomes corresponding to staffing vector $\mathbf{z}$

We note here, that while the sequences defined above are functions of the staffing vector $\mathbf{z}$, we will suppress the $\mathbf{z}$ in the following development to avoid clutter, and the staffing vector $\mathbf{z}$ will be implied by the context. We now assume a set of feasible staffing vectors for each $i = 1, \cdots, n$, given by

$$\mathbf{Z}_i = \{\mathbf{z} : l_{ij} \leq z_j \leq u_{ij}, z_j \in \mathbb{Z}_+ \forall j\} \tag{3.3}$$

where, for $j = 1, \cdots, J$, l_{ij} and u_{ij} are respectively, suitable lower and upper bounds for the staffing level in the staffing interval j. Now, for each $\mathbf{z} \in \mathbf{Z}_i$, if $\mathbf{z} = \mathbf{y}_i$, then the cost of staffing and abandonment is readily available from the data. We denote it by $C_i(\mathbf{z})$, and it is given by

$$C_i(\mathbf{z}) = c_S \cdot \sum_{j=1}^{J} y_{ij} + c_A \cdot \sum_{k=1}^{|K_i|} (1 - r_{ik}) \tag{3.4}$$

If, however, $\mathbf{z} \neq \mathbf{y}_i$, then the cost of staffing and abandonment is not actually observed. We define
$\tilde{C}_i(\mathbf{z})$:- the random cost of staffing and abandonment for staffing vector $\mathbf{z} \neq \mathbf{y}_i$ We can

still compute the staffing cost in this case. We denote the staffing cost corresponding to the staffing vector $\mathbf{z}$ by $C_i^W(\mathbf{z})$, and it is given by

$$C_i^W(\mathbf{z}) = c_S \cdot \sum_{j=1}^{J} z_j \tag{3.5}$$

Next, we denote by $\tilde{C}_i^H(\mathbf{z})$, the random abandonment cost realized due to the staffing vector $\mathbf{z}$. Clearly, we have, for $\mathbf{z} \neq \mathbf{y}_i$,

$$\tilde{C}_i(\mathbf{z}) = C_i^W(\mathbf{z}) + \tilde{C}_i^H(\mathbf{z}) \tag{3.6}$$

$$\Rightarrow \mathbb{E}[\tilde{C}_i(\mathbf{z})] = C_i^W(\mathbf{z}) + \mathbb{E}[\tilde{C}_i^H(\mathbf{z})] \tag{3.7}$$

Thus, we need a way to estimate the expected abandonment cost $\mathbb{E}[\tilde{C}_i^H(\mathbf{z})]$ in order to estimate the expected total cost $\mathbb{E}[\tilde{C}_i(\mathbf{z})]$.

Cost Estimation Algorithm

We are interested in the expected abandonment cost given by $\mathbb{E}[\tilde{C}_i^H(\mathbf{z})]$, and we estimate it as follows, by generating a large number of queueing system sample paths (L). For each $j = 1, \cdots, J$, we define the set

$$B_{ij} = \{m : t_{ijm} = 1\} \tag{3.8}$$

This set defines the collection of agents who were actually working during the staffing interval j of day i. Now, for $l = 1, \cdots, L$, where l is the sample path number, we do the following. If $z_j \leq y_{ij}$, we randomly pick

$$D_{ij}^{(l)} \subset B_{ij} \text{ subject to } |D_{ij}^{(l)}| = z_j \tag{3.9}$$

This just means $D_{ij}^{(l)}$ is a subset of the collection of the agents who were actually working during the staffing interval j of day i. On the other hand, if $z_j > y_{ij}$, we randomly pick

$$D_{ij}^{(l)} \subset \{1, 2, \cdots, A\} \text{ subject to } B_{ij} \subset D_{ij}^{(l)} \text{ and } |D_{ij}^{(l)}| = z_j \tag{3.10}$$

This means $D_{ij}^{(l)}$ is a subset of all available agents that includes the collection of the agents who were actually working during the staffing interval j of day i.

We now define

$\mathbf{e} \in \mathbb{R}^A$:- the vector of earliest epochs of availability of all A agents

$\mathbf{e}(m)$:- the mth element of $\mathbf{e}$, or the earliest epoch of availability of the mth agent

$\mathbf{e}^{(l,k)}$:- the updated version of $\mathbf{e}$ in the lth sample path, after the kth call, where $k = 1, \cdots, |K_i|$

All sample paths are assumed to be independent of each other and are generated as such. At the outset, we assume

$$\mathbf{e}^{(l,0)}(m) = 0 \text{ for all } m \tag{3.11}$$

Then for $j = 1, \cdots, J$, we let $h = \min K_{ij}$, and we proceed as follows, for $k = h, h+1, \cdots, h+|K_{ij}|-1$ (i.e., for all call indices during staffing interval j of day i),

$$\mathbf{e}^{(l,k-1)}(m) = jT/J \text{ for } m \notin D_{ij}^{(l)} \tag{3.12}$$

Equation (3.12) sets the earliest epoch of availability to the end of the staffing interval, for all agents not picked to work during the staffing interval j for sample path l.

$$f^{(k)} = \arg\min_m \mathbf{e}^{(l,k-1)}(m) \tag{3.13}$$

Equation (3.13) picks the agent with the earliest epoch of availability, amongst all agents working during interval j in sample path l, to attend to the kth call. Now, depending on when the earliest epoch of availability occurs relative to the actual duration of the kth call, we have the following cases described in the equations below.

$$(\mathbf{e}^{(l,k-1)}(f^{(k)}) \leq a_{ik} + q_{ik}, r_{ik} = 1) \Rightarrow \mathbf{e}^{(l,k)}(f^{(k)}) = a_{ik} + \max\{\mathbf{e}^{(l,k-1)}(f^{(k)}) - a_{ik}, 0\} + s_{ik}, \tag{3.14}$$

Equation (3.14) corresponds to the case in which the agent becomes available for a call that was actually served, before it was served. In this case, the updated epoch of the serving agent is given by adding the actual service time s_{ik} of the call to the earliest epoch of availability or at the time the call was actually served, which ever came later.

$$(\mathbf{e}^{(l,k-1)}(f^{(k)}) \leq a_{ik}+q_{ik}, r_{ik} = 0) \Rightarrow \mathbf{e}^{(l,k)}(f^{(k)}) = a_{ik}+\max\{\mathbf{e}^{(l,k-1)}(f^{(k)})-a_{ik}, 0\}+\tilde{s}_{ik}^{(l)} \tag{3.15}$$

where $\tilde{s}_{ik}^{(l)} \sim S$. Equation (3.15) corresponds to the case in which the agent becomes available for a call that was actually hung up, before it was hung up. In this case, the agent is in a position to serve a call, for which there is no actual service time. The service time in this case ($\tilde{s}_{ik}^{(l)}$) is hence sampled from the empirical estimated distribution of the service time S. This sampled service time is then used to obtain the updated epoch of the agent, as in Equation (3.14). Next, for $\tilde{q}_{ik}^{(l)} \sim R|\tilde{q}_{ik}^{(l)} > q_{ik}$,

$$(\mathbf{e}^{(l,k-1)}(f^{(k)}) > a_{ik} + q_{ik}, \mathbf{e}^{(l,k-1)}(f^{(k)}) \leq a_{ik} + \tilde{q}_{ik}^{(l)}, r_{ik} = 1) \Rightarrow \mathbf{e}^{(l,k)}(f^{(k)}) = \mathbf{e}^{(l,k-1)}(f^{(k)}) + s_{ik} \tag{3.16}$$

Equation (3.16) corresponds to the case in which the agent becomes available for a call that was actually served, after its service was actually begun. In this case, we need to sample a new queue time for the call $\tilde{q}_{ik}^{(l)}$ from the empirical distribution of R, such that the new queue time is greater than the actual queue time q_{ik}. In this case, if the new queue time is such that the call is still waiting to be served (i.e. not hung up) by the time the agent is available, the updated epoch of the agent is given by adding the actual service time s_{ik} to the earliest epoch of availability. In all other cases, the call has been hung up before the agent is available, and the updated epoch of the agent remains the same as the earliest epoch of availability, as stated in Equation (3.17).

$$\mathbf{e}^{(l,k)}(f^{(k)}) = \mathbf{e}^{(l,k-1)}(f^{(k)}) \tag{3.17}$$

It then follows that,

$$\begin{aligned} \tilde{r}_{ik}^{(l)}(\mathbf{z}) &= 0 \quad \text{if } \mathbf{e}^{(l,k)}(f^{(k)}) = \mathbf{e}^{(l,k-1)}(f^{(k)}) \\ &= 1 \quad \text{otherwise} \end{aligned} \tag{3.18}$$

The abandonment cost for the sample path l is now given by

$$\tilde{C}_i^{H(l)}(\mathbf{z}) = c_A \cdot \sum_{k=1}^{|K_i|} (1 - \tilde{r}_{ik}^{(l)}(\mathbf{z})) \tag{3.19}$$

Now, by the weak law of large numbers, we have, as $L \to \infty$,

$$\frac{1}{L} \sum_{l=1}^{L} \tilde{C}_i^{H(l)}(\mathbf{z}) \xrightarrow{P} \mathbb{E}_{\mathbf{z}}[\tilde{C}_i^H(\mathbf{z})] \tag{3.20}$$

Hence,

$$\frac{1}{L} \sum_{l=1}^{L} \tilde{C}_i^{H(l)}(\mathbf{z}) \text{ is a consistent estimator of } \mathbb{E}[\tilde{C}_i^H(\mathbf{z})] \tag{3.21}$$

Call Profile Forecasting

Following this, we want to obtain a one day-ahead forecast of the call profile using the detailed call information we have available for the n days in our data set. For this we use the approach described in Shen and Huang [77]. The approach is as follows:

Let $\mathbf{X} = (x_{ij})$ be an $n \times p$ matrix that records the call volumes for n days, with each day having p time periods. Note that these time periods need not be the same as the staffing intervals described previously. The rows and columns of $\mathbf{X}$ correspond respectively to days and time periods within a day. The ith row of $\mathbf{X}$, denoted as $\mathbf{x}_i^T = (x_{i1}, \cdots, x_{ip})$, is referred to as the intraday call volume profile of the ith day. The intraday profiles, $\mathbf{x}_1, \mathbf{x}_2, \cdots$, form a vector-valued time series (TS) taking values in $\mathbb{R}^p$. We want to build a TS model for this series and use it for forecasting. However, commonly used multivariate TS models such as vector autoregressive models and more general vector autoregressive and moving average models (Reinsel [67]) are not directly applicable because of the large dimensionality of the TS we consider.

The approach starts from dimension reduction. A few basis vectors are sought, denoted as $\mathbf{f}_k, k = 1, \cdots, K$, such that all elements in the TS $\{\mathbf{x}_i\}$ can be represented (or approximated well) by these basis vectors. The number of the basis vectors K should be much smaller than the dimensionality m of the TS. Specifically, the following decomposition is considered (see Shen and Huang [77]),

$$\mathbf{x}_i = \beta_{i1}\mathbf{f}_1 + \cdots + \beta_{iK}\mathbf{f}_K + \epsilon_i, \ i = 1, \cdots, n \tag{3.22}$$

where $\mathbf{f}_1, \cdots, \mathbf{f}_K \in \mathbb{R}^m$ are the basis vectors and $\epsilon_1, \cdots, \epsilon_n \in \mathbb{R}^m$ are the error terms. It is expected that the main features of $\mathbf{x}_i$ can be summarized by a linear combination of the basis vectors so that the error terms in (3.22) would be small in magnitude. This can be achieved by solving the following minimization problem for fixed K (see Shen and Huang [77]):

$$\min_{\beta_{i1},\cdots,\beta_{iK},\mathbf{f}_1,\cdots,\mathbf{f}_K} \sum_{i=1}^{n} \|\epsilon_i\|^2 = \min_{\beta_{i1},\cdots,\beta_{iK},\mathbf{f}_1,\cdots,\mathbf{f}_K} \sum_{i=1}^{n} \|\mathbf{x}_i - (\beta_{i1}\mathbf{f}_1 + \cdots + \beta_{iK}\mathbf{f}_K)\|^2 \tag{3.23}$$

For identifiability, it is required in (3.23) that $\mathbf{f}_i^T\mathbf{f}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta, which equals 1 for $i = j$ and 0 otherwise. The solution to this problem is actually given by the singular value decomposition (SVD) of the matrix X as shown below (see Shen and Huang [77]).

The SVD of the matrix X can be expressed as

$$X = USV^T \tag{3.24}$$

where $U = (u_{ij})$ is an $n \times p$ matrix with orthonormal columns, S is an $p \times p$ diagonal matrix, and V is an $p \times p$ orthogonal matrix. The diagonal elements of S are the singular values, which are usually ordered decreasingly. Let $S = \text{diag}(s_1, \cdots, s_p)$ and $r = \text{rank}(X)$. It then follows from (3.24) that

$$\mathbf{x}_i = s_1 u_{i1}\mathbf{v}_1 + \cdots + s_r u_{ir}\mathbf{v}_r$$

where $\mathbf{v}_1, \cdots, \mathbf{v}_r$ are columns of V. Keeping only the terms associated with the largest K singular values, we have the following approximation:

$$\mathbf{x}_i \simeq s_1 u_{i1}\mathbf{v}_1 + \cdots + s_K u_{iK}\mathbf{v}_K$$

This K-term approximation is an optimal solution for the minimization problem (3.23) (see Shen and Huang [77]). More precisely, $\beta_{ik} = s_k u_{ik}$ and $\mathbf{f}_k = \mathbf{v}_k$, $i = 1, \cdots, n, k = 1, \cdots, K$, solve (3.23), and the solution is unique up to a sign change to $\mathbf{f}_k$ (Eckart and Young [30]). Thus, the decomposition (3.22) is formally obtained using the SVD of $\mathbf{X}$.

Now, to estimate the model (3.22), SVD is applied to X and the first K pairs of singular vectors are extracted, along with the corresponding singular values, which then lead to the intraday feature vectors $\mathbf{f}_1, \cdots, \mathbf{f}_K$, and the interday feature series $\{\beta_{i1}\}, \cdots, \{\beta_{iK}\}$. This SVD-based dimension reduction is closely related to principal components analysis (PCA) when principal components (PCs) are calculated from the covariance matrix (Jolliffe [49]). If the data matrix X is column centered such that the columns (viewed as variables) have a mean of zero, then X^TX is proportional to the sample covariance matrix of the columns of X. According to (3.24), $X^TX = VS^2V^T$, which means that the columns $\mathbf{v}_k$ of V, or the intraday feature vectors $\mathbf{f}_k$, are indeed the PC-loading vectors, or $\mathbf{X}\mathbf{v}_k$ are the PCs; and the squared singular values s_k^2 are proportional to the variances of the PCs. In PCA, the quantity $s_k^2 / \sum_{i=1}^n s_i^2$ measures the relative importance of the kth PC and can be plotted versus k in a scree plot. To determine the number of PCs, one usually looks for an "elbow" in the scree plot, formed by a steep drop followed by a relatively flat tail of small values. The number of PCs needed then corresponds to the integer prior to the elbow. In this way, K is determined (see Shen and Huang [77]).

Thus, by making use of the model (3.22), forecasting an m-dimensional TS $\{\mathbf{x}_i\}$ reduces to forecasting K one-dimensional interday feature series $\{\beta_{i1}\}, \cdots, \{\beta_{iK}\}$. Because of the way the SVD is constructed, $(\beta_{1k}, \cdots, \beta_{nk})$ is orthogonal to $(\beta_{1l}, \cdots, \beta_{nl})$ for $k \neq l$. This lack of contemporaneous correlation suggests that the cross-correlations at nonzero lags are likely to be small. Therefore, it is reasonable to believe that forecasting each series separately using univariate TS methods, for $k = 1, \cdots, K$, is adequate (see Shen and Huang [77]).

Kernel Smoothing of Cost Function

Having applied the univariate models to each of the K series, we obtain a one day-ahead forecast for each series. We denote these by $\hat{\beta} = (\hat{\beta}_1, \cdots, \hat{\beta}_K)$. We now recall, from Section 3.1, that if the dynamics of the call center conform to the "modeled", multiple server queueing system, then the cost computed using the empirically generated sample paths for the actual staffing level will be the cost observed that day. This is usually not the case in practice (for example, see Mok and Shanthikumar [65]). In practice, the staff may not be available to serve a waiting call immediately after completing the service to another call (depending on the number of calls waiting) and supervisors may take calls when excessive waiting is seen. These are usually followed in an ad-hoc manner, making it difficult to model them by a set of well defined rules. In such cases, an operationally "adjusted" queueing model needs to be used to match the cost computed using the model and the actual cost. In this case we assume that the cost function for any day is computed using this "adjusted" queueing model.

We adjust the queueing model as follows. We define the clairvoyant cost associated with a particular staffing vector $\mathbf{z}$ to be the actual cost associated with $\mathbf{z}$, when all the call arrival times are known beforehand. For any day i and any staffing vector $\mathbf{z}$, we denote the clairvoyant cost by $C_i^C(\mathbf{z})$. When $\mathbf{z} = \mathbf{y}_i$, it follows that $C_i^C(\mathbf{z}) = C_i(\mathbf{y}_i)$, for all i. Except for the actual staffing vector $\mathbf{y}_i$, for $i = 1, \cdots, n$, the clairvoyant cost associated with any other staffing vector $\mathbf{z} \in \mathbf{Z}_i, \mathbf{z} \neq \mathbf{y}_i$ is at best, a theoretical quantity, as it has not actually been observed. We now suggest a somewhat intuitive heuristic to estimate the clairvoyant cost associated with a staffing vector that has not actually been observed. For $\mathbf{z} \in \mathbf{Z}_i, \mathbf{z} \neq \mathbf{y}_i$, we denote the estimated clairvoyant cost by $\hat{C}_i^C(\mathbf{z})$. For $i = 1, \cdots, n$, we define a correction factor g_i given by

$$g_i = \frac{c_A \cdot \sum_{k=1}^{|K_i|} (1 - r_{ik})}{\frac{1}{L} \sum_{l=1}^{L} \tilde{C}_i^{H(l)}(\mathbf{y}_i)} \tag{3.25}$$

Now, for each $i = 1, \cdots, n$, and $\mathbf{z} \in \mathbf{Z}_i, \mathbf{z} \neq \mathbf{y}_i$, we define the estimated clairvoyant cost to be

$$\hat{C}_i^C(\mathbf{z}) = C_i^W(\mathbf{z}) + g_i \cdot \frac{1}{L}\sum_{l=1}^{L} \tilde{C}_i^{H(l)}(\mathbf{z}) \tag{3.26}$$

We now define a cost function $\phi(\hat{\beta}, \mathbf{z})$ given by

$$\phi(\hat{\beta}, \mathbf{z}) = \mathbb{E}[C_i^C(\mathbf{z})|\beta_i = \hat{\beta}] \tag{3.27}$$

We want to construct an objective operational estimate for $\phi(\hat{\beta}, \mathbf{z})$, which we denote by

$$\hat{\phi}_n(\hat{\beta}, \mathbf{z})$$

To construct this estimate, we have actually observed data of the form

$$(\beta_i, \mathbf{y}_i, C_i^C(\mathbf{y}_i)) \text{ for } \mathbf{z} = \mathbf{y}_i,$$

and also "extended" (estimated) data of the form

$$(\beta_i, \mathbf{z}, \hat{C}_i^C(\mathbf{z})) \text{ for } \mathbf{z} \neq \mathbf{y}_i.$$

Our aim is to appropriately 'smooth' these two kinds of data in our objective operational estimate of the cost function. Toward this end, we define our smoothed objective operational estimate $\hat{\phi}_n(\hat{\beta}, \mathbf{z})$ as:

$$\hat{\phi}_n(\hat{\beta}, \mathbf{z}) = \frac{\sum_{i=1}^n K_1(\hat{\beta}, \beta_i, h_1) C_i^S(\mathbf{z})}{\sum_{i=1}^n K_1(\hat{\beta}, \beta_i, h_1)} \tag{3.28}$$

where

$$C_i^S(\mathbf{z}) := K_2(\mathbf{z}, \mathbf{y}_i, h_2) C_i^C(\mathbf{y}_i) + (1 - K_2(\mathbf{z}, \mathbf{y}_i, h_2)) \cdot \hat{C}_i^S(\mathbf{z}) \tag{3.29}$$

where

$$\hat{C}_i^S(\mathbf{z}) := \frac{\sum_{\mathbf{w} \in \mathbf{Z}_i, \mathbf{w} \neq \mathbf{y}_i} K_3(\mathbf{z}, \mathbf{w}, h_3) \hat{C}_i^C(\mathbf{w})}{\sum_{\mathbf{w} \in \mathbf{Z}_i, \mathbf{w} \neq \mathbf{y}_i} K_3(\mathbf{z}, \mathbf{w}, h_3)} \tag{3.30}$$

Here, K_1, K_2, and K_3 are appropriate kernels and h_1, h_2, h_3 are corresponding smoothing parameters such that for all i, $K_i : (\mathbf{x}, \mathbf{y}, h) \to [0, 1]$, are decreasing in $\|\mathbf{x} - \mathbf{y}\|$, and have the property that $K_i(\mathbf{x}, \mathbf{x}, \cdot) = 1$ for all $\mathbf{x}$. A Gaussian kernel $K(\mathbf{x}, \mathbf{y}, h) = \exp\{-h\|\mathbf{x} - \mathbf{y}\|^2\}$ is an example of a kernel with the required properties. Furthermore, in the above expression for $\hat{\phi}_n(\hat{\beta}, \mathbf{z})$, the set of indices $i = 1, \cdots, n$ refer to all the previous n days in the data. Having constructed the smoothed objective operational estimate of the cost function, for a given forecast $\hat{\beta}$, the optimal staffing vector $\mathbf{z}^*(\hat{\beta})$, is given by

$$\mathbf{z}^*(\hat{\beta}) = \arg\min_{\mathbf{z}} \hat{\phi}(\hat{\beta}, \mathbf{z}) \tag{3.31}$$

We now consider the asymptotic behaviour of $\hat{\phi}(\hat{\beta}, \mathbf{z})$ with respect to sample size. We make the following assumptions about the call arrival process.

(A1) The sequence $\{\beta_n, \{a_{nk}\}\}$ converges in distribution, i.e, for any continuous, bounded functional $\eta(\cdot,\cdot)$, we have

$$\mathbb{E}[\eta(\{\beta_n, \{a_{nk}\}\})] \rightarrow \mathbb{E}[\eta(\{\beta, \{a_k\}\})] \text{ as } n \rightarrow \infty \tag{3.32}$$

(A2) For each continuous, bounded functional $\eta(\cdot,\cdot)$, we have that

$$\lim_{n\rightarrow\infty} \frac{\sum_{i=1}^{n} \eta(\beta_i, \{a_{ik}\}) 1\{\beta_i = \beta\}}{\sum_{i=1}^{n} 1\{\beta_i = \beta\}} = \mathbb{E}[\eta(\{\beta, \{a_k\}\})] \text{ a.s.} \tag{3.33}$$

An example of a process that satisfies A1 and A2 is a semi-regenerative process (see Çinlar [23] for details). Furthermore, we make the following assumptions as well.

(A3) There exists γ such that $|C_i^S(\mathbf{z})| \leq \gamma < \infty$ for all $\mathbf{z}$.

(A4) $h_1(n) \rightarrow \infty$ and $nh_1(n)^{-K/2} \rightarrow \infty$ as $n \rightarrow \infty$.

(A5) There exist positive numbers r, c_1, c_2 such that for all $h, \mathbf{x}$ and $\mathbf{y}$,

$$c_1 1\{h^{-1/2}\|\mathbf{x}-\mathbf{y}\| \leq r\} \leq K_1(\mathbf{x},\mathbf{y};h) \leq c_2 1\{h^{-1/2}\|\mathbf{x}-\mathbf{y}\| \leq r\} \tag{3.34}$$

(A6) $h_2(n) \rightarrow \infty$ as $n \rightarrow \infty$.

(A7) $\mathbb{E}[|(\hat{C}_i^S(\mathbf{z}) - C_i^C(\mathbf{z}))||\beta_i = \hat{\beta}] \rightarrow 0$ for almost all $(\hat{\beta}, \mathbf{z})$ as $n \rightarrow \infty$.

Theorem 1. *Assume that A1 - A7 hold. Then* $\mathbb{E}|\hat{\phi}_n(\hat{\beta},\mathbf{z}) - \phi(\hat{\beta},\mathbf{z})| \rightarrow 0$ *as* $n \rightarrow \infty$ *for almost all* $(\hat{\beta}, \mathbf{z})$.

Theorem 1 is proved in Appendix A.

3.4 Numerical Experiments

In our numerical experiments, we aim to do a dynamic out-of-sample testing of our proposed algorithm, as well as another approach in the literature, with respect to an optimal clairvoyant cost. For $\mathbf{z} \in \mathbf{Z}_i, \mathbf{z} \neq \mathbf{y}_i$, we denote the estimated clairvoyant cost by $\hat{C}_i^C(\mathbf{z})$. For notational simplicity, for $\mathbf{z} = \mathbf{y}_i$ as well, we denote the clairvoyant cost by $\hat{C}_i^C(\mathbf{z})$, where it is understood that $\hat{C}_i^C(\mathbf{z}) = C_i^C(\mathbf{z})$. The optimal clairvoyant cost is the clairvoyant cost associated with an optimal staffing vector (which need not be the actually observed staffing vector). We now pick a subset of our data (comprising of say $I < n$ days) as a 'test' set. For each day i in this test set, we use all previous days in the data to construct the objective operational estimate of the cost function using the forecast of the call profile for the day in question. This objective operational estimate is then optimized, with respect to a set of feasible staffing vectors to determine the optimal objective operational staffing vector $\mathbf{z}_{iO}^*$. The clairvoyant cost $\hat{C}_i^C(\mathbf{z}_{iO}^*)$ associated with the staffing vector $\mathbf{z}_{iO}^*$ is then compared with the optimal clairvoyant solution $\hat{C}_i^C(\mathbf{z}^*) = \min_{\mathbf{z}\in\mathbf{Z}_i} \hat{C}_i^C(\mathbf{z})$ for the same day i. This is done by computing the relative regret for $i = 1, \cdots, I$, given by

$$L_i^O = \frac{\hat{C}_i^C(\mathbf{z}_{iO}^*) - \hat{C}_i^C(\mathbf{z}^*)}{\hat{C}_i^C(\mathbf{z}^*)}$$

We also test an adapted version of the approach described in Bassamboo and Zeevi [12] to determine the corresponding optimal staffing vector $\mathbf{z}_{iB}^*$ for $i = 1, \cdots, I$. The clairvoyant cost $\hat{C}_i^C(\mathbf{z}_{iB}^*)$ associated with the staffing vector $\mathbf{z}_{iB}^*$ is then compared with the optimal clairvoyant solution $\hat{C}_i^C(\mathbf{z}^*)$ for the same day i (as defined above). This is done by computing the relative regret for $i = 1, \cdots, I$, given by

$$L_i^B = \frac{\hat{C}_i^C(\mathbf{z}_{iB}^*) - \hat{C}_i^C(\mathbf{z}^*)}{\hat{C}_i^C(\mathbf{z}^*)}$$

We now describe the particular data used by us in the numerical experiments.

The Data

We consider data from a small call center for one of Israel's banks. This data was obtained from the website of Prof. Avi Mandelbaum of the Technion, Haifa. The data archives all the calls handled by the call center, over the period of 12 months from January 1999 to December 1999. This center provides several types of basic services, as well as others, including stock trading and technical support, for users of the banks Internet site. On weekdays (Sunday to Thursday in Israel) the center is open from 7 AM to midnight. During working hours, at most 13 regular agents, 5 Internet agents, and 1 shift supervisor may be working.

A simplified description of the path that each call follows through the center is as follows. A customer calls one of several telephone numbers associated with the call center, with the number depending on the type of service sought. Except for rare busy signals, the customer is then connected to a Voice Response Unit (VRU) and identifies herself. While using the VRU, the customer receives recorded information, both general and customized (e.g., an account balance). It is also possible for the customer to perform some self-service transactions here, and 65% of the bank's customers actually complete their service via the VRU. The other 35% indicate the need to speak with an agent. If an agent is free who is capable of performing the desired service, then the customer and the agent are matched to start service immediately. Otherwise, the customer joins the tele-queue.

Customers in the tele-queue are nominally served on a first-come, first-served (FCFS) basis (which agrees with our assumption), and customers' positions in queue are distinguished by the times when they arrive. While waiting, each customer periodically receives information on his or her progress in the queue. More specifically, he or she is told the amount of time that the first person in queue has been waiting, as well as his or her approximate location in the queue. The announcement is replayed every 60 seconds or so, with music, news, or commercials intertwined.

In each of the 12 months of 1999, roughly $100,000$ - $120,000$ calls arrived to the system, with $65,000$ - $85,000$ of these terminating in the VRU. The remaining $30,000$ to $40,000$ calls per month involved callers who exited the VRU indicating a desire to speak to an agent. These calls are the focus of our study. About 80% of those requesting service were in fact served, and about 20% were abandoned before being served.

Each call that proceeds past the VRU can be thought of as passing through up to three stages, each of which generates distinct data. The first of these is the *arrival* stage, which is triggered by the call's exit from the VRU and generates a record of an *arrival* time. If no appropriate server is available, then the call enters the *queueing* stage. Three pieces of data are recorded for each call that queues: the time it entered the queue, the time it exited the queue, and the manner in which it exited the queue, by being served or abandoning. The time spent in the queue is computed from this data. In the last stage, *service*, the data recorded are, the starting and ending times of the service, the service time, and the name of the agent who served the call. Note that calls that are served immediately skip the queueing stage, and calls that are abandoned never enter the service stage.

In addition to these time stamps, each call record in the database includes a categorical description of the type of service requested. The main call types are regular (PS or PE, for service in Hebrew or English respectively), stock transaction (NE), new or potential customer (NW), and Internet assistance (IN).

Over the year, two important operational changes occurred. First, in January to July, all calls were served by the same group of agents, but beginning in August, Internet (IN) customers were served by a separate pool of agents. Thus, in August to December, the center can be considered to be two separate service systems, one for IN customers and another for all other types. Second, one aspect of the recording of service time data changed at the end of October. In several instances, our illustrative example is based on only the November and December data. November and December were also convenient, because they contained no Israeli holidays. In our example, we also restrict the data to include only regular weekdays, i.e. Sunday to Thursday, 7 AM to midnight; because these are the hours of full operation of the center.

Testing

In our numerical study, on account of the changes in the data from the month of August, we restricted our attention to 110 regular weekdays from August to December. Even among these days, there were 3 days which exhibited very erratic call arrival patterns. We excluded these days from our study, and considered a set of $n = 107$ days. For each of these n days, we considered $J = 17$ staffing intervals of interest, which corresponded to the hours of operation of the call center on a regular weekday. We assumed a set of $A = 50$ agents were available for staffing, although the total number of agents we actually observed was 28. We considered a set $|\mathbf{Z}_i|$ of 81 staffing vectors for each day i in our data set, which included the actual staffing vector $\mathbf{y}_i$ observed on that day. Now, for the cost estimation algorithm, we considered $L = 50$ sample paths for each of the staffing vectors under consideration for each day in the data set. The empirical time to abandonment and service time distributions were estimated using all the $n = 107$ days in the data set.

Having generated cost estimates for each of the feasible staffing vectors for each of the $n = 107$ days in our data set, we defined a test set of $I = 44$ days, comprising the regular weekdays in November and December. It was then observed that the actual cost for the actual staffing vector was severely underestimated by the corresponding estimated cost. We needed an estimate of the clairvoyant cost for each feasible staffing vector for each day in the test set. While the clairvoyant cost for the actual staffing vector would just be given by the actual cost observed, for any other feasible staffing vector, the generated cost estimates would not do. We therefore used the heuristic (described in Section 3.3) to estimate the clairvoyant cost for feasible staffing vectors $\mathbf{z} \neq \mathbf{y}_i$.

For each of the days in the test set, we used the call arrival information from all of the previous days to generate a forecast of the call profile for that day. In this case, we applied the approach of Shen and Huang [77], described in the previous section. For all days in the test set, it was found that a single univariate time series $\{\beta_{i1}\}$ from the singular value decomposition of the call matrix, sufficed to approximate the call count information reasonably well. For this single univariate time series, on plotting the partial autocorrelation function, an AR(2) time series model was found to be appropriate, for all days in the test set.

For each $i = 1, \cdots, I$, we now performed smoothing of the estimated cost function and determine the optimal objective operational staffing vector $\mathbf{z}^*_{iO}$. All the smoothing parameters h_1, h_2, and h_3 were set to 1. The clairvoyant cost $\hat{C}^C_i(\mathbf{z}^*_{iO})$ associated with the staffing vector $\mathbf{z}^*_{iO}$ was then compared with the optimal clairvoyant solution $\hat{C}^C_i(\mathbf{z}^*) = \min_{\mathbf{z}\in\mathbf{Z}_i} \hat{C}^C_i(\mathbf{z})$ for the same day i. This was done by computing the relative regret for $i = 1, \cdots, I$, given by

$$L^O_i = \frac{\hat{C}^C_i(\mathbf{z}^*_{iO}) - \hat{C}^C_i(\mathbf{z}^*)}{\hat{C}^C_i(\mathbf{z}^*)}$$

We also tested an adaptation of the approach described in Bassamboo and Zeevi [12] to determine the corresponding optimal staffing vector $\mathbf{z}^*_{iB}$ for $i = 1, \cdots, I$. The approach was adapted as follows. In Bassamboo and Zeevi [12], the abandonment cost corresponding to a particular instantaneous arrival rate λ and staffing level z_j was determined as the optimal value of a linear program. This abandonment cost was then summed over the staffing period of interest j and its expectation was taken with respect to the distribution of the arrival rate process, $\Lambda_j = (\Lambda_j(t) : 0 \leq t \leq T_j)$. Here T_j was the length of the staffing interval of interest (one hour). The following cumulative distribution function (c.d.f.) was defined:

$$G_j(\lambda) = \frac{1}{T_j}\int_0^{T_j} \mathbb{P}(\Lambda_j(s) \leq \lambda)ds \tag{3.35}$$

where $\lambda \in \mathbb{R}_+$ and $G_j(\lambda)$ was interpreted as the expected fraction of time (within the staffing period of interest $[0, T_j]$) during which $\Lambda_j(\cdot) \leq \lambda$. The cost function $V(\cdot)$ was then expressed as follows:

$$V(z_j) = c_S z_j + T_j \int_{\lambda \in \mathbb{R}_+} c_A \max(0, \lambda - \mu_j z_j) dG_j(\lambda) \tag{3.36}$$

where μ_j is the mean service rate associated with the staffing period of interest. Now, the main idea was to use historical call arrival observations to approximate the distribution G_j given in (3.35). A straightforward nonparametric method for estimating the arrival rate was based on counting the number of arrivals over a small window and dividing by the window length. In our case, we divided each staffing period of interest into 10 windows of size ds in each of which, the arrival rate was assumed to be constant. We now formed the empirical counterpart to (3.35) using the rate estimators described as follows:

$$\hat{G}_{j_n}(\lambda) = \frac{1}{T_j} \int_0^{T_j} \frac{1}{n} \sum_{l=1}^{n} 1\{\hat{\Lambda}_j^l(s) \leq \lambda\} ds \tag{3.37}$$

where $\hat{\Lambda}_j^l(s)$ is the constant arrival rate in the window ds. We also estimated the mean service rate by taking the reciprocal of the sample mean of the service times of all previous days. Based on the above empirical distribution, we constructed the empirical counterpart of $V(\cdot)$,

$$\hat{V}_n(z_j) = c_S z_j + T_j \int_{\lambda \in \mathbb{R}_+} c_A \max(0, \lambda - \hat{\mu}_j z_j) d\hat{G}_{j_n}(\lambda)$$

The optimal staffing vector $\mathbf{z}_{iB}^*$ in the adapted Bassamboo-Zeevi approach then had the following components, for each $j = 1, \cdots, J$,

$$z_j^* = \arg \min_{z_j | \mathbf{z} \in \mathbf{Z}_i} \hat{V}_n(z_j) \tag{3.38}$$

The clairvoyant cost $\hat{C}_i^C(\mathbf{z}_{iB}^*)$ associated with the staffing vector $\mathbf{z}_{iB}^*$ was then compared with the optimal clairvoyant solution $\hat{C}_i^C(\mathbf{z}^*)$ for the same day i (as defined above). This was done by computing the relative regret for $i = 1, \cdots, I$, given by

$$L_i^B = \frac{\hat{C}_i^C(\mathbf{z}_{iB}^*) - \hat{C}_i^C(\mathbf{z}^*)}{\hat{C}_i^C(\mathbf{z}^*)}$$

Figure 3.1 shows a plot of L_i^O and L_i^B over the test set. Figure 3.2 also shows a plot of L_i^O and L_i^B over the test set, but with the smoothing parameters h_1 and h_2 set to 10, and h_3 still set to 1. We note that our approach does provide improvement for almost all days in the test set, over the adapted Bassamboo Zeevi approach. We admit that our particular adaptation of the approach of Bassamboo and Zeevi deserves critical scrutiny, and it is certainly not our intention to show their approach in a poor light. Indeed, we acknowledge the importance of their work and the inspiration it provided us for our own work. Yet, the plot does strongly suggest that our approach is highly worthy of serious consideration. In the next section, we consider the sensitivity of the relative regret to the smoothing parameters.

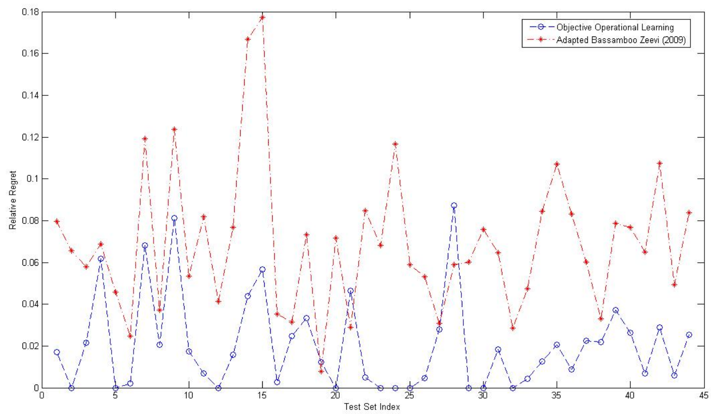

Figure 3.1: Relative regret over test set ($h_1 = h_2 = h_3 = 1$)

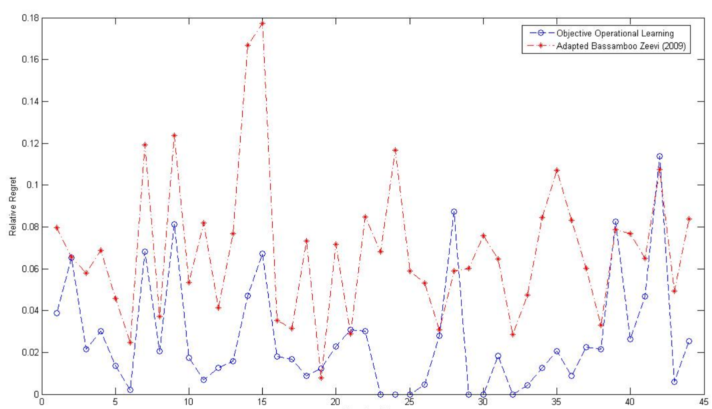

Figure 3.2: Relative regret over test set ($h_1 = h_2 = 10, h_3 = 1$)

Table 3.1: Mean relative regret over test set

h_1	h_2	h_3	mean relative regret
1	1	1	0.020384
1	1	10	0.2046
1	1	100	0.16349
1	10	1	0.020384
1	10	10	0.2046
1	10	100	0.16349
1	100	1	0.020384
1	100	10	0.2046
1	100	100	0.16349
10	1	1	0.02675
10	1	10	0.2046
10	1	100	0.16349
10	10	1	0.02675
10	10	10	0.2046
10	10	100	0.16349
10	100	1	0.02675
10	100	10	0.2046
10	100	100	0.16349
100	1	1	0.032076
100	1	10	0.20425
100	1	100	0.16349
100	10	1	0.032076
100	10	10	0.20425
100	10	100	0.16349
100	100	1	0.032076
100	100	10	0.20425
100	100	100	0.16349

3.5 Discussion

We now look at how the relative regret varies for different values of h_1, h_2, and h_3. Figures 3.3, 3.4, and 3.5 show how the relative regret varies when each of the parameters is varied in turn. Figure 3.6 shows how the relative regret varies when all the smoothing parameters are set to 100. Table 3.5 shows the mean relative regret over the test set for all combinations of h_1, h_2, h_3 taking one of the three values 1, 10, or 100.

We can make the following observations about the sensitivity of relative regret with respect to the smoothing parameters h_1, h_2, and h_3:

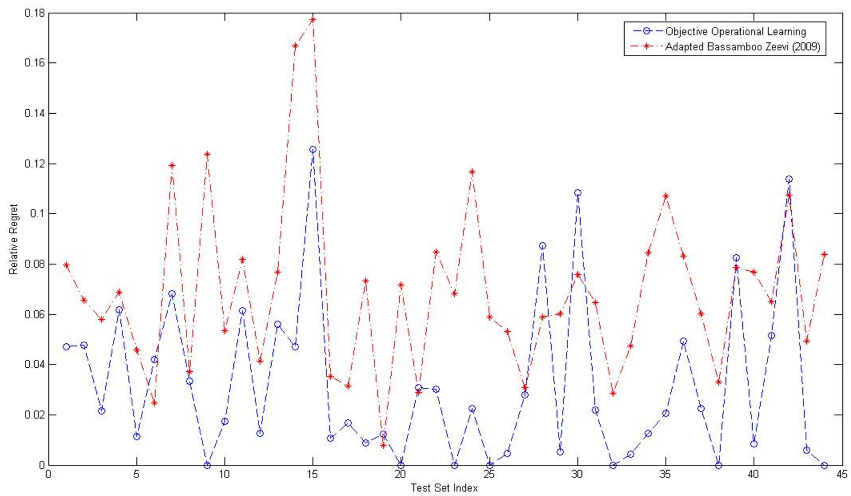

Figure 3.3: Relative regret over test set $(h_1 = 100, h_2 = h_3 = 1)$

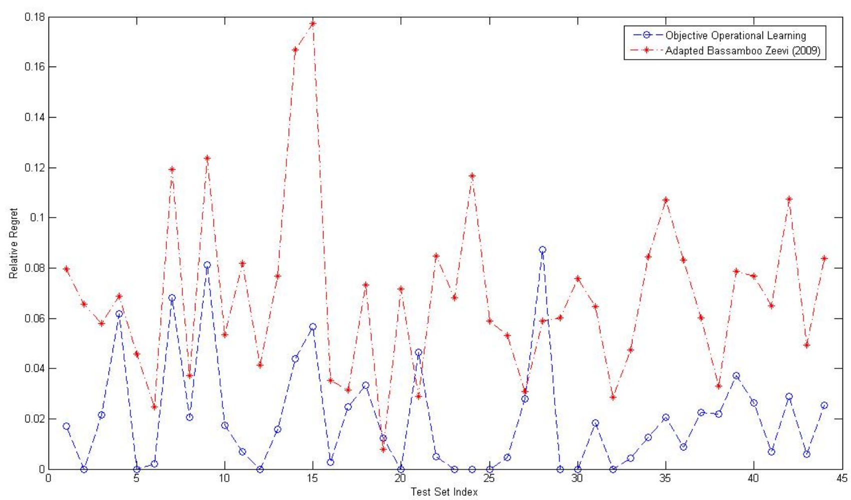

Figure 3.4: Relative regret over test set $(h_2 = 100, h_1 = h_3 = 1)$

- The relative regret increases as we increase the value of h_1 through 1, 10 and 100.
- The relative regret is insensitive to changes in the value of h_2 through 1, 10 and 100.
- As we increase the value of h_3 from 1 to 10, the relative regret increases, and then decreases as we increase h_3 from 10 to 100. Furthermore, we note that for $h_3 = 10$, or $h_3 = 100$, our approach performs worse than the adapted Bassamboo Zeevi approach

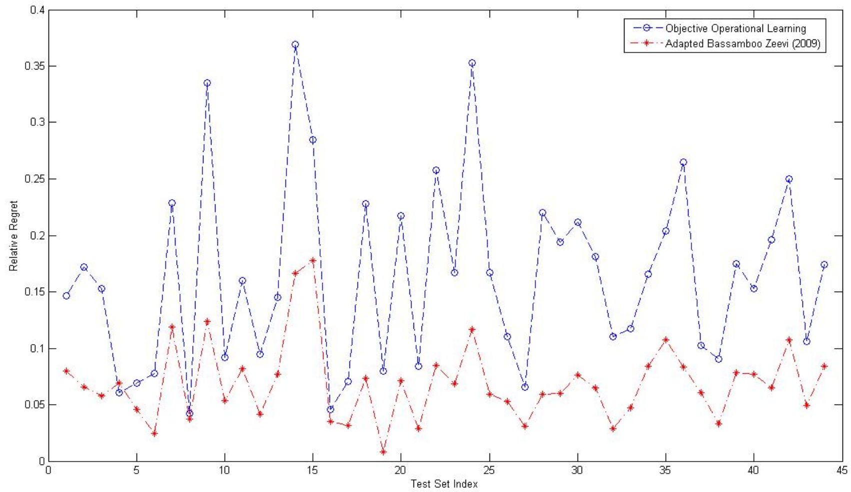

Figure 3.5: Relative regret over test set ($h_3 = 100, h_1 = h_2 = 1$)

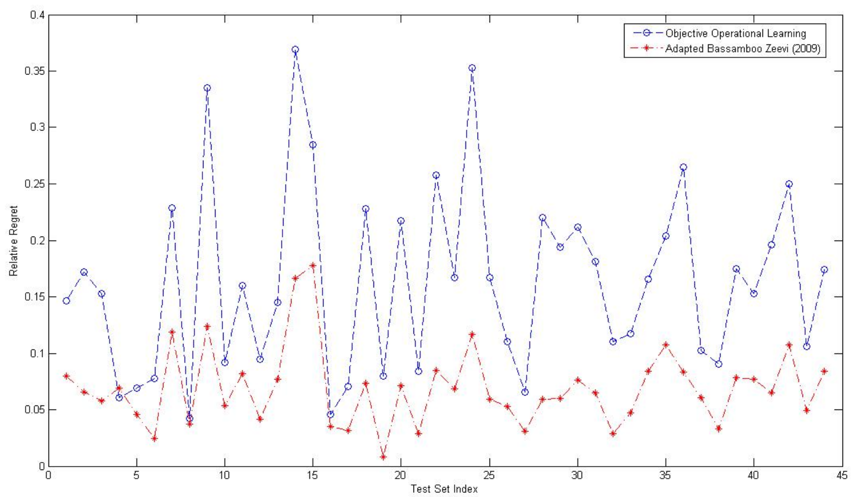

Figure 3.6: Relative regret over test set ($h_1 = h_2 = h_3 = 100$)

over the test set (for any values of h_1 and h_2). Only for $h_3 = 1$, is our approach better than the adapted Bassamboo Zeevi approach over the test set (for any values of h_1 and h_2). Hence, it is necessary to optimally set the value of h_3 in order to achieve best results.

Finally, we consider the effect of the forecasting approach used, on the relative regret. In addition to the approach of Shen and Huang [77], we use a simpler approach that considers a time series of the total number of call arrivals per day, which gives us a univariate time series. For this univariate time series, on plotting the partial autocorrelation function, an AR(2) time series model is found to be appropriate, for all days in the test set. We find that the relative regret is quite insensitive to the particular forecasting approach used, regardless of the smoothing parameter values. While this may simply be a consequence of the particular data set we used, it is comforting to note that we may use the best forecasting approach available to us, without seriously compromising performance. Figure 3.7 illustrates the same. Here, 'Forecast Method 1' refers to the approach of Shen and Huang [77], while 'Forecast Method 2' refers to the approach that considers a time series of the total number of call arrivals per day.

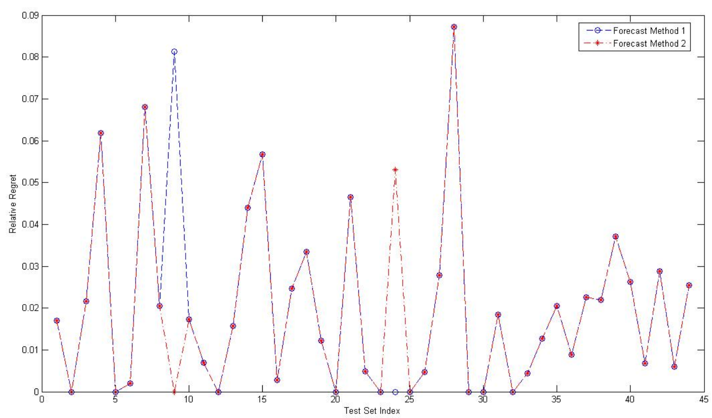

Figure 3.7: Relative regret over test set ($h_1 = h_2 = h_3 = 1$)

Chapter 4

Dynamic Assortment Planning

4.1 Introduction

Selecting an optimal assortment of products for display is one of the most important decisions that face retailers. In recent times, companies have been able to revise their product assortment decisions as demand information became available. Hence, it is common now to try to infer consumer preferences and update product assortments offered accordingly. Often, due to a variety of capacity related considerations, the retailer is unable to simultaneously display every possible product to prospective customers. One of the primary decisions is then to determine which products to include in the retailer's product assortment, in order to maximize the expected revenues. This problem is referred to as the *assortment planning problem*; see Kök et al. [54] for an overview. In this paper, we are interested in *dynamic* instances of this problem. In this case, the revenue distributions of the products are unknown, and the assortment planning decisions are revisited in every period and updated based on demand information derived from assortments offered in previous periods. This is referred to as the *dynamic assortment planning* problem. Following are two motivating examples that arise in quite different contexts:

Example 1: Fast fashion. In recent years "fast" fashion companies such as Zara, Mango or World Co have implemented supply chains that allow them to make and revisit most product design and assortment decisions during the selling season. Customers visiting one of their stores will only see a fraction of the potential products that the retailer has to offer, and their purchase decisions will effectively depend on the specific assortment presented at the store. Fashion retail basically involves offering new products for which no demand information is available, and hence the ability to revisit these decisions at a high frequency is key to the "fast fashion" business model. Every season there is a need to learn the current fashion trend by exploring with different styles and colors, and to exploit such knowledge before the season is over.

Example 2: On-line advertising. This emerging area of business is the single most important source of revenues for thousands of web sites. Giants such as Yahoo and Google, depend almost completely on on-line advertisement for their revenues. One of the most prevalent business models here builds on the cost-per-click statistic: advertisers pay the web site (a "publisher") only when a user clicks on their ads. Upon each visit, users are presented with a finite set of ads, on which they may or may not click depending on what is being presented. Roughly speaking, the publishers objective is to learn ad click-through-rates (and their dependence on the set of ads being displayed) and present the set of ads that maximizes revenues within the life span of the contract with the advertiser.

The above motivating applications share common features. For products/ads for which little or no demand information is available a priori, retailers/publishers must learn their desirability/effectiveness by dynamically adjusting their product/ad offering and observing customer behavior. It is natural to think that any good assortment strategy should gather some information on consumer preferences before committing to assortments that are thought to be profitable. This is the classical "exploration versus exploitation" trade-off: on the one hand, the longer a retailer/publisher spends learning consumer preferences, the less time remains to exploit that knowledge and optimize profits. On the other hand, less time spent on studying consumer behavior translates into more residual uncertainty, which could hamper revenue maximization objective. Moreover, demand information must be gathered carefully as product/ad profitability depends on the assortments offered: the retailer/publisher may learn consumer preferences more effectively by experimenting with a particular set of assortments.

The purpose of this chapter is to study a family of stylized dynamic assortment problems in which the retailer needs to devise an assortment policy to maximize revenues over the relevant time horizon by properly adapting the offered assortment based on observed customer purchase decisions and subject to capacity constraints that limit the size of the assortment. Our main focus in this work is on the impact of learning consumer behavior via suitable assortment experimentation, and doing this in a manner that guarantees minimal revenue loss over the selling horizon. To shed light on this facet of the problem, we ignore other effects such as inventory considerations, additional costs (such as assortment switching costs), operational constraints (e.g. restrictions on the sequence of offered assortments), and finally, we assume that product prices are fixed throughout the selling season. Returning to the motivating examples we discussed earlier, it is worth noting that such considerations are absent almost altogether from the on line advertisement problem, and are often ignored in the fast fashion setting; see, for example, the work of Caro and Gallien [21].

4.2 Literature Review

Static Assortment Planning

The static assortment planning literature focuses on finding an optimal assortment that is held unchanged throughout the entire selling season. Customer behavior is assumed to be known a priori, but inventory decisions are considered. Kök et al. [54] review the considerable literature and industry practice in static assortment planning. A seminal paper for choice modeling is McFadden [63], which introduced the popular multinomial logit model (MNL). Mahajan and van Ryzin [71] consider a category of product variants for which a retailer must construct an assortment. They analyze this problem using a multinomial logit model to describe the consumer choice process and a newsboy model to represent the retailer's inventory cost. They show that the optimal assortment has a simple structure and provide insights on how various factors affect the optimal level of assortment variety. Smith and Agrawal [78] develop a probabilistic demand model for items in an assortment that captures the effects of substitution and a methodology for selecting item inventory levels so as to maximize total expected profit, subject to given resource constraints. Illustrative examples are solved to provide insights concerning the behavior of the optimal inventory policies, using the negative binomial demand distribution.

Mahajan and van Ryzin [61] analyze a single-period, stochastic inventory model in which a sequence of heterogeneous customers dynamically substitute among product variants within a retail assortment when inventory is depleted. The customer choice decisions are based on a natural and classical utility maximization criterion. Faced with such substitution behavior, the retailer must choose initial inventory levels for the assortment to maximize expected profits. Using a sample path analysis, they analyze structural properties of the expected profit function. Talluri and van Ryzin [79] analyze a single-leg reserve management problem in which the buyers' choice behavior is modeled explicitly. The choice model is very general, simply specifying the probability of purchase for each fare product as a function of the set of fare products offered. The control problem is to decide which subset of fare products to offer at each point in time. They show that the optimal policy for this problem has a simple form which consists of identifying an ordered family of "efficient" subsets.

Gaur and Honhon [35] consider a single-period assortment planning and inventory management problem for a retailer, using a locational choice model to represent consumer demand. They first determine the optimal variety, product location, and inventory decisions under static substitution, and show that the optimal assortment consists of products equally spaced out such that there is no substitution among them regardless of the distribution of consumer preferences. They then obtain bounds on profit when customers dynamically substitute, using the static substitution for the lower bound, and a retailer-controlled substitution for the upper bound. They thus define two heuristics to solve the problem under dynamic substitution and numerically evaluate their performance. This analysis shows the value of modeling dynamic substitution and identifies conditions in which the static substitution solution serves as a good approximation. Kök and Fisher [53] study an assortment planning model in which consumers might accept substitutes when their favorite product is unavailable. They develop an algorithmic process to help retailers compute the best assortment for each store. First, they present a procedure for estimating the parameters of substitution behavior and demand for products in each store, including the products that have not been previously carried in that store. Second, they propose an iterative optimization heuristic for solving the assortment planning problem.

Goyal et al. [37] consider a single-period joint assortment and inventory planning problem under dynamic substitution with stochastic demands, and provide complexity and algorithmic results as well as insightful structural characterizations of near-optimal solutions for important variants of the problem. The work of Rusmevichientong et al. [70] identifies a polynomial-time algorithm for the static optimization problem when consumer preferences are represented using the MNL model. Vulcano et al. [83] consider a method for estimating substitute and lost demand when only sales and product availability data are observable, not all products are available in all periods (e.g., due to stock-outs or availability controls imposed by the seller), and the seller knows its market share. The model combines a multinomial logit (MNL) choice model with a non-homogeneous Poisson model of arrivals over multiple periods. The key idea is to view the problem in terms of primary (or first-choice) demand; that is, the demand that would have been observed if all products were available in all periods. They then apply the expectation-maximization (EM) method to this model, and treat the observed demand as an incomplete observation of primary demand. This leads to an efficient, iterative procedure for estimating the parameters of the model, which provably converges to a stationary point of the incomplete data log-likelihood function. Finally, Honhon et al. [44] present an efficient dynamic programming algorithm to determine the optimal assortment and inventory levels in a single-period problem with stockout-based substitution.

Dynamic Assortment Planning

This problem setting allows us to revisit assortment decisions at each point in time as more information is collected about initially unknown demand/consumer preferences. To the best of our knowledge, Caro and Gallien [21] were the first to study this type of problem, motivated by applications in fast fashion. Focusing on a stylized version of this problem, they study a finite horizon multiarmed bandit model with several plays per stage and Bayesian learning. Their analysis involves the Lagrangian relaxation of weakly coupled dynamic programs (DPs), results contributing to the emerging theory of DP duality, and various approximations. It yields a closed-form dynamic index policy capturing the key exploration versus exploitation trade-off and associated suboptimality bounds. The simplicity of their policy enables extensions to more realistic versions of the motivating dynamic assortment problem that include implementation delays, switching costs, and demand substitution effects. Bertsimas and Mersereau [14] present a Bayesian formulation of the problem in which decisions are made for batches of customers simultaneously, although decisions may vary within a batch. This extends the classical multiarmed bandit problem for sampling one-by-one from a set of reward populations. Their solution methods include a Lagrangian decomposition-based approximate dynamic programming approach and a heuristic based on a known asymptotic approximation to the multiarmed bandit solution. Rusmevichientong at al. [70] consider a stylized model of a dynamic assortment optimization problem, where given a limited capacity constraint, they must decide the assortment of products to offer to customers to maximize the profit. They assume that each customer chooses to purchase the product (or to click on the ad) that maximizes her utility, and use the multinomial logit choice model to represent the unknown demand. They present an adaptive policy for joint parameter estimation and assortment optimization. and show that the running average expected profit generated by their policy converges to the benchmark profit and establish its convergence rate. Saure and Zeevi [72] study a family of stylized assortment planning problems, where arriving customers make purchase decisions among offered products based on maximizing their utility. Given limited display capacity and no a priori information on consumers utility, the retailer must select which subset of products to offer. By offering different assortments and observing the resulting purchase behavior, the retailer learns about consumer preferences, but this experimentation should be balanced with the goal of maximizing revenues. They develop a family of dynamic policies that judiciously balance the aforementioned tradeoff between exploration and exploitation, and prove that their performance cannot be improved upon in a precise mathematical sense. One salient feature of these policies is that they “quickly” recognize, and hence limit experimentation on, strictly suboptimal products. Finally, motivated by the real world problems of identifying the ‘right’ model of comsumer choice, Jagabathula et al. [47] visit the following problem: For a ‘generic’ model of consumer choice (namely, distributions over preference lists) and a limited amount of data on how consumers actually make decisions (such as marginal information about these distributions), how may one predict revenues from offering a particular assortment of choices? They present a framework to answer such questions and design a number of tractable algorithms from a data and computational standpoint for the same. This paper thus takes a significant step towards automating the crucial task of choice model selection in the context of operational decision problems.

Multi-armed Bandit Problems

In the canonical multi-armed bandit problem the decision maker can select in each period to pull a single arm out of a set of K possible arms, where each arm delivers a random reward whose distribution is not known a priori, and the objective is to maximize the expected revenue over a finite horizon. Robbins [68] motivated this problem, and later Lai and Robbins [55] gave a classical formulation for the problem. Following this, the multiarmed bandit problem with multiple plays was formulated in two papers by Anantharam et al. [3, 4]. In the first [3], at each instant of time we are required to sample a fixed number $m \geq 1$ out of N i.i.d. processes whose distributions belong to a family suitably parameterized by a real number θ. The objective is to maximize the long run total expected value of the samples. Following Lai and Robbins [55], the learning loss of a sampling scheme corresponding to a configuration of parameters $C = (\theta_1, \cdots, \theta_N)$ is quantified by the *regret* $R_n(C)$. This is the difference between the maximum expected reward at time n that could be achieved if C were known and the expected reward actually obtained by the sampling scheme. They provide a lower bound for the regret associated with any uniformly good scheme, and construct a scheme which attains the lower bound for every configuration C. The lower bound is given explicitly in terms of the Kullback-Liebler number between pairs of distributions. In the second paper [4], the authors consider the same problem when the reward processes are Markovian. We provide further details of these particular works in a subsequent section.

Lai [56] proposes a class of simple adaptive allocation rules for the multi-armed bandit problem where reward densities belong to an exponential family. The rules are shown to be asymptotically optimal from both Bayesian and frequentist points of view. Agrawal [1] considers a non-Bayesian infinite horizon version of the multi-armed bandit problem with the objective of designing simple policies whose regret increases slowly with time. In this paper he constructs index policies that depend on the rewards from each arm only throught their sample mean. These policies are computationally much simpler and are also applicable much more generally. They achieve a $O(\log n)$ regret with a constant that is also based on the Kullback-Leibler number. Auer et al. [8] make no statistical assumptions whatsoever about the nature of the process generating the payoffs of the slot machines. They give a solution to the bandit problem in which an adversary, rather than a well-behaved stochastic process, has complete control over the payoffs. In addition, they prove results about the rate of convergence of the expected per-round payoff of their algorithm and performance bounds on general algorithms in their setting. Brezzi and Lai [19] study how and how much active experimentation is used in discounted or finite-horizon optimization problems with an agent who chooses actions sequentially from a finite set of actions, with rewards depending on unknown parameters associated with the actions. Finally, Auer et al. [9] show that the optimal logarithmic regret of the multi-armed bandit problem is also achievable uniformly over time, with simple and efficient policies, and for all reward distributions with bounded support.

Motivation

From the foregoing discussion, we see that most models for dynamic assortment optimization incorporate various models for consumer choice. A popular model is the Multinomial Logit Choice model. While this model is widely used, it does possess a somewhat restrictive property known as the *independence from irrelavant alternatives* (IIA) property, which basically says that the relative likelihood of choosing between two alternatives is independent of the choice set containing these alternatives. This property is not realistic, however, if the choice set contains alternatives that can be grouped such that alternatives within a group are more similar than alternatives outside the group because adding a new alternative reduces the probability of choosing similar alternatives more than dissimilar alternatives (see Talluri and van Ryzin [80]). More generally, using any parametric demand model for consumer choice requires estimating parameters of the model first, and then incorporating these estimates into an optimization algorithm. Considering that, in practice, prior data is often insufficient to justify any particular parametric model, using estimates of such parametric models for optimizing assortments is likely to reduce the reliability, or increase the variance, of solutions found. Hence, we would like to eschew, as far as possible, any particular parametric demand model and make minimal assumptions with respect to it. Toward this end, we seek to develop an adaptive, non-parametric approach to dynamic assortment optimization that directly maps assortments to revenues without estimating any parameters in the process.

Furthermore, it may be seen from the discussion above that there is a thematic connection between multi-armed bandits and assortment planning problems, in the sense that both look to balance exploration and exploitation. However, as noted in Rusmevichientong et al. [70], in the single play formulation of the multiarmed bandit problem (see Lai and Robbins [55], Auer et al. [9]) the number of possible decisions grows exponentially with the capacity constraint and corresponding algorithms have running times and convergence rates that scale linearly with the number of decisions. Hence, it is curious, that while some papers have cited Anantharam et al. [3, 4] and their multiple play formulation, none of them have sought to apply the formulation to the dynamic assortment optimization problem. Moreover, none of the papers have even offered a convincing argument for eschewing the approach of Anantharam et al. [3, 4], although it seems tailor-made for this problem. It is our intention to explore in this chapter, the applicability of the multiple play multi-armed bandit approach in the dynamic assortment planning problem. Toward this end, we detail in the following section, the contributions of Anantharam et al. [3, 4]. Thereafter, we outline an adaptive, non-parametric approach to dynamic assortment planning that incorporates some ideas from Anantharam et al. [3, 4]. The critical difference between our approach and that of Anantharam et al. [3, 4], is that our apparoach takes into account correlations and substitution effects between products, as we will see in a later section.

4.3 Model Formulation

Multiarmed Bandit Problem with Multiple Plays

In this section, we summarize the contributions of Anantharam et al. [3, 4]. We focus on the results in Anantharam et al. [3], where the reward processed are i.i.d. The results in Anantharam et al. [4] are similar, and they differ only in that reward processes are Markovian.

I.I.D. Rewards

Anantharam et al. [3] consider a version of the multiarmed bandit problem with multiple plays. They are given a one-parameter family of reward distributions with densities $f(x, \theta)$ with respect to some measure ν on $\mathbf{R}$. θ is a real valued parameter. There are N arms X_j, $j = 1, \cdots, N$ with parameter configuration $C = (\theta_1, \cdots, \theta_N)$. When arm j is played, it gives a reward with distribution $f(x, \theta_j)d\nu(x)$. Successive plays of arm j produce i.i.d. rewards. At each stage they are required to play a fixed number, m, of the arms, $1 \leq m \leq N$.

The distributions of the individual rewards are assumed to be known. To maximize the total expected reward up to any stage, one must play the arms with the m highest means. However, if the parameters θ_j, are unknown, the poorer arms are forced to be played in order to learn about their means from the observations. The aim is to minimize, in some sense, the total expected loss incurred in the process of learning for every possible parameter configuration.

The setup is as follows. The actual values θ that can arise as parameters of the arms are known a priori to belong to a subset $\Theta \subseteq \mathbf{R}$. The rewards are assumed to be integrable.

$$\int_{-\infty}^{\infty} |x| f(x,\theta) d\nu(x) < \infty \tag{4.1}$$

and the mean reward

$$\mu(\theta) = \int_{-\infty}^{\infty} x f(x,\theta) d\nu(x)$$

is a strictly monotone increasing function of the parameter θ. In general, $0 \leq I(\theta,\lambda) \leq \infty$, where $I(\theta,\lambda)$ is the Kullback-Liebler number. It is assumed that

$$0 < I(\theta,\lambda) < \infty \text{ if } \lambda > \theta \tag{4.2}$$

and

$$I(\theta,\lambda) \text{ is continuous in } \lambda > \theta \text{ for fixed } \theta \tag{4.3}$$

Initially, the following denseness condition on Θ is imposed:

$$\text{for all } \lambda \in \Theta \text{ and } \delta > 0, \text{ there is } \lambda' \in \Theta \text{ s.t. } \mu(\lambda) < \mu(\lambda') < \mu(\lambda) + \delta \tag{4.4}$$

(For details on the case when the denseness condition is removed, the reader is referred to Anantharam et al. [3]). $Y_{j1}, Y_{j2}, \cdots$ denote successive rewards from arm j. $\mathbf{F}_t(j)$ denotes the σ-algebra generated by $Y_{j1}, \cdots, Y_{jt}$, $F_\infty(j) = \vee_t \mathbf{F}_t(j)$, and $\mathbf{G}(j) = \vee_{i \neq j} \mathbf{F}_\infty(i)$. An adaptive allocation rule is a rule for deciding which m arms to play at time $t+1$ based only on knowledge of the past rewards $Y_{j1}, \cdots, Y_{jT_t(j)}$, $j = 1, \cdots, N$ and the past decisions. For an adaptive allocation rule Φ, the number of plays made of arm j by time t, $T_t(j)$, is a stopping time of $\{\mathbf{F}_s(j) \vee \mathbf{G}(j), s \geq 1\}$. By Wald's lemma, if S_t denotes the total reward received up to time t,

$$ES_t = \sum_{j=1}^{N} \mu(\theta_j) ET_t(j) \tag{4.5}$$

For a configuration $(\theta_1, \cdots, \theta_N)$, the loss associated to a rule is a function of the number of plays t which gives the difference between the expected reward that could have been achieved with prior knowledge of the parameters and the expected reward actually achieved under the rule. Following Lai and Robbins [55], this function is called the *regret*. σ is a permutation of $\{1, \cdots, N\}$ such that

$$\mu(\theta_{\sigma(1)}) \geq \mu(\theta_{\sigma(2)}) \geq \cdots \geq \mu(\theta_{\sigma(N)})$$

Then, the regret is

$$R_t(\theta_1, \cdots, \theta_N) = t \sum_{i=1}^{m} \mu(\theta_{\sigma(i)}) - ES_t \tag{4.6}$$

The problem is to minimize the regret in some sense. A rule is called *uniformly good* if for every parameter configuration $R_t(\theta_1, \cdots, \theta_N) = o(t^\alpha)$ for every real $\alpha > 0$. Any rule that is not uniformly good is considered uninteresting. The arms are assumed to have parameter configuration $C = (\theta_1, \cdots, \theta_N)$ and σ is assumed to be a permutation of $\{1, \cdots, N\}$ such that

$$\mu(\theta_{\sigma(1)}) \geq \mu(\theta_{\sigma(2)}) \geq \cdots \geq \mu(\theta_{\sigma(N)})$$

(a) If $\mu(\theta_{\sigma(m)}) > \mu(\theta_{\sigma(m+1)})$, arms $\sigma(1), \cdots, \sigma(m)$ are called the distinctly m-best arms and $\sigma(m+1), \cdots, \sigma(N)$ the distinctly m-worst arms.

(b) If $\mu(\theta_{\sigma(m)}) = \mu(\theta_{\sigma(m+1)})$, and if $0 \leq l < m$ and $m \leq n \leq N$ are such that

$$\mu(\theta_{\sigma(1)}) \geq \cdots \mu(\theta_{\sigma(l)}) > \mu(\theta_{\sigma(l+1)}) = \cdots \mu(\theta_{\sigma(m)}) = \cdots = \mu(\theta_{\sigma(n)}) > \mu(\theta_{\sigma(n+1)}) \geq \cdots \geq \mu(\theta_{\sigma(N)})$$

Then arms

$$\sigma(1), \cdots, \sigma(l)$$

are called the distinctly m-best arms, and arms

$$\sigma(n+1), \cdots, \sigma(N)$$

the distinctly m-worst arms.

(c) The arms with mean equal to $\mu(\theta_{\sigma(m)})$ are called the m-border arms. Note that in (a) $\sigma(m)$ is both a distinctly m-best arm and an m-border arm. In (b) the m-border arms are the arm j, $l+1 \leq j \leq n$.

Φ is assumed to be an adaptive allocation rule. Then Φ is uniformly good iff for every distinctly m-best arm j

$$E(t - T_t(j)) = o(t^\alpha)$$

and for every distinctly m-worst arm j

$$E(T_t(j)) = o(t^\alpha)$$

for every real $\alpha > 0$.

Theorem 2. *(see Anantharam et al. [3]) Now, the family of reward distributions is assumed to satisfy conditions* (4.2), (4.3), *and* (4.4). *Φ is given to be a uniformly good rule. If the arms have parameter configuration $C = (\theta_1, \cdots, \theta_N)$, then for each distinctly m-worst arm j and each $\epsilon > 0$*

$$\lim_{t\to\infty} P_C \left\{ T_t(j) \geq \frac{(1-\epsilon)\log t}{I(\theta_j, \theta_{\sigma(m)})} \right\} = 1$$

so that

$$\liminf_{t\to\infty} \frac{E_C T_t(j)}{\log t} \geq \frac{1}{I(\theta_j, \theta_{\sigma(m)})}$$

where σ is a permutation of $\{1, \cdots, N\}$ such that

$$\mu(\theta_{\sigma(1)}) \geq \mu(\theta_{\sigma(2)}) \geq \cdots \geq \mu(\theta_{\sigma(N)})$$

Consequently,

$$\liminf_{t\to\infty} \frac{R_t(\theta_1, \cdots, \theta_N)}{\log t} \geq \sum_{j \text{ is } m\text{-worst}} \frac{\mu(\theta_{\sigma(m)}) - \mu(\theta_j)}{I(\theta_j, \theta_{\sigma(m)})}$$

for every configuration $C = (\theta_1, \cdots, \theta_N)$.

Motivated by Theorem 2, an adaptive allocation rule is called asymptotically efficient if for each configuration $(\theta_1, \cdots, \theta_N)$,

$$\limsup_{t\to\infty} \frac{R_t(\theta_1, \cdots, \theta_N)}{\log t} \leq \sum_{j \text{ is } m\text{-worst}} \frac{\mu(\theta_{\sigma(m)}) - \mu(\theta_j)}{I(\theta_j, \theta_{\sigma(m)})}$$

To construct an asymptotically efficient rule a technique is required for deciding when we need to experiment, i.e., when to play an arm in order to learn more about its parameter value from the additional sample. At time t there are $T_t(j)$ samples from arm j from which can be estimated θ_j by various methods, e.g., sample mean, maximum likelihood estimate, sample median. The decision to be made at time $t+1$ is whether to play the m arms whose estimated parameter values are the largest - "play the winners" rule - or to experiment by playing some of the apparently inferior arms. To do this, a family of statistics $g_{ta}(Y_1, \cdots, Y_a)$, $1 \leq a \leq t$, $t = 1, 2, \cdots$ is constructed, so that when $g_{tT_t(j)}$ is larger than any of the m best estimated parameter values, this indicates the need to experiment with arm j. Such statistics are constructed in Lai and Robbins [55] for exponential families of distributions, based on results of Pollak and Siegmund [66]. Anantharam et al. [3] use a similar technique to construct $g_{ta}(Y_1, \cdots, Y_a)$ under the following assumptions:

$$\log f(x, \theta) \text{ is concave in } \theta \text{ for each fixed } x, \tag{4.7}$$

$$\int x^2 f(x, \theta) d\nu(x) < \infty \text{ for each } \theta \in \mathbf{R}. \tag{4.8}$$

Theorem 3. *(see Anantharam et al. [3]) $Y_1, Y_2, \cdots$ are given to be the sequence of rewards from an arm. It is given that*

$$W_a(\theta) = \int_{-\infty}^{0} \prod_{b=1}^{a} \frac{f(Y_b, \theta + t)}{f(Y_b, \theta)} h(t) dt,$$

where $h : (-\infty, 0) \to \mathbf{R}_+$ is a strictly positive continuous function with $\int_{-\infty}^{0} h(t)dt = 1$. For any $K > 0$

$$U(a, Y_1, \cdots, Y_a, K) = \inf\{\theta | W_a(\theta) \geq K\}. \tag{4.9}$$

Then for all $\lambda > \theta > \eta$

(1) $P_\theta\{\eta < U(a, Y_1, \cdots, Y_a, K) \text{ for all } a \geq 1\} \geq 1 - \frac{1}{K}$,

(2) $\lim_{K\to\infty} \frac{1}{\log K} \sum_{a=1}^{\infty} P_\theta\{U(a, Y_1, \cdots, Y_a, K) \geq \lambda = \frac{1}{I(\theta,\lambda)}$.

Having observed samples $Y_1, \cdots, Y_a$, for any $\theta \in \mathbf{R}$, $W_a(\theta)$ is a natural statistic to test the compound hypothesis that the samples have been generated by a parameter value less than θ against the hypothesis that they have been generated by θ. By the log concavity assumption (4.7), $W_a(\theta)$ is increasing in θ. Therefore, for fixed K, for any $\theta > U(a, Y_1, \cdots, Y_a, K)$, it is more likely that the samples have been generated by parameter values below θ than by θ, whereas, for any $\theta < U(a, Y_1, \cdots, Y_a, K)$, it is more likely that the samples have been generated by θ than by parameter values below θ. When $U(a, Y_1, \cdots, Y_a, K)$ is used to decide if there is a need to experiment, K is chosen appropriately - the larger K is, the more certainty there will be that the samples have been generated by parameter values below θ before the possibility that they may have been generated by θ. This heuristic was suggested by Pollak and Siegmund [66].

Theorem 4. *(see Anantharam et al. [3])*
It is assumed that $g_{ta}(Y_1, \cdots, Y_a) = \mu[U(a, Y_1, \cdots, Y_a, t(\log t)^p)]$ *for some* $p > 1$. *Then for any* $\lambda > \theta > \eta$

$$\text{(1) } P_\theta\{g_{ta}(Y_1, \cdots, Y_a) > \mu\eta \text{ for all } a \leq t\} = 1 - O(t^{-1}(\log t)^{-p}); \tag{4.10}$$

$$\text{(2) } \limsup_{t\to\infty} \frac{\sum_{a=1}^{t} P_\theta\{g_{ta}(Y_1, \cdots, Y_a) \geq \mu\lambda\}}{\log t} \leq \frac{1}{I(\theta, \lambda)} \tag{4.11}$$

$$\text{(3) } g_{ta} \text{ is nondecreasing in } t \text{ for fixed } a. \tag{4.12}$$

The sample mean is taken as an estimate for the mean reward of an arm.

$$h_a(Y_1, \cdots, Y_a) = \frac{Y_1 + \cdots + Y_a}{a}.$$

Now, the N arms correspond to $(\theta_1, \cdots, \theta_N)$. It is assumed that the arms have been reindexed so that

$$\mu(\theta_1) \geq \cdots \geq \mu(\theta_N)$$

With g_{ta} and h_a as above, the following adaptive allocation rule is considered. We denote it as the "Anantharam-Varaiya-Walrand (AVW) Approach" for future reference.

(1) In the first N steps, each of the arms is sampled m times in some order to establish an initial sample.

(2) δ is chosen such that $0 < \delta < 1/N^2$. Consider the situation when we are about to decide which m arms to sample at time $t+1$. Clearly, whatever the preceding decisions, at least m among the arms have been sampled at least δt times. Among these "well-sampled" arms the m-leaders are chosen at stage $t+1$, namely the arms with the m best values of the statistic $\mu_t(j)$, $j = 1, \cdots, N$, where

$$\mu_t(j) = h_{T_t(j)}(Y_{j1}, \cdots, Y_{jT_t(j)}).$$

Let $j \in \{1, \cdots, N\}$ be the arm for which $t+1 \equiv j \mod N$. The statistic $U_t(j)$ is calculated, where

$$U_t(j) = g_{tT_t(j)}(Y_{j1}, \cdots, Y_{jT_t(j)})$$

(a) If arm j is already one of the m-leaders, then at stage $t+1$ the m-leaders are played.

(b) If arm j is not among the m-leaders, and $U_t(j)$ is less than $\mu_t(k)$ for every m-leader k, then again, the m-leaders are played.

(c) If arm j is not among the m-leaders, and $U_t(j)$ equals or exceeds the μ_t statistic of the least best of the m-leaders, then the $m-1$ best of the m-leaders and the arm j are played at stage t.

Note that in any case the $m-1$ best of the m-leaders always get played.

Theorem 5. *(see Anantharam et al. [3]) The rule above is asymptotically efficient.*

Adaptive Non-parametric Approach

We consider an online retailer that has a collection of N different products, each with a deterministic unit revenue of p_i, for $i = 1, \cdots, N$. Now, we have space for only $m \leq N$ products to be displayed. We assume that a random $X_i \in \mathbb{Z}_+$ number of people purchase a product i, for each $i = 1, \cdots, N$. The problem is now to choose $D \subset \{1, \cdots, N\}$ such that $|D| \leq m$ and $\sum_{i \in D} p_i \mathbb{E}[X_i]$ is maximized.

We desire an adaptive, non-parametric approach for choosing assortments, that maximizes the long run total expected revenues. A simple adaptation of the multiarmed bandit approach described above would require us to independently compute mean statistics of the revenues from each product, regardless of the particular assortments it was offered in. Such an approach would be ignoring correlations and substitution effects between products, that have been found to be important, as we noted from Gaur et al. [37] and Honhon et al. [44]. In order to take correlations between products into account, we propose the following revenue measure. We define an $N \times N$ matrix called the revenue score matrix, and denote it by S. We suppose that the assortments $D_1, \cdots, D_n$ were displayed in the last n periods. For each product i, we let $X_{i1}, \cdots, X_{in}$ denote the respective number of sales in the last n assortments. For each $i, j = 1, \cdots, N$, we define

$$S_{ij} = \frac{\sum_{k=1}^{n} \mathbf{1}\{i, j \in D_k\} p_i X_{ik}}{\sum_{k=1}^{n} \mathbf{1}\{i, j \in D_k\}} \tag{4.13}$$

We observe that the diagonal elements of the revenue score matrix are exactly equal to the average revenue generated by each product. Moreover, for each i, the diagonal element S_{ii} is positively correlated with S_{ij} for each $j = 1, \cdots, N$. We now describe an efficient procedure to determine the mth principal minor with the maximum sum. The procedure which we call "Best m", is as follows:

- Sort the rows of the revenue score matrix in decreasing order of the diagonal elements and call the new matrix P. Denote the sum of the leading $m$the principal minor by M_0. Let J_0 denote the ordered list of indices corresponding to the first C diagonal elements of P.

- For each $i = 1, \cdots, m$, sort the $i$the row of P in descending order, and denote by J_i, the ordered list of indices correponding to the first m elements of the sorted row. Check if $i \in J_i$. If yes, determine the principal minor corresponding to J_i. If not, let $J_i(m) = J_i(m-1)$, and $J_i(m-1) = i$. Here, $J_i(k)$ denotes the kth element of the ordered list J_i. Determine the principal minor corresponding to J_i. Denote the sum of the chosen principal minor by M_i.

- Find the maximum $M = \max_i\{M_i : i = 0, 1, \cdots, m\}$ and the corresponding principal minor Q. Let $J = J_l$, where $l = \arg\max_i\{M_i : i = 0, 1, \cdots, m\}$.

Proposition 1. *The worst-case time complexity of the procedure "Best m" for matrix S is $O(N \log N)$, when $m > 2$ is an integer much less than N.*

Proof: The proposition follows easily when we note that the procedure first involves the sorting of the N diagonal elements. Following that m lists of length N are sorted. The worst case time complexity of sorting a list of N elements when using a comparison-based sort is known to be $O(N \log N)$ (see Cormen et al. [25]). In our case, the worst-case complexity is given by $O((m+1)N \log N) \equiv O(N \log N)$ when m is a constant much less than N.

Furthermore, for each $i, j = 1, \cdots, N$, we define

$$\hat{\sigma}_{ij} = \sqrt{\frac{\sum_{k=1}^{n} \left(\mathbf{1}\{i,j \in D_k\} p_i X_{ik} - S_{ij}\right)^2}{\left(\sum_{k=1}^{n} \mathbf{1}\{i,j \in D_k\}\right) - 1}} \tag{4.14}$$

and

$$U_{ij} = S_{ij} + B\hat{\sigma}_{ij} \tag{4.15}$$

where B is a suitable parameter. We now propose the following objective, adaptive, non-parametric approach for choosing assortments. We denote it as the "Objective Operational Approach" for future reference.

(1) To begin with, sample atleast m times from each of the products in different assortments, such that every pair of products is in at atleast one of the assortments. This is our exploratory sample, whose size we denote by $N_E(m)$. Update S in each of the initial steps.

(2) Choose $0 < \delta < 1/N^2$. Consider the situation when we are about to decide which m products to sample in our assortment at time $t+1$. Clearly, whatever the preceding decisions, at least m among the products have been sampled at least δt times. Consider the sub-matrix S' of S corresponding to these "well-sampled" products and determine

the $m$the principal minor of S' with the maximum sum at stage $t+1$, using the "Best m" procedure described above. Denote the principal minor by Q and the corresponding set of indices by J, as in the "Best m" procedure.

Let $i \in \{1, \cdots, N\}$ be the product index for which $t+1 \equiv i \mod N$. If $i \in J$, then at stage $t+1$ choose the products given by J. If not, let $k = J(m)$. Calculate the statistic $U_t(i)$ where

$$U_t(i) = U_{ii} + \sum_{j \in J/\{k\}} (U_{ij} + U_{ji})$$

Also, let

$$V_t(k) = Q_{kk} + \sum_{j \in J/\{k\}} (Q_{kj} + Q_{jk})$$

(a) If $U_t(i) < V_t(k)$, then again choose the products given by J.

(b) If $U_t(i) \geq V_t(k)$, then choose product i and all products given by the ordered list J, except for k.

Now, let J^* be the ordered set of indices corresponding to the theoretical best assortment. For all $k = 1, \cdots, N$, we define

$$\mu_k = \mathbb{E}[V_t^{J^*}(k)]$$

where

$$V_t^{J^*}(k) = Q_{kk} + \sum_{j \in J^*/\{k\}} (Q_{kj} + Q_{jk}) \text{ if } k \in J^*$$

and

$$V_t^{J^*}(k) = Q_{kk} + \sum_{j \in J^*/\{J^*(m)\}} (Q_{kj} + Q_{jk}) \text{ otherwise}$$

We now make the following assumptions.

(A1) For all $k = 1, \cdots, N$, for any $\mu_l < \mu_k$, for some $p > 1$,

$$P\{U_t(k) > \mu_l\} = 1 - O(t^{-1}(\log t)^{-p})$$

(A2) For any $0 < \delta < 1$ and $\epsilon > 0$

$$P\{\max_{\delta t \leq a \leq t} |V_a(k) - \mu_k| > \epsilon\} = o(t^{-1})$$

for any $k = 1, \cdots, N$.

(A3) For all $j = 1, \cdots, N$, for any $\mu_m > \mu_j$,

$$\limsup_{t \to \infty} \frac{\sum_{a=1}^{t} P\{U_a(j) \geq \mu_m\}}{\log t} \leq \mathcal{K}_{jm}$$

for some constant $\mathcal{K}_{jm}$.

Theorem 6. *Assume that A1 - A3 hold. Then, the rule above is asymptotically efficient.*

Theorem 6 is proved in Appendix B.

4.4 Numerical Experiments

In this section, we report the results of our numerical experiments. First,we describe the data set and the model of the mean utilities. We then consider the dynamic assortment planning problem and compare the performance of our approach with that of an adapted AVW Approach.

Data Set and Model

Before we can evaluate the performance of our approach, we need to identify a set of products and specify their mean utilities. To help us understand the range of utility values that we might encounter in actual applications, we use the utilities estimated in Rusmevichientong et al. [70], using data on DVD sales at a large online retailer. They consider DVDs that are sold during a three-month period from July 1, 2005 through September 30, 2005. During this period, the retailer sold over over 4.3 million DVDs, spanning across 51,764 DVD titles.

To simplify their analysis, Rusmevichientong et al. [70] restrict their attention to customers who purchased DVDs that account for the top 33% of the total sales, and they assume that each customer purchases at most one DVD. This gives them a total of 1,409,261 customers in their data set. The products correspond to the 200 best-selling DVDs that account for about 65% of the total sales among their customers. They assume that all 200 DVDs are available for purchase, and when customers do not purchase these DVDs, they assign them to the no-purchase alternative. They observe that the best-selling DVD in their data set was purchased by only about 2.6% of the customers. In fact, among the top 10 best-selling DVDs, each one was sold to only around $1.1\% - 2.6\%$ of the customers. Thus, only a small fraction of the customers purchased each DVD.

Rusmevichientong et al. [70] assume a linear-in parameters utility model. The attributes of each DVD considered are the selling price (averaged over three months of data), customer reviews, total votes received by the reviews, running time, and the number of discs in the DVD collection. They obtain data on customer reviews and the number of discs of each DVD from the Amazon.com website through a publicly available interface via Amazon.com E-Commerce Services (http://aws.amazon.com). Each visitor to the Amazon.com website can provide a review and a rating for each DVD. The rating is on a scale of 1 to 5, with 5 representing the most favorable review. Each review can be voted by other visitors as either "helpful" or "not helpful". For each DVD, they consider all reviews up until June 30, 2005, and compute features such as the average rating, the proportion of reviews that give a 5 rating, the average number of helpful votes received by each review, and so on.

Under the linear-in-parameters utility model, for $i \in \{1, 2, \cdots, 200\}$, the mean utility μ_i of DVD i is given by

$$\mu_i = \alpha_0 + \sum_{k=1}^{F} \alpha_k \phi_{i,k},$$

where $(\phi_{i,1}, \cdots, \phi_{i,F})$ denotes the features of DVD i, and $\mu_0 = 0$. Rusmevichientong et al. [70] used the software BIOGEME developed by Bierlaire [16] to determine the most relevant DVD features and estimate the corresponding coefficients. It turned out that the two most relevant attributes were the total number of votes received by the reviews of each DVD and the price per disc (computed as the selling price divided by the number of discs in the DVD collection). They estimated that for each DVD $i = 1, \cdots, 200$,

$$\mu_i = -4.31 + (3.54 \times 10^{-5} \times \phi_{i,1}) - (0.038 \times \phi_{i,2}), \tag{4.16}$$

where

$$\begin{aligned} \phi_{i,1} &= \text{Total Number of Votes Received by All Reviews of DVD } i \\ \phi_{i,2} &= \text{Price Per Disc Associated with DVD } i \end{aligned}$$

Assuming a Multinomial Logit Model for the utilities, we assume each utility has a random component with a standard Gumbel distribution. We now need need a way to map a given utility to the revenue received from the product. For the purpose of our experiments, we use a linear mapping as follows. Over all 200 products, we denote the minimum and maximum mean utilites by μ_{min} and μ_{max} respectively. Denoting the correponding revenues by R_{min} and R_{max} respectively, for a given utility μ for a product in a particular assortment, the revenue received R from that product in that assortment is given by

$$R = R_{min} + \frac{\mu - \mu_{min}}{\mu_{max} - \mu_{min}} \cdot (R_{max} - R_{min}) \tag{4.17}$$

Computational Results

We consider assortment sizes ranging from $m = 3$ to $m = 20$. For each assortment size, we generate an initial set of assortments such that each product is sampled at least m times, and every pair of products appears in at least one assortment. In our experiments, we use $R_{min} = 10$, $R_{max} = 100$, and $B = 2$. Using the initial set of assortments, we generate the matrices S and U as described in Section 4.3. Following this exploratory phase, we generate assortments using two different approaches: the Objective Operational Approach, and the adapted AVW approach. In the adapted AVW Approach, at each time step t, we merely compute the sample mean of the revenue from each product, without taking pairs of products into account, as in the Objective Operational Approach. Moreover, in the adapted AVW approach, the upper bound for the revenue of each product we experiment with in any step, is given by the upper bound for the corresponding sample mean. The quantities computed at each time step t in the adapted AVW Approach are, for each product $j = 1, \cdots, N$:

$$\mu_t(j) = \frac{\sum_{k=1}^{t-1} \mathbf{1}\{j \in D_k\} p_j X_{jk}}{\sum_{k=1}^{t-1} \mathbf{1}\{j \in D_k\}}$$

$$\hat{\sigma}_t(j) = \sqrt{\frac{\sum_{k=1}^{t-1} \left(\mathbf{1}\{j \in D_k\} p_j X_{jk} - \mu_t(j)\right)^2}{\left(\sum_{k=1}^{t-1} \mathbf{1}\{j \in D_k\}\right) - 1}} \tag{4.18}$$

and

$$U_t(j) = \mu_t(j) + B\hat{\sigma}_t(j) \tag{4.19}$$

Now, let $\sigma(1), \cdots, \sigma(N)$ denote the permutation of the products corresponding to the sorted descending order of the mean utilities. For a given assortment size m, the maximum expected revenue $Z^*(m)$ at each time period is given by

$$Z^*(m) = \sum_{i=1}^{m} R_{\sigma(i)}$$

where R_i is the expected revenue corresponding to the mean utility μ_i for product i. Denoting the cumulative revenues received upto time t, using the Objective Operational Approach and the adapted AVW Approach, by S_{1t} and S_{2t} respectively, we define the cumulative regret L_{it}, for $i = 1, 2$ by

$$L_{it}(m) = t \cdot Z^*(m) - \mathbb{E}[S_{it}]$$

Similarly, we define the cumulative relative regret L_{it}^{rel}, for $i = 1, 2$ by

$$L_{it}^{rel}(m) = \frac{t \cdot Z^*(m) - \mathbb{E}[S_{it}]}{t \cdot Z^*(m)}$$

In the following, we plot the cumulative regret and relative regret functions for some representative assortment sizes, for both approaches, over a time horizon of $T = 1,000,000$ periods. Figures 4.1, 4.2, 4.3, and 4.4 show that the Objective Operational Approach dominates the adapted AVW Approach in terms of cumulative regret. Similarly, Figures 4.5, 4.6, 4.7, and 4.8 show that the Objective Operational Approach also dominates the adapted AVW Approach in terms of cumulative relative regret. Finally, Figures 4.9 and 4.10 respectively plot the cumulative regret and cumulative relative regret for several assortment sizes over the entire time period. It may be seen that the cumulative relative regret converges to zero faster as the assortment size increases.

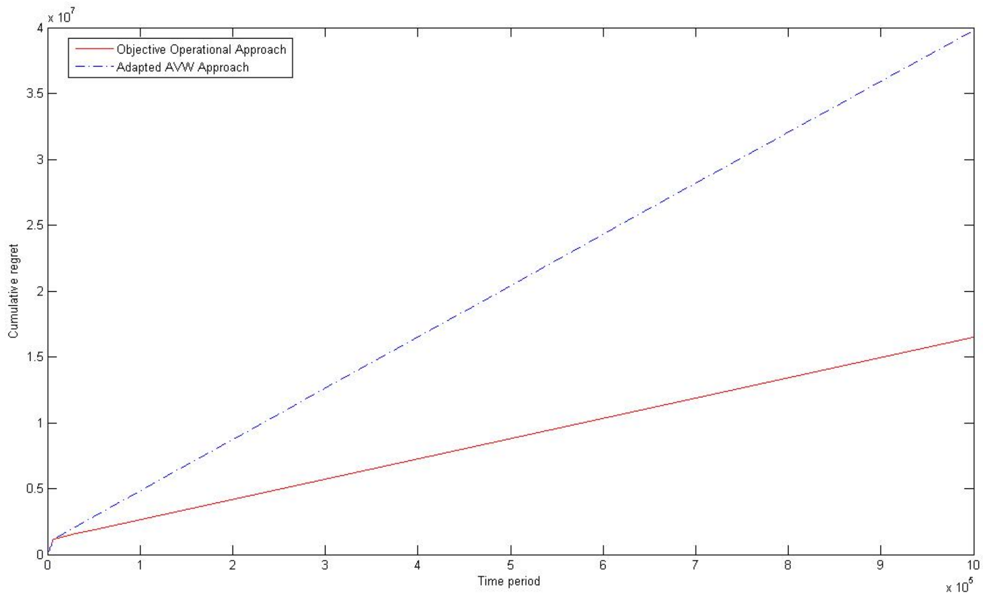

Figure 4.1: Cumulative Regret: $m = 4$

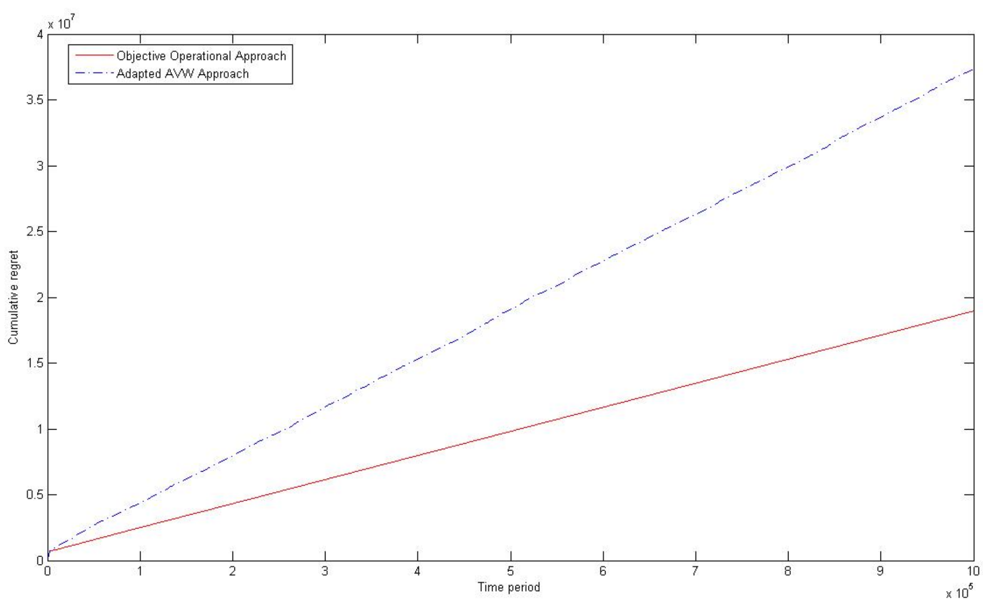

Figure 4.2: Cumulative Regret: $m = 8$

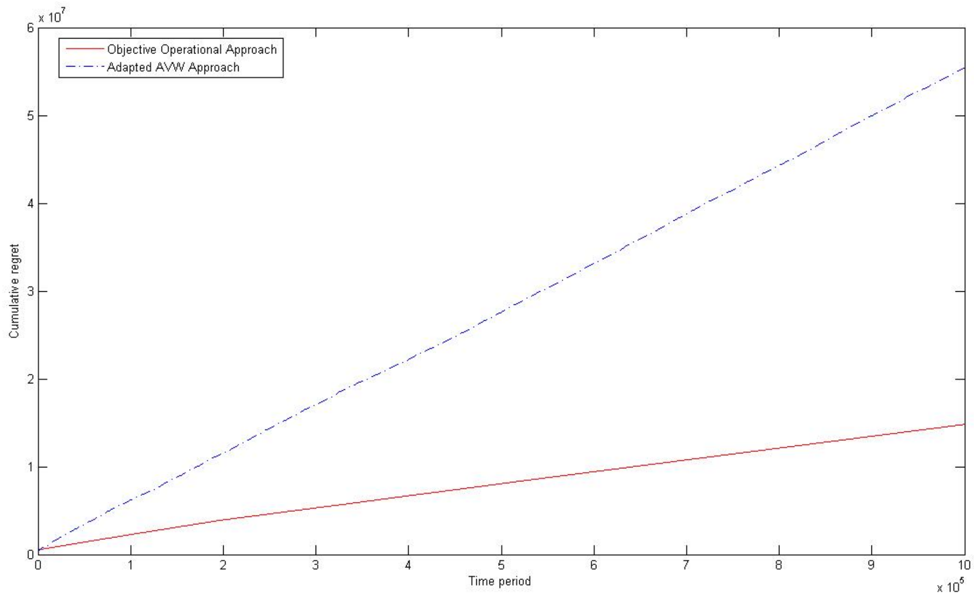

Figure 4.3: Cumulative Regret: $m = 12$

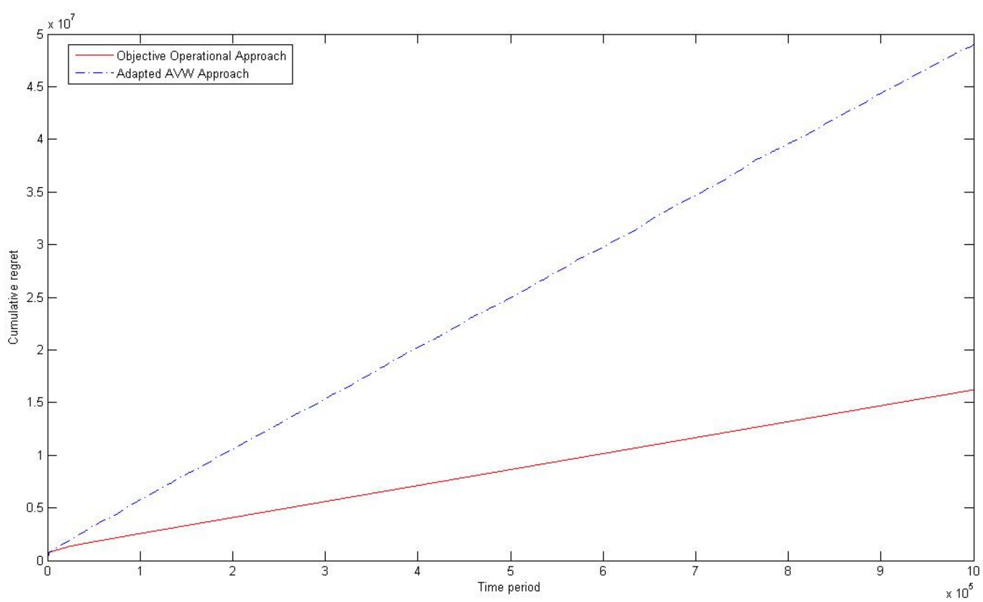

Figure 4.4: Cumulative Regret: $m = 16$

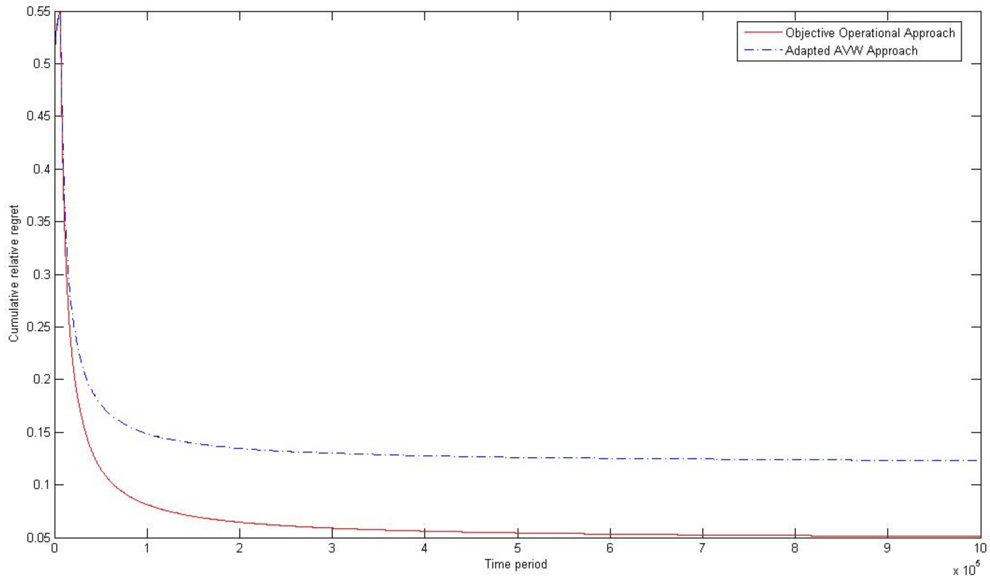

Figure 4.5: Cumulative Relative Regret: $m = 4$

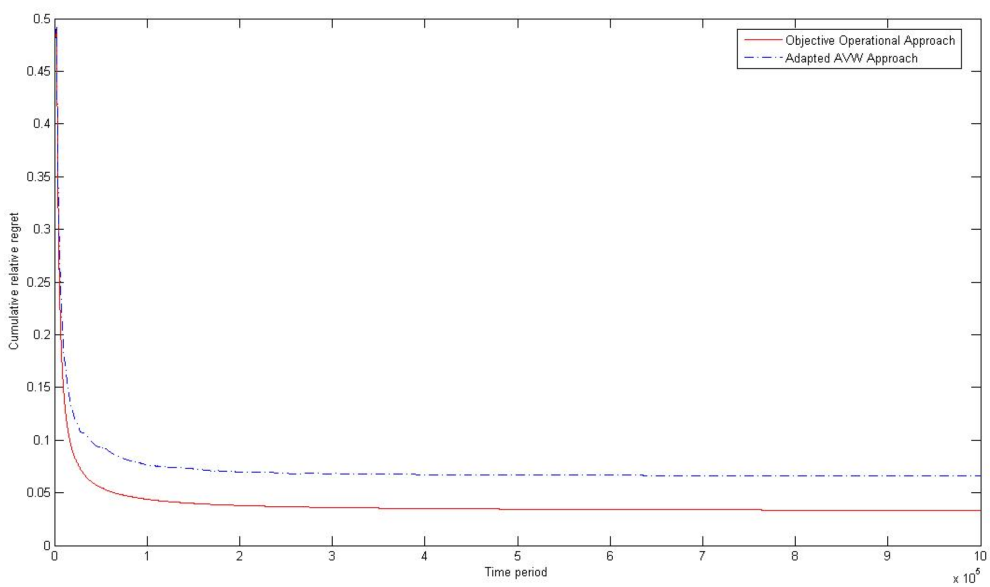

Figure 4.6: Cumulative Relative Regret: $m = 8$

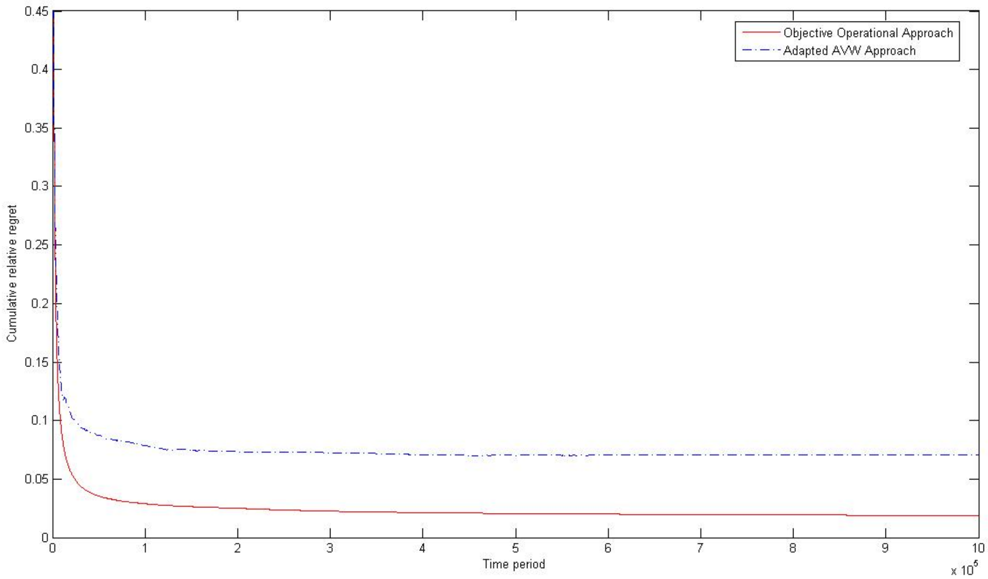

Figure 4.7: Cumulative Relative Regret: $m = 12$

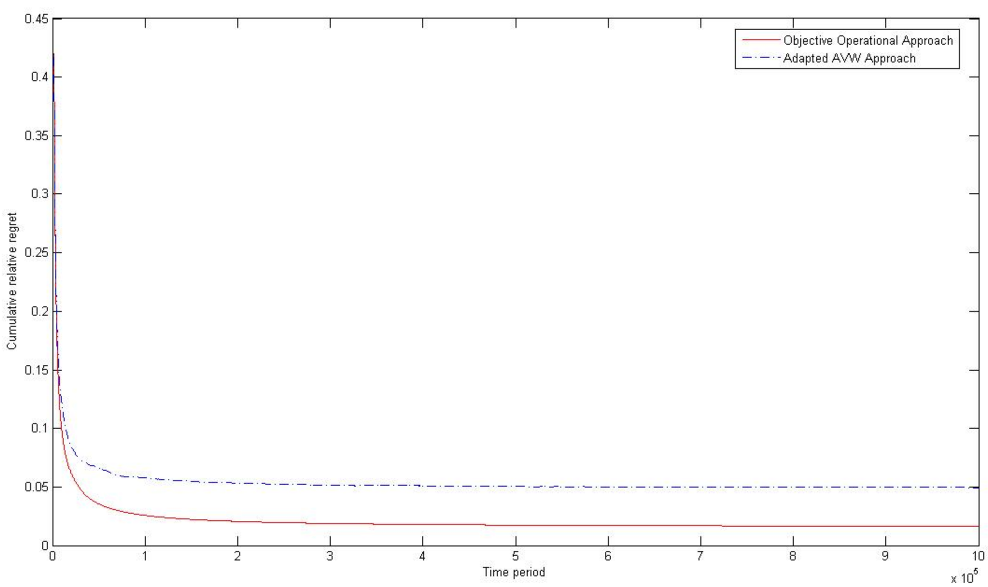

Figure 4.8: Cumulative Relative Regret: $m = 16$

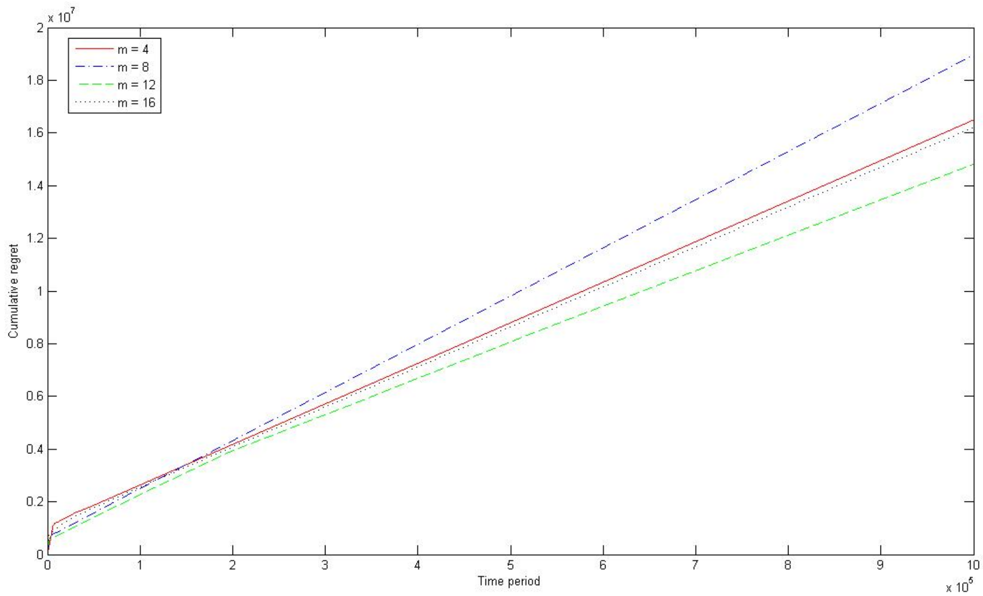

Figure 4.9: Cumulative Regret

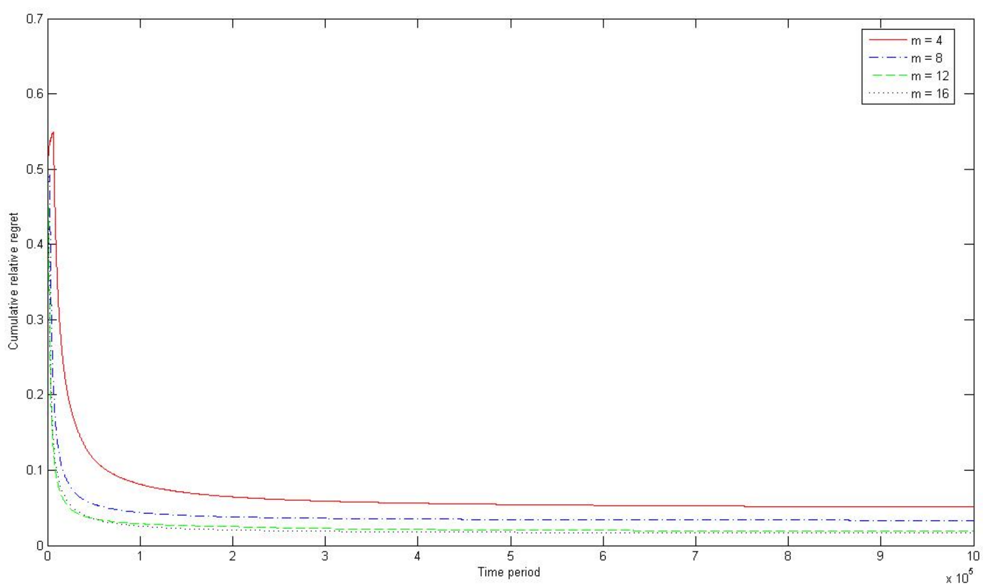

Figure 4.10: Cumulative Relative Regret

Chapter 5

Conclusion

In this thesis, we developed novel learning approaches to specific problems in inventory control, call center staffing and dynamic assortment optimization. We tested these approaches computationally, and provided strong evidence for the adoption of our general approach in tackling model uncertainty in operations management problems.

In Chapter 2, we focussed on data-driven approaches to inventory control problems which used past demand data, often limited, in order to devise efficient control policies. We considered the approach of operational statistics to inventory control. We reviewed related results which enabled us to maximize the expected profit in the single period newsvendor problem, when the demand distribution was known up to a location and scale parameter. Following that, we considered the problem posed by an unknown shape parameter. Since most distributions can be fully characterized by specifying their location, scale and shape parameters, it was desirable to obtain results analogous to those reported by Chu et al. [22] for the shape parameter. However, this proved to be difficult due to the nature of the shape parameter. Consequently, we proposed heuristics to obtain improved control policies in the newsvendor problem when the shape parameter was unknown. First, we proposed a heuristic based on operational statistics when the shape parameter was unknown. In this case, our computational study using the Pareto distribution revealed that this heuristic was an improvement over traditional approaches. In less tractable cases, such as in the case of the Gamma distribution, we proposed heuristics based on linear correction of estimates and support vector regression. In our computational study using the Gamma distribution, we found that in most cases, linear correction of estimates yielded significant improvement over traditional approaches when the value of s/c tended to 1, for small sample sizes. Moreover, support vector regression using only a rough tuning of model parameters, yielded improvement in 2 out of the 3 cases considered. This showed potential for greater improvement with better tuning of model parameters. This also suggested that other non-parametric approaches such as boosting and bagging may be profitably applied to inventory control problems. Extending the above study to multi-period models is another potential line for future examination.

In Chapter 3, we proposed an objective operational learning approach to optimal staffing in a call center. Our primary objective in this approach was making minimal assumptions about the distributions of call arrivals, customer waiting times and service times, and using empirical estimates wherever possible. Our broader goal was to estimate the objective function values for various staffing levels, and in the long run, as more data are available, to eliminate any errors introduced by using empirical estimates of parameters. We did so by probabilistically extending our data set, constructing cost estimates for unobserved staffing levels, by sampling from empirically estimated distributions for waiting times and times to abandonment. Then we proposed a novel smoothing approach to appropriately weight actual data and "extended" data to construct our objective operational estimate of cost, as a function of the one-day ahead feature forecast, and the staffing level. We considered some structural properties of the cost function and the asymptotic behaviour of the objective operational estimate of the cost function as the size of the data increased. We compared our approach to another recently published approach for optimal staffing in call centers and found improved performance in several of the cases considered. Finally, we also saw that our approach was not sensitive to the particular forecasting approach used. All of this suggested that the novel approach proposed was worthy of strong consideration for optimally staffing call centers in a practical setting.

Finally, in Chapter 4, we studied a family of stylized dynamic assortment problems in which the retailer needs to devise an assortment policy to maximize revenues over the relevant time horizon by properly adapting the offered assortment based on observed customer purchase decisions and subject to capacity constraints that limit the size of the assortment. Our main focus in this work was on the impact of learning consumer behavior via suitable assortment experimentation, and doing this in a manner that guarantees minimal revenue loss over the selling horizon.

Toward this end, we proposed an adaptive, non-parametric approach which incorporated ideas from asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays (see Anantharam et al. [3, 4]). Our approach sought to build on the basic multiarmed bandit approach by taking into account correlations and substitution effects between products in an assortment. Furthermore, we eschewed any particular parametric model for the demand for products in an assortment.

In our computational study, we used the utilities estimated in Rusmevichientong et al. [70] for 200 products, using data on DVD sales at a large online retailer, and considered assortment sizes ranging from $m = 3$ to $m = 20$. We found that our proposed approach easily outperformed the standard multiarmed bandit approach proposed by Anantharam et al. [3, 4], both in terms of cumulative regret and cumulative relative regret. Morever, while the profit computation was different in the work of Rusmevichientong et al. [70], the cumulative relative regret we obtained with our approach for identical assortment sizes compared favorably with the results they obtained. All this clearly suggested that approaches based on multiarmed bandit problems with multiple plays were worthy of serious consideration for problems in dynamic assortment optimization.

As far as future work is concerned, it would be fruitful to consider further interesting adaptations of the multiarmed bandit approach, and study their performance for the dynamic assortment planning problem. The mean utilities of products that we used for our computational study were generated assuming a Multinomial Logit Model for demand. It would be interesting to study our proposed approach on data that was generated using more general random utlity models, to study their robustness.

Our model also assumes that the cost of changing an assortment from one customer to the next is negligible. This assumption is reasonable in the online setting where the cost of changing the ads or product recommendations on the Web page is minimal. However, in settings where there are significant costs associated with switching product assortments, our model might not be appropriate. In addition, we implicitly assume that we have enough supply of each product to ignore all inventory considerations. Incorporating inventory constraints is an exciting direction for future research.

Appendix A

A.1 Proof of Theorem 1

We define

$$\tilde{\phi}(\hat{\beta}, \mathbf{z}) = \mathbb{E}[C_i^S(\mathbf{z})|\beta_i = \hat{\beta}] \tag{A.1}$$

We have

$$|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \leq |\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \tilde{\phi}(\hat{\beta}, \mathbf{z})| + |\tilde{\phi}(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \tag{A.2}$$

Taking expectations,

$$\mathbb{E}|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \leq \mathbb{E}|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \tilde{\phi}(\hat{\beta}, \mathbf{z})| + \mathbb{E}|\tilde{\phi}(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \tag{A.3}$$

We now state without proof, two lemmas. For the correponding general lemmas and their proofs, the reader is referred to Devroye [26] (Lemma 2.1 and Lemma 2.2). In what follows, we use the following notation:

$$w_{ni}(x) = \frac{K_1(X_i, x, h_1)}{\sum_{j=1}^{n} K_1(X_j, x, h_1)} \tag{A.4}$$

Lemma 1. *Suppose $f \in L^1(\mu)$, that is, $\int |f(x)|\mu(dx) < \infty$, where μ is a probability measure over x. Also suppose assumptions A4 and A5 hold. Then*

$$\mathbb{E}\left[\sum_{i=1}^{n} w_{ni(x)} \left|f(X_i) - f(x)\right|\right] \to 0 \text{ as } n \to \infty \tag{A.5}$$

In what follows, we will use the symbol S_r for the closed ball of radius r centered at x.

Lemma 2. *Let $h_1(n)$ be a sequence of positive numbers satisfying assumptions A4. For all $c > 0$, we have*

$$n\mu\left(S_{\left(ch_1^{-1/2}\right)}\right) \to \infty \text{ as } n \to \infty, \text{ almost all } x(\mu) \tag{A.6}$$

where μ is a probability measure over x.

Now, by Minkowski's inequality, for $p = 1$, we have

$$\mathbb{E}\left|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \tilde{\phi}(\hat{\beta}, \mathbf{z})\right| \leq \mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(C_i^S(\mathbf{z}) - \tilde{\phi}(\beta_i, \mathbf{z}))\right| + \mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(\tilde{\phi}(\beta_i, \mathbf{z}) - \tilde{\phi}(\hat{\beta}, \mathbf{z}))\right| \tag{A.7}$$

Now, there is a possibility that $w_{ni}(\hat{\beta}) = 0$ for all i; in that case, a third term should be added on the RHS of (A.7), namely $|\tilde{\phi}(\hat{\beta}, \mathbf{z})| P\left(\sum_{i=1}^{n} w_{ni}(\hat{\beta}) = 0\right)$. Now, by Lemma 2, we have

$$P\left(\sum_{i=1}^{n} w_{ni}(\hat{\beta}) = 0\right) = \left(1 - \mu\left(S_{\left(ch_1^{-1/2}\right)}\right)\right)^n \leq \exp\left\{-n\mu\left(S_{\left(ch_1^{-1/2}\right)}\right)\right\} \to 0 \text{ as } n \to \infty \tag{A.8}$$

for almost all $\hat{\beta}(\mu)$, where μ is the probability measure of $\hat{\beta}$. Furthermore, by assumption A1, $|\tilde{\phi}(\hat{\beta}, \mathbf{z})|$ is finite, and hence the third term goes to 0 for almost all $\hat{\beta}(\mu)$ as $n \to \infty$. Similarly, the second term on the RHS of (A.7) tends to 0 as $n \to \infty$ for almost all $\hat{\beta}(\mu)$ by Lemma 1. It now remains to show that the first term on the RHS of (A.7) tends to 0 for almost all $\hat{\beta}(\mu)$ as $n \to \infty$. Define for integer $t > 0$, $Z_i' = C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}1\{C_i^S(\mathbf{z}) \leq t\}$, $Z_i'' = C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - Z_i'$, $\tilde{\phi}'(\hat{\beta}, \mathbf{z}) = \mathbb{E}[Z_i']$, $\tilde{\phi}''(\hat{\beta}, \mathbf{z}) = \mathbb{E}[Z_i'']$. Thus,

$$\mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(C_i^S(\mathbf{z}) - \tilde{\phi}(\beta_i, \mathbf{z}))\right| \leq \mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(Z_i' - \tilde{\phi}'(\beta_i, \mathbf{z}))\right| + \mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(Z_i'' - \tilde{\phi}''(\beta_i, \mathbf{z}))\right| \tag{A.9}$$

We have

$$\mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(Z_i'' - \tilde{\phi}''(\beta_i, \mathbf{z}))\right| \leq 2\mathbb{E}\left(\sum_{i=1}^{n} w_{ni}(\hat{\beta})|Z_i''|\right) = 2\mathbb{E}\left(\sum_{i=1}^{n} w_{ni}(\hat{\beta}) f_t(\beta_i)\right) \tag{A.10}$$

where $f_t(\hat{\beta}) = \mathbb{E}[|Z_i''|]$. Let G_t be the set of all $\hat{\beta}$ for which the first term on the RHS of (A.9) tends to 0 and $\mathbb{E}\left(\sum_{i=1}^{n} w_{ni}(\hat{\beta}) f_t(\beta_i)\right) \to f_t(\hat{\beta})$ as $n \to \infty$. It follows, by a straightforward application of Lemma 1, that for each fixed t, $\mu(G_t) = 1$. Let H be the set of all $\hat{\beta}$ such that $f_t(\hat{\beta}) \to 0$ as $t \to \infty$. Clearly, $\mu(H) = 1$, since $\mathbb{E}[f_t(\beta)] \to 0$ as $t \to 0$ and f_t is monotone in t. For all $\hat{\beta}$ in $H \cap (\cap_t G_t)$, we claim that the RHS of (A.9) tends to 0: first pick t large enough so that $f_t(\hat{\beta})$ is small, and then let n grow large. Since this set has μ-measure 1, it follows that

$$\mathbb{E}\left|\sum_{i=1}^{n} w_{ni}(\hat{\beta})(C_i^S(\mathbf{z}) - \tilde{\phi}(\beta_i, \mathbf{z}))\right| \to 0 \tag{A.11}$$

for almost all $\hat{\beta}(\mu)$ as $n \to \infty$. Hence, we have now shown that all the terms on the RHS of (A.7) tend to 0 for almost all $\hat{\beta}(\mu)$ as $n \to \infty$. It then follows that

$$\mathbb{E}\left|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \tilde{\phi}(\hat{\beta}, \mathbf{z})\right| \to 0 \tag{A.12}$$

for almost all $\hat{\beta}(\mu)$ as $n \to \infty$. Now, we have that

$$\mathbb{E}|\tilde{\phi}(\hat{\beta},\mathbf{z})-\phi(\hat{\beta},\mathbf{z})| \leq \mathbb{E}\left|\tilde{\phi}(\hat{\beta},\mathbf{z}) - \frac{\sum_{i=1}^n C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}}{\sum_{i=1}^n 1\{\beta_i = \hat{\beta}\}}\right| + \mathbb{E}\left|\frac{\sum_{i=1}^n C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}}{\sum_{i=1}^n 1\{\beta_i = \hat{\beta}\}} - \phi(\hat{\beta},\mathbf{z})\right| \tag{A.13}$$

Letting

$$C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} = \eta(\beta_i, \{a_{ik}\}) \tag{A.14}$$

in assumption A2, it easily follows from Equation (A.1) that

$$\mathbb{E}\left|\tilde{\phi}(\hat{\beta},\mathbf{z}) - \frac{\sum_{i=1}^n C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}}{\sum_{i=1}^n 1\{\beta_i = \hat{\beta}\}}\right| \to 0 \tag{A.15}$$

for almost all $(\hat{\beta},\mathbf{z})$ as $n \to \infty$. We now define $B_n = \{i \leq n : \beta_i = \hat{\beta}\}$ and we assume that $|B_n| > 0$. We also define $D_n = \{i \in B_n : \mathbf{z} \neq \mathbf{y}_i\}$. By assumption A6, it follows that

$$\frac{\sum_{i=1}^n C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}}{\sum_{i=1}^n 1\{\beta_i = \hat{\beta}\}} \to \frac{\left(\sum_{i\in D_n^c} C_i^C(\mathbf{z}) + \sum_{i \in D_n} \hat{C}_i^S(\mathbf{z})\right)1\{\beta_i = \hat{\beta}\}}{|B_n|} \tag{A.16}$$

Hence, as $n \to \infty$, the second term on the RHS of Equation (A.13) tends to

$$\mathbb{E}\left|\frac{|D_n^c|}{|B_n|}(C_i^C(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})) + \frac{|D_n|}{|B_n|}(\hat{C}_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z}))\right|$$

$$\leq \frac{|D_n^c|}{|B_n|}\mathbb{E}\left|C_i^C(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| + \frac{|D_n|}{|B_n|}\mathbb{E}\left|\hat{C}_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| \tag{A.17}$$

Clearly, by Lemma 1, we have that

$$\mathbb{E}\left|C_i^C(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| \to 0 \tag{A.18}$$

for almost all $(\hat{\beta},\mathbf{z})$ as $n \to \infty$. Moreover, we have that

$$\mathbb{E}\left|\hat{C}_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| \leq \mathbb{E}\left|\hat{C}_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - C_i^C(\mathbf{z})1\{\beta_i = \hat{\beta}\}\right| + \mathbb{E}\left|C_i^C(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| \tag{A.19}$$

By assumption A7, the first term on the RHS of Equation (A.19) tends to 0 for almost all $(\hat{\beta},\mathbf{z})$ as $n \to \infty$, while the second term tends to 0 by Lemma 1. Hence, it follows that

$$\mathbb{E}\left|\hat{C}_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\} - \phi(\hat{\beta},\mathbf{z})\right| \to 0 \tag{A.20}$$

for almost all $(\hat{\beta},\mathbf{z})$ as $n \to \infty$. Hence, by Equations (A.17), (A.18), and (A.20), it follows that

$$\mathbb{E}\left|\frac{\sum_{i=1}^n C_i^S(\mathbf{z})1\{\beta_i = \hat{\beta}\}}{\sum_{i=1}^n 1\{\beta_i = \hat{\beta}\}} - \phi(\hat{\beta},\mathbf{z})\right| \to 0 \tag{A.21}$$

for almost all $(\hat{\beta}, \mathbf{z})$ as $n \to \infty$. From Equations (A.13), (A.15), and (A.21), it follows that

$$\mathbb{E}|\tilde{\phi}(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \to 0 \tag{A.22}$$

for almost all $(\hat{\beta}, \mathbf{z})$ as $n \to \infty$. From Equations (A.3), (A.12), and (A.22), we have

$$\mathbb{E}|\hat{\phi}_n(\hat{\beta}, \mathbf{z}) - \phi(\hat{\beta}, \mathbf{z})| \to 0 \tag{A.23}$$

for almost all $(\hat{\beta}, \mathbf{z})$ as $n \to \infty$, and we are done.

Appendix B

B.1 Proof of Theorem 6

The proof consists of three main steps, which we now summarize. First, define $0 \leq l \leq m-1$ and $m \leq n \leq N$ by

$$\mu_1 \geq \cdots \geq \mu_l > \mu_{l+1} = \cdots = \mu_m = \cdots = \mu_n > \mu_{n+1} \geq \cdots \geq \mu_N$$

Throughout the proof, fix $\epsilon > 0$, satisfying $\epsilon < \mu_1 - \mu_m/2$ if $l > 0$ and $\epsilon < \mu_n - \mu_{n+1}/2$ if $n < N$.

- **Step A:** This step is required only if $l > 0$. We need to show

$$\mu_j \geq \mu_l \Longrightarrow \mathbb{E}[t - T_t(j)] = o(\log t)$$

- **Step B:** This step is required only if $n < N$. Define the increasing sequence of integer-valued random variables B_t by

$$B_t = \#\{N_E(m) \leq a \leq t | \text{ for some } j \geq n+1, j \text{ is one of the } m\text{-leaders at stage } a+1\}$$

  where $\#\{\}$ denotes the number of elements in $\{\}$. Here, the m-leaders are the products chosen by the “Best m” procedure. Then, $\mathbb{E}[B_t] = o(\log t)$.

- **Step C:** This step is required only if $n < N$. For each $j \geq n+1$ define the increasing sequence of integer-valued random variables $S_t(j)$ by

$$S_t(j) = \#\{N_E(m) \leq a \leq t | \text{Condition } S_t(j) \text{ holds}\}$$

  where Condition $S_t(j)$:- All the m-leaders at stage $a+1$ are among the arms k with $\mu_k \geq \mu_n$ and for each m-leader at stage $a+1$, $|V_a(k) - \mu_k| < \epsilon$, but still the rule plays arm j at stage $a+1$. Then, for each $\rho > 0$ we can choose $\epsilon > 0$ small enough so that

$$\mathbb{E}[S_t(j)] \leq (1 + \rho + o(1))\mathcal{K}_{jm} \log t$$

We now prove the individual steps.

Proof of Step A: This step is required only if $l > 0$. Pick a positive integer c, satisfying $c > (1 - N^2\delta)^{-1}$. This choice of c implies that

$$\frac{t - c^{r-1}}{N} > N\delta t \text{ for } t > c^r$$

Lemma 3. *Let r be a positive integer. Define the sets*

$$A_r = \bigcap_{1 \le j \le N} \{ \max_{\delta c^{r-1} \le t \le c^{r+1}} |V_t(j) - \mu_j| \le \epsilon \},$$

$$B_r = \bigcap_{k \le l} \{ U_a(k) \ge \mu_l - \epsilon \text{ for } 1 \le a \le \delta t \text{ and } c^{r-1} \le t \le c^{r+1} \}.$$

Then $P_C(A_r^c) = o(c^{-r})$ and $P_C(B_r^c) = o(c^{-r})$ where A_r^c and B_r^c denote the complements of A_r and B_r, respectively.

Proof: From Assumption A2, we immediately get $P_C(A_r^c) = o(c^{-r})$. From Assumption A1, we see that $P_C(B_r^c) = O(c^{-r}r^{-p}) = o(c^{-r})$.

Lemma 4. *On the event $A_r \cap B_r$, if $t + 1 \equiv k \mod N$ for some $k \le l$ and $c^{r-1} \le t \le c^{r+1}$, the objective operational approach selects product k.*

Proof: On A_r, the $V_a(\cdot)$ statistics of the m-leaders are all within ϵ of the means $\mu_{(\cdot)}$. If product k is one of the m-leaders at stage $t+1$, then according to the approach, it is included in the assortment. Suppose product k is not an m-leader at stage $t + 1$. On A_r the least best of the m-leaders at stage $t + 1$, say j_t, has

$$V_t(j_t) < \mu_l - \epsilon$$

In case $T_t(k) \ge \delta t$, we have on A_r,

$$\mu_l - \epsilon \le V_t(k)$$

hence, the approach will include product k in the assortment since it will already be one of the m-leaders at stage $t + 1$. In case $T_t(k) < \delta t$, we have on B_r,

$$\mu_l - \epsilon \le U_t(k)$$

so in any case, product k will be selected.

By Lemma 4, on the event $A_r \cap B_r$, for $c^r \le t \le c^{r+1}$, the number of times product k has been selected, for $k \le l$, exceeds

$$N^{-1}(t - c^{r-1} - 2N)$$

which exceeds $N\delta t$ if $r \ge r_0$ for some r_0.

Lemma 5. *If $r \geq r_0$, then on the event $A_r \cap B_r$, for every $c^r \leq t \leq c^{r+1}$, the objective operational approach selects product k, where $k \leq l$.*

Proof: By Lemma 4, on $A_r \cap B_r$, and $c^r \leq t \leq c^{r+1}$, $r \geq r_0$, all products k, $k \leq l$, are well sampled. Since on A_r, every well-sampled product has its $V_a(\cdot)$ statistic ϵ close to $\mu_{(\cdot)}$, all products k, $k \leq l$ must be among the m-leaders. Further, they cannot be replaced by a nonleading arm because none of them is the least best of the m-leaders.

It follows that for $r \geq r_0$, the expected number of times product k, $k \leq l$, is not played during $c^r \leq t \leq c^{r+1}$ is less than

$$\sum_{c^r \leq t \leq c^{r+1}} P_C(A_r^c) + P_C(B_r^c) = o(1)$$

Hence, the expected number of times product k, $k \leq l$, is not selected in t steps is $o(\log t)$.

Proof of Step B: This step is required only if $n < N$. The proof is identical in form to that of Step A and proceeds as follows.

Lemma 6. *Let A_r be as in Lemma 3 and let*

$$Z_r = \bigcap_{k \leq n} \{U_a(k) \geq \mu_k - \epsilon \text{ for } 1 \leq a \leq \delta t \text{ and } c^{r-1} \leq t \leq c^{r+1}\}.$$

Then $P_C(A_r^c) = o(c^{-r})$ and $P_C(Z_r^c) = o(c^{-r})$.

Proof: The proof is identical to the proof of Lemma 3.

Lemma 7. *On the event $A_r \cap Z_r$, if $t+1 \equiv k \mod N$ for some $k \leq n$ and $c^{r-1} \leq t \leq c^{r+1}$, the objective operational approach only selects a product with index $\leq n$ at stage $t+1$.*

Proof: Suppose not. Then k is not one of the m-leaders and the least best of the m-leaders has index $j_t > n$ on the event A_r with $V_t(j_t) < \mu_n - \epsilon$. If $T_t(k) \geq \delta t$,

$$\mu_n - \epsilon \leq V_t(k)$$

on A_r, hence our approach will select product k; in fact product k will already be one of the m-leaders at stage $t+1$. If $T_t(k) < \delta t$,

$$\mu_n - \epsilon \leq U_t(k)$$

on Z_r, hence, our approach will select product k.

Let r_0 be defined as in the proof of Step A. We now show that on $A_r \cap Z_r$, for $r \geq r_0 + 1$ and $c^{r-1} \leq t \leq c^{r+1}$, $m - l$ of the m-border products have been selected δt times. Here, m-border follows the definition given in Section 4.3.

(1) First consider the case $n = m$. For each of the m-border products j with indices $l+1 \leq j \leq n$, there are at least $t - c^{r-1} - 2N/N > N\delta t$ times prior to t at which $t+1 \equiv j \mod N$. Choose δt of these times. By Lemma 7, on the event $A_r \cap Z_r$, each of the products that is selected at this time has index $\leq m$. But this means that the product j is selected at this time. Thus, we see that at stage $t+1$, all m-border products are well sampled, and there are $m - l$ of them.

(2) Suppose $n > m$ and that fewer than $m - l$ of the m-border products have been well sampled. Let j be one of the products that is not well sampled, $l+1 \leq j \leq n$. There are at least $t - c^{r-1} - 2N/N > N\delta t$ times prior to t at which $t+1 \equiv j \mod N$. Choose $N\delta t$ of these times. Since product j is not well sampled, we can choose $(N-1)\delta t$ of these times at which the approach selects only products whose indices are $\geq n$, by Lemma 7 above. We know by Lemma 5 that at each of these times the approach selects all products whose indices are $\leq l$ on the event $A_r \cap B_r$, which contains the event $A_r \cap Z_r$. Thus, $(m-l)(N-1)\delta t$ selections of m-border products with index $\neq j$ are made at these times. Note that there are $n - l - 1 \geq m - l$ such arms. Also note that at these $(N-1)\delta t$ times, not one of these products can undergo more than $(N-1)\delta t$ selections. Suppose that only $p < m - l$ of these $n - l - 1$ products undergo δt selections or more at these times. Then the total number of selctions of these products at these times is strictly less than

$$p(N-1)\delta t + (n - l - 1 - p)\delta t \leq (m - l)(N-1)\delta t$$

which gives a contradiction.

The analog of Lemma 5 is as follows.

Lemma 8. *If $r \geq r_0 + 1$, then on the event $A_r \cap Z_r$, for every $c^r \leq t \leq c^{r+1}$, the m-leaders are among the products k, $k \leq n$.*

Proof: On A_r, a well-sampled product has its $V_a(\cdot)$ statistic ϵ close to $\mu_{(\cdot)}$. By the above reasoning, at least m of the $k, k \leq n$, are well sampled at stage $t+1$, hence the m-leaders are constituted of such products.

Step B now follows from Lemmas 6 and 8.

Proof of Step C: This step is again required only if $n < N$. Let $j \geq n+1$. Then observe that

$$S_t(j) \leq \#\{N_E(m) \leq a \leq t | U_a(j) \geq \mu_m - \epsilon\}$$

Taking expectations,

$$\mathbb{E}_C[S_t(j)] \leq \mathbb{E}_C[\#\{N_E(m) \leq a \leq t | U_a(j) \geq \mu_m - \epsilon\}]$$

$$\Rightarrow \mathbb{E}_C[S_t(j)] \leq \sum_{a=1}^{t} P\{U_a(j) \geq \mu_m - \epsilon\}$$

But by Assumption A3 we can, for each $\rho > 0$, choose ϵ small enough so that

$$\sum_{a=1}^{t} P\{U_a(j) \geq \mu_m - \epsilon\} \leq (1 + \rho + o(1))\mathcal{K}_{jm} \log t$$

which establishes Step C. Hence Theorem 6 is proved.

Bibliography

[1] Rajeev Agrawal. “Sample Mean Based Index Policies with O(log n) Regret for the Multi Armed Bandit Problem”. In: *Advances in Applied Probability* 27.4 (1995), pp. 1054–1078.

[2] Zeynep Aksin, Mor Armony, and Vijay Mehrotra. “The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research”. In: *Production and Operations Management* 16.6 (2007), pp. 665–688.

[3] V. Anantharam, P. Varaiya, and J. Walrand. “Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-Part I: I.I.D. rewards”. In: *Automatic Control, IEEE Transactions on* 32.11 (1987), pp. 968–976.

[4] V. Anantharam, P. Varaiya, and J. Walrand. “Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-Part II: Markovian rewards”. In: *Automatic Control, IEEE Transactions on* 32.11 (1987), pp. 977–982.

[5] Mor Armony and Itai Gurvich. “When Promotions Meet Operations: Cross-Selling and Its Effect on Call Center Performance”. In: *Manufacturing and Service Operations Management* 12.3 (2010), pp. 470–488.

[6] Mor Armony, Erica Plambeck, and Sridhar Seshadri. “Sensitivity of Optimal Capacity to Customer Impatience in an Unobservable M/M/S Queue (Why You Shouldn’t Shout at the DMV)”. In: *Manufacturing and Service Operations Management* 11.1 (2009), pp. 19–32.

[7] Kenneth J. Arrow, Theodore Harris, and Jacob Marschak. “Optimal Inventory Policy”. In: *Econometrica* 19.3 (1951), pp. 250–272.

[8] P. Auer et al. “Gambling in a rigged casino: The adversarial multi-armed bandit problem”. In: 1995, pp. 322–331.

[9] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. “Finite-time Analysis of the Multiarmed Bandit Problem”. In: *Machine Learning* 47.2 (2002), pp. 235–256.

[10] Athanassios N. Avramidis, Alexandre Deslauriers, and Pierre L’Ecuyer. “Modeling Daily Arrivals to a Telephone Call Center”. In: *Management Science* 50.7 (2004), pp. 896–908.

[11] Katy S. Azoury. “Bayes Solution to Dynamic Inventory Models under Unknown Demand Distribution”. In: *Management Science* 31.9 (1985), pp. 1150–1160.

[12] Achal Bassamboo and Assaf Zeevi. "On a Data-Driven Method for Staffing Large Call Centers". In: *Operations Research* 57.3 (2009), pp. 714–726.

[13] Dimitris Bertsimas and Xuan Vinh Doan. "Robust and data-driven approaches to call centers". In: *European Journal of Operational Research* 207.2 (2010), pp. 1072–1085.

[14] Dimitris Bertsimas and Adam J. Mersereau. "A Learning Approach for Interactive Marketing to a Customer Segment". In: *Operations Research* 55.6 (2007), pp. 1120–1135.

[15] Dimitris Bertsimas and Aurélie Thiele. "A Robust Optimization Approach to Inventory Theory". In: *Operations Research* 54.1 (2006), pp. 150–168.

[16] Michel Bierlaire. "BIOGEME: A free package for the estimation of discrete choice models". In: *Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland*. 2003.

[17] Sem Borst, Avi Mandelbaum, and Martin I. Reiman. "Dimensioning Large Call Centers". In: *Operations Research* 52.1 (2004), pp. 17–34.

[18] David J. Braden and Marshall Freimer. "Informational Dynamics of Censored Observations". In: *Management Science* 37.11 (1991), pp. 1390–1404.

[19] Monica Brezzi and Tze Leung Lai. "Optimal learning and experimentation in bandit problems". In: *Journal of Economic Dynamics and Control* 27.1 (2002), pp. 87–108.

[20] Lawrence Brown et al. "Statistical Analysis of a Telephone Call Center". In: *Journal of the American Statistical Association* 100.469 (2005), pp. 36–50.

[21] Felipe Caro and Jérémie Gallien. "Dynamic Assortment with Demand Learning for Seasonal Consumer Goods". In: *Management Science* 53.2 (2007), pp. 276–292.

[22] Leon Yang Chu, J.George Shanthikumar, and Zuo-Jun Max Shen. "Solving operational statistics via a Bayesian analysis". In: *Operations Research Letters* 36.1 (2008), pp. 110–116.

[23] Erhan Çinlar. *Introduction to Stochastic Processes*. New Jersey: Prentice-Hall, 1975.

[24] S. A. Conrad. "Sales Data and the Estimation of Demand". In: *Operations Research Quarterly* 27.1 (1976), pp. 123–127.

[25] Thomas H. Cormen et al. *Introduction to Algorithms*. Cambridge, MA: MIT Press, 2009.

[26] Luc Devroye. "On the Almost Everywhere Convergence of Nonparametric Regression Function Estimates". In: *The Annals of Statistics* 9.6 (1981), pp. 1310–1319.

[27] Xiaomei Ding, Martin L. Puterman, and Arnab Bisi. "The Censored Newsvendor and the Optimal Acquisition of Information". In: *Operations Research* 50.3 (2002), pp. 517–527.

[28] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. "The Inventory Problem: I. Case of Known Distributions of Demand". In: *Econometrica* 20.2 (1952), pp. 187–222.

[29] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. “The Inventory Problem: II. Case of Unknown Distributions of Demand”. In: *Econometrica* 20.3 (1952), pp. 450–466.

[30] Carl Eckart and Gale Young. “The approximation of one matrix by another of lower rank”. In: *Psychometrika* 1.3 (1936), pp. 211–218.

[31] Guillermo Gallego and Ilkyeong Moon. “The Distribution Free Newsboy Problem: Review and Extensions”. In: *The Journal of the Operational Research Society* 44.8 (1993), pp. 825–834.

[32] Noah Gans, Ger Koole, and Avishai Mandelbaum. “Telephone Call Centers: Tutorial, Review, and Research Prospects”. In: *Manufacturing and Service Operations Management* 5.2 (2003), pp. 79–141.

[33] Noah Gans et al. *Parametric Stochastic Programming Models for Call-Center Workforce Scheduling*. Working Paper. University of Washington Foster School of Business, 2009.

[34] O. Garnet, A. Mandelbaum, and M. Reiman. “Designing a Call Center with Impatient Customers”. In: *Manufacturing and Service Operations Management* 4.3 (2002), pp. 208–227.

[35] Vishal Gaur and Dorothée Honhon. “Assortment Planning and Inventory Decisions under a Locational Choice Model”. In: *Management Science* 52.10 (2006), pp. 1528–1543.

[36] Gregory A. Godfrey and Warren B. Powell. “An Adaptive, Distribution-Free Algorithm for the Newsvendor Problem with Censored Demands, with Applications to Inventory and Distribution”. In: *Management Science* 47.8 (2001), pp. 1101–1112.

[37] Vineet Goyal, Retsef Levi, and Danny Segev. “Near-Optimal Algorithms for the Assortment Planning Problem under Dynamic Substitution and Stochastic Demand”. Submitted to *Operations Research*. 2009.

[38] Itai Gurvich, James Luedtke, and Tolga Tezcan. “Staffing Call Centers with Uncertain Demand Forecasts: A Chance-Constrained Optimization Approach”. In: *Management Science* 56.7 (2010), pp. 1093–1115.

[39] Itay Gurvich, Mor Armony, and Constantinos Maglaras. “Cross-Selling in a Call Center with a Heterogeneous Customer Population”. In: *Operations Research* 57.2 (2009), pp. 299–313.

[40] Itay Gurvich, Mor Armony, and Avishai Mandelbaum. “Service-Level Differentiation in Call Centers with Fully Flexible Servers”. In: *Management Science* 54.2 (2008), pp. 279–294.

[41] Shlomo Halfin and Ward Whitt. “Heavy-Traffic Limits for Queues with Many Exponential Servers”. In: *Operations Research* 29.3 (1981), pp. 567–588.

[42] J. Michael Harrison and Assaf Zeevi. "A Method for Staffing Large Call Centers Based on Stochastic Fluid Models". In: *Manufacturing and Service Operations Management* 7.1 (2005), pp. 20–36.

[43] R. H. Hayes. "Statistical Estimation Problems in Inventory Control". In: *Management Science* 15.11 (1969), pp. 686–701.

[44] Dorothée Honhon, Vishal Gaur, and Sridhar Seshadri. "Assortment Planning and Inventory Decisions Under Stockout-Based Substitution". In: *Operations Research* 58.5 (2010), pp. 1364–1379.

[45] Woonghee Tim Huh and Paat Rusmevichientong. "A Nonparametric Asymptotic Analysis of Inventory Planning with Censored Demand". In: *Mathematics of Operations Research* 34.1 (2009), pp. 103–123.

[46] Donald L. Iglehart. "The Dynamic Inventory Problem with Unknown Demand Distribution". In: *Management Science* 10.3 (1964), pp. 429–440.

[47] Srikanth Jagabathula, Vivek F. Farias, and Devavrat Shah. "A Nonparametric Approach to Modeling Choice with Limited Data". Submitted. 2010.

[48] Elleke Janssen, Leo Strijbosch, and Ruud Brekelmans. "Assessing the effects of using demand parameters estimates in inventory control and improving the performance using a correction function". In: *International Journal of Production Economics* 118.1 (2009), pp. 34–42.

[49] I. T. Jolliffe. *Principal Component Analysis*. New York: Springer-Verlag, 2002.

[50] Samuel Karlin. "Dynamic Inventory Policy with Varying Stochastic Demands". In: *Management Science* 6.3 (1960), pp. 231–258.

[51] Robert E. Kass and Larry Wasserman. "The Selection of Prior Distributions by Formal Rules". In: *Journal of the American Statistical Association* 91.435 (1996), pp. 1343–1370.

[52] Mark Keaton. "Using the Gamma Distribution to Model Demand when Lead Time is Random". In: *Journal of Business Logistics* 16.1 (1995), pp. 107–131.

[53] A. Gürhan Kök and Marshall L. Fisher. "Demand Estimation and Assortment Optimization Under Substitution: Methodology and Application". In: *Operations Research* 55.6 (2007), pp. 1001–1021.

[54] A. Gürhan Kök, Marshall L. Fisher, and Ramnath Vaidyanathan. "Assortment planning: Review of literature and industry practice". In: *Retail Supply Chain Management: Quantitative Models and Empirical Studies*. Springer, 2008.

[55] T.L Lai and Herbert Robbins. "Asymptotically efficient adaptive allocation rules". In: *Advances in Applied Mathematics* 6.1 (1985), pp. 4–22.

[56] Tze Leung Lai. "Adaptive Treatment Allocation and the Multi-Armed Bandit Problem". In: *The Annals of Statistics* 15.3 (1987), pp. 1091–1114.

[57] Martin A. Lariviere and Evan L. Porteus. "Stalking Information: Bayesian Inventory Management with Unobserved Lost Sales". In: *Management Science* 45.3 (1999), pp. 346–363.

[58] Retsef Levi, Robin O. Roundy, and David B. Shmoys. "Provably Near Optimal Sampling Based Policies for Stochastic Inventory Control Models". In: *Mathematics of Operations Research* 32.4 (2007), pp. 821–839.

[59] Andrew E. B. Lim, J. George Shanthikumar, and Zuo-Jun Max Shen. "Model Uncertainty, Robust Optimization and Learning". In: *Tutorials in Operations Research: Models, Methods, and Applications for Innovative Decision Making.* M.P. Johnson, B. Norman, N. Secomandi (Eds.), 2006.

[60] Liwan H. Liyanage and J.George Shanthikumar. "A practical inventory control policy using operational statistics". In: *Operations Research Letters* 33 (2005), pp. 341–348.

[61] Siddharth Mahajan and Garrett van Ryzin. "Stocking Retail Assortments under Dynamic Consumer Substitution". In: *Operations Research* 49.3 (2001), pp. 334–351.

[62] Avishai Mandelbaum and Sergey Zeltyn. "Staffing Many-Server Queues with Impatient Customers: Constraint Satisfaction in Call Centers". In: *Operations Research* 57.5 (2009), pp. 1189–1205.

[63] Daniel McFadden. "Econometric Models for Probabilistic Choice Among Products". In: *The Journal of Business* 53.3 (1980), S13–S29.

[64] Vijay Mehrotra, Ozgür Ozlük, and Robert Saltzman. "Intelligent Procedures for Intra-Day Updating of Call Center Agent Schedules". In: *Production and Operations Management* 19.3 (2010), pp. 353–367.

[65] Stephen K. Mok and J.George Shanthikumar. "A transient queueing model for Business Office with standby servers". In: *European Journal of Operational Research* 28.2 (1987), pp. 158–174.

[66] M. Pollak and D. Siegmund. "Approximations to the Expected Sample Size of Certain Sequential Tests". In: *The Annals of Statistics* 3.6 (1975), pp. 1267–1282.

[67] G. C. Reinsel. *Elements of Multivariate Time Series Analysis.* New York: Springer, 2003.

[68] Herbert Robbins. "Some aspects of the sequential design of experiments". In: *Bulletin of the American Mathematical Society* 58.5 (1952), pp. 527–535.

[69] Guillaume Roels and Georgia Perakis. "The "Price of Information": Inventory Management with Limited Information about Demand". In: *Manufacturing and Service Operations Management* 8.1 (2006), pp. 102–104.

[70] Paat Rusmevichientong, Zuo-Jun Max Shen, and David B. Shmoys. "Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint". In: *Operations Research* 58.6 (2010), pp. 1666–1680.

[71] Garrett van Ryzin and Siddharth Mahajan. "On the Relationship between Inventory Costs and Variety Benefits in Retail Assortments". In: *Management Science* 45.11 (1999), pp. 1496–1509.

[72] Denis Saure and Assaf Zeevi. "Optimal Dynamic Assortment Planning". Columbia GSB Working Paper. 2009.

[73] Herbert Scarf. "A min-max solution of an inventory problem". In: *Studies in The Mathematical Theory of Inventory and Production*. Stanford University Press, 1958.

[74] Herbert Scarf. "Bayes Solutions of the Statistical Inventory Problem". In: *The Annals of Mathematical Statistics* 30.2 (1959), pp. 490–508.

[75] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels*. Cambridge, Massachusetts: MIT Press, 2002.

[76] Bernhard Schölkopf et al. "New support vector algorithms". In: *Neural Computation* 12.5 (2000), pp. 1207–1245.

[77] Haipeng Shen and Jianhua Z. Huang. "Interday Forecasting and Intraday Updating of Call Center Arrivals". In: *Manufacturing and Service Operations Management* 10.3 (2008), pp. 391–410.

[78] Stephen A. Smith and Narendra Agrawal. "Management of Multi-Item Retail Inventory Systems with Demand Substitution". In: *Operations Research* 48.1 (2000), pp. 50–64.

[79] Kalyan Talluri and Garrett van Ryzin. "Revenue Management under a General Discrete Choice Model of Consumer Behavior". In: *Management Science* 50.1 (2004), pp. 15–33.

[80] Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. New York: Springer, 2004.

[81] Vladimir Vapnik. *The Nature of Statistical Learning Theory*. New York: Springer, 1995.

[82] Jr. Veinott Arthur F. "Optimal Policy for a Multi-Product, Dynamic, Nonstationary Inventory Problem". In: *Management Science* 12.3 (1965), pp. 206–222.

[83] Gustavo Vulcano, Garrett van Ryzin, and Richard Ratliff. "Estimating Primary Demand for Substitutable Products from Sales Transaction Data". Columbia Business School Research Paper. 2011.

[84] Michael R. Wagner. "Fully Distribution-Free Profit Maximization: The Inventory Management Case". In: *Mathematics of Operations Research* 35.4 (2010), pp. 728–741.

[85] Rodney B. Wallace and Ward Whitt. "A Staffing Algorithm for Call Centers with Skill-Based Routing". In: *Manufacturing and Service Operations Management* 7.4 (2005), pp. 276–294.

[86] Jonathan Weinberg, Lawrence D. Brown, and Jonathan R. Stroud. "Bayesian Forecasting of an Inhomogeneous Poisson Process With Applications to Call Center Data". In: *Journal of the American Statistical Association* 102.480 (2007), pp. 1185–1198.

[87] Jinfeng Yue, Bintong Chen, and Min-Chiang Wang. "Expected Value of Distribution Information for the Newsvendor Problem". In: *Operations Research* 54.6 (2006), pp. 1128–1136.

[88] Jinfeng Yue, Min-Chiang Wang, and Bintong Chen. "Mean-range based distribution-free procedures to minimize "overage" and "underage" costs". In: *European Journal of Operational Research* 176.2 (2007), pp. 1103–1116.