Data-Driven Methods for Personalized Product Recommendation Systems

by

Anna Papush

B.A., Cornell University (2012)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2018

Signature redacted

Author ....
Sloan School of Management
January 12, 2018

Signature redacted

Certified by..
Georgia Perakis
William F. Pounds Professor of Management
Sloan School of Management
Thesis Supervisor

Signature redacted

Certified by..
Pavithra Harsha
Research Staff Member
IBM T.J. Watson Research Center
Thesis Supervisor

Signature redacted

Accepted by ....
Dimitris Bertsimas
Boeing Professor of Operations Research
Co-Director, Operations Research Center

Data-Driven Methods for Personalized Product Recommendation Systems

by

Anna Papush

Submitted to the Sloan School of Management
on January 12, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

The online market has expanded tremendously over the past two decades across all industries ranging from retail to travel. This trend has resulted in the growing availability of information regarding consumer preferences and purchase behavior, sparking the development of increasingly more sophisticated product recommendation systems. Thus, a competitive edge in this rapidly growing sector could be worth up to millions of dollars in revenue for an online seller.

Motivated by this increasingly prevalent problem, we propose an innovative model that selects, prices and recommends a personalized bundle of products to an online consumer. This model captures the trade-off between myopic profit maximization and inventory management, while selecting relevant products from consumer preferences. We develop two classes of approximation algorithms that run efficiently in real-time and provide analytical guarantees on their performance. We present practical applications through two case studies using: (i) point-of-sale transaction data from a large U.S. e-tailer, and, (ii) ticket transaction data from a premier global airline. The results demonstrate that our approaches result in significant improvements on the order of 3-7% lifts in expected revenue over current industry practices.

We then extend this model to the setting in which consumer demand is subject to uncertainty. We address this challenge using dynamic learning and then improve upon it with robust optimization. We first frame our learning model as a contextual non-linear multi-armed bandit problem and develop an approximation algorithm to solve it in real-time. We provide analytical guarantees on the asymptotic behavior of this algorithm's regret, showing that with high probability it is on the order of $O(\sqrt{T})$. Our computational studies demonstrate this algorithm's tractability across various numbers of products, consumer features, and demand functions, and illustrate how it significantly out performs benchmark strategies. Given that demand estimates inherently contain error, we next consider a robust optimization approach under row-wise demand uncertainty. We define the robust counterparts under both polynomial and ellipsoidal uncertainty sets. Computational analysis shows that robust optimization is critical in highly constrained inventory settings, however the price of robustness

drastically grows as a result of pricing strategies if the level of conservatism is too high.

Thesis Supervisor: Georgia Perakis
Title: William F. Pounds Professor of Management
Sloan School of Management

Thesis Supervisor: Pavithra Harsha
Title: Research Staff Member
IBM T.J. Watson Research Center

Acknowledgments

First and foremost, I would like to thank my advisor Georgia Perakis for all of her guidance, patience, time and unwavering support throughout this entire academic journey. You have not only been an exemplary advisor in research and teaching, but also an infinitely kind and caring mentor. Thank you for dedicating the time to teaching me so many invaluable lessons in research, hard work, presentation skills, paper writing and teaching. You are truly the embodiment of a professor who is equally passionate about their work and their students and I aspire to some day be even a fraction of the mentor you have been to me. Thank you for the countless hours spent discussing interesting avenues of research, showing me innovative ways to approach new problems, and believing in my success even when I was not sure I could obtain it myself. I could not have been more fortunate in my academic advisor and I will be forever grateful for all that you have taught me and done for me.

I would also like to thank my mentor and second advisor, Pavithra Harsha, for so many great years of collaboration. Thank you for teaching me so much about the research process, paper editing, and taking pride in your work. I so appreciate the many hours you spent helping me with drafts, working on presentations, and teaching me how to tackle new problems. You are truly a top notch researcher and such a kind and generous person, and I have learned so much from you over the past several years. Thank you for also giving me the opportunity to work extensively at IBM and come into contact with so many interesting people and practical problems, I am so grateful to have been able to work with you. To both of my advisors, thank you for all of your support and for making me feel like family.

There are many faculty members who have had a significant impact on my time at the ORC. I would like to thank Tauhid Zaman for serving on my thesis committee and for all of your very interesting and insightful input which has led me to consider new directions and approaches to my research problems. Thank you also to Dimitris Bertsimas for all of your guidance, support and willingness to allow me to prove myself throughout my time at MIT. I am very grateful for all that you have done for me and for all of your generosity. Thank you to both Dimitris and Patrick Jaillet for constantly striving to make the ORC a better place and for working so hard to incorporate us students into that growth. Thank you also to all of the professors whose courses I was fortunate enough to take, you continue to inspire and provide an exceptional education for all of us. I would also like to sincerely thank Laura Rose and Andrew Carvalho for taking the time to answer an infinite number of questions and for always being available to help with any logistical or administrative challenges. Thank you to Laura for all of your help during some of the more challenging times. Also, I would like to thank my undergraduate research project advisor Steven Strogatz for inspiring me to pursue applied mathematics, mentoring me both during and after my first research experience, and ultimately putting me on the track to graduate school, which has been one of the most fulfilling experiences of my life.

I would also like to thank members of the IBM T.J. Watson Research Center, in addition to Pavithra, for giving me the opportunity to collaborate with you and work in such a stimulating and interesting environment. Thank you to Markus Ettl, it was such a pleasure to work in your group and thank you for so many interesting discussions, during which you always had great business insights. Thank you also to Arun Hampapur for your support of our work and for taking the time to discuss our project. Also, I owe a big thank you to Ajay Deshpande for all of your patience and time spent helping me understand SPSS and learning to work with client data. Thanks to all of you, my time at IBM was one of the hi-lights of my PhD journey.

It goes without saying that my time at MIT would not have been the same without the friendships I was fortunate enough to have made along the way. To my roommates Joey Huchette, Adam Elmachtoub and Stefano Traca, thank you for so many years of being my family and my home away from home. To Stefano Traca, I am so grateful to have had you along for this journey with me. To Charles Thraves, thank you for becoming another older brother to me, for such an incredible trip in Chile, and for so many good times spent together. To Andrew Li, thanks for being our social kingpin for so many years. To the fantastic four, Stefano Traca, Alex Sahyoun and Daniel Smilkov, thank you for so many laughs, late nights, study sessions, and memories.

To Andrew Li, Charles Thraves, Stefano Traca, Alex Remorov, Velibor Misic and the rest of our first year group, thank you for making the transition to Boston so easy. To all of my other ORC friends: Max Biggs, Max Burq, Zach Owen, Hugh Carson, Sebastien Martin, Arthur Flajolet, Virgile Galle, Mathieu Dahan, Mariapaola Testa, Cecile Casses, and Ludovica Rizzo, thank you for so many memorable times and adventures, these five years would not have been the same without you. To my partner TA's Park Sinchaisri, Max Biggs and Lennart Baardman, it was so great working with all of you! To my fellow INFORMS officers Velibor Misic, Alex Weinstein and Alex Sahyoun, thank you for so many fun and successful events. To Matthieu Monsch, Sean Whipple and Divya Singhvi, thank you for all of your help and collaboration. To my second family: Erica Rhodin, Sofya Muftishvili and Sasha Perepelova, and to Melanie Sovern, thank you for all of your love and help over the years, I am so fortunate to have such a strong foundation of support. To Tom Clark, thank you for your unwavering support, help and patience during all the most difficult parts of this final stretch, I am so grateful for you and all you do for me.

Finally, I would like to above all sincerely thank my family for being a never ending source of support, love, motivation and guidance throughout my entire life. To my parents, brother, grandparents, aunt, and the rest of my extended family, thank you for showing me the real meaning of hard work, perseverance and resilience. Watching all of you overcome the challenges of immigration and truly live the American dream has been the most humbling and valuable lesson of my life. I would not have completed this degree or any other achievement without so much support from all of you, thank you for teaching me the value of family and what really matters in life. To my parents, Dmitry and Gina, there are no words with which I can adequately thank you for everything you have done and continue to do for me every day. You are my role models and I am so grateful to both of you for teaching me everything that I know and for continuing to patiently provide me with so much advice during every chapter of my life. To my brother Andrew, thank you for always being my best friend and for supporting me in everything. To my parents and brother, you have been my foundation and my guiding lights in life and I dedicate this thesis to you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The rapid expansion of the online market over the past two decades has directly resulted in an increasing focus on the development of research in the field of revenue management. More specifically, the growing availability of consumer data across all industries operating an online platform has sparked a great deal of interest in the modeling and analysis of personalized assortment optimization, constrained assortment planning, dynamic pricing and online demand learning. Furthermore, the utilization of the above techniques in practice has resulted in significant increases in revenue and sales volume for businesses that have leveraged this growing source of information successfully in their operational strategies. In this thesis, we study the areas of inventory planning, strategic pricing and demand learning in the context of product recommendation systems for bundles as follows: in Chapter 2 we develop a model for personalized bundle pricing and recommendation and propose approximation approaches with analytical guarantees for real-time implementation; in Chapter 3 we present a detailed analysis of the performance of these algorithms through two data-driven case studies from the retail and travel industries; and, in Chapter 4 we extend our modeling framework to include potential demand uncertainty, which we capture using dynamic learning through a contextual non-linear multi-armed bandit formulation and robust optimization under various uncertainty sets.

1.1 Motivation

Over the past few decades, the expansion of the online channel has experienced dramatic development across all industries ranging from retail to travel to personalized shopping and styling. In retail alone, the online market is predicted to increase by over 34% in sales over the four year period from 2016 to 2020, from a recorded $373 billion to an estimated more than $500 billion. This growth has generated an increasingly availability of information regarding consumer demographics and preferences, resulting in a greater emphasis among online sellers on personalizing customer experiences in order to improve retention and satisfaction. Stemming from this trend, effective personalized recommendation services across businesses in the online sector have experienced great prosperity over the past several years. As a direct example of this, Dollar Shave Club, a startup specializing in male grooming that sends its customers monthly personalized boxes of razors and shaving supplies, recorded $152 million in revenue in 2015, which was projected increase to over $200 million by end of 2016 and ultimately led to a $1 billion acquisition by Unilever. These success stories demonstrate that consumers are both interested in and willing to pay for customized experiences. Thus, a more sophisticated product recommendation system can provide the necessary competitive edge for any online seller, making the difference on the order of millions in profits.

These current industry practices and research developments in revenue management directly motivate the primary goal of Chapter 2, which is the development of a personalized model that selects, prices, and recommends a bundle of related products to a consumer during their online session. This model presents a framework for constructing dynamic bundle offers and combines diverse recommendations with personalized discounts by leveraging consumer profiles and in-session context, while considering the trade-off between myopic profit maximization and long-term profitability under inventory constraints. The resulting dynamic programming problem is both structurally complex and practically challenging to implement, therefore we develop two classes of approximation algorithms in order to utilize this model efficiently in real-time in an online setting. We develop analytical guarantees on both classes of algorithms relative to their respective "optimal" strategies and analyze the empirical performance of these bounds depending on initial instance parameter settings using real data. We also develop and analyze a multi-level pricing setting in which our model and resulting algorithms align an upper-level periodic nominal price trajectory problem with our lower-level personalized discounted offers, and extend our analytical guarantees to this context.

While we develop a new modeling framework for simultaneously considering personalization, inventory balancing and long-term profit maximization within the context of the online bundle recommendation problem, we also aim to demonstrate the practical relevance of our proposed algorithms. In addition to providing analytical guarantees, we want to show that both classes of approximation methodologies solve efficiently in real-time and provide quality solutions to the original dynamic programming problem. Thus, in Chapter 3 we implement these algorithms and analyze their performance through two in-depth case studies: (i) using point-of-sale transaction data from a major U.S. e-tailer, and, (ii) using ticket transaction data from a premier global airline. These case studies demonstrate that our approaches result in significant improvement on the order of 3-7% lifts in revenue over existing industry practices and provide efficient solutions that obtain up to 98% of the expected profit of a full-information offline benchmark strategy. In retail and travel industries which operate on razor-thin margins and are increasingly utilizing personalized strategies in the online channel, these gains are very substantial and can scale up to millions of dollars in revenue.

In the first two chapters we consider the personalized bundle pricing and recommendation problem under the assumption that there is sufficient historical consumer data to adequately estimate consumer preferences and demand. However, when considering the online setting in practice, consumer purchasing information is often unavailable in a multitude of realistic cases such as the following: where no (or few) product bundles have ever been offered historically, where new products are introduced into the market with no previous purchase history known to the seller, and finally, where first time shoppers frequently appear in the market without any purchase history or preferences known to the seller. Businesses such as StitchFix, which offers personalized clothing styling and achieved $250M in revenue in 2015 that nearly tripled to $730M by the end of 2016, obtain their success by collecting, analyzing and effectively leveraging consumer feedback data that is traditionally not available to brick-and-mortar businesses in these uncertain demand scenarios. Thus, motivated by these industry successes, we consider these challenging and relevant online scenarios in Chapter 4 by extending our prior model to account for demand uncertainty. We study this by considering two approaches: (i) dynamic learning, and, (ii) robust optimization. We initially formulate our learning model under setting (i) as a contextual non-linear multi-armed bandit problem. We propose an approximation method based on estimating the first-order Taylor series of the expected reward function, which requires no assumptions on the functional form of the demand other than it being differentiable. We establish analytical guarantees on the asymptotic behavior of this algorithm's regret compared to an oracle strategy and the empirical results show that it performs well across various demand functions, numbers of products and consumer features relative to relevant benchmarks from the existing literature. However, we also find that in certain cases dynamic demand learning innately captures a significant amount of error, which leads us extend our modeling approach to setting (ii), where we consider a robust optimization approach under row-wise demand uncertainty. We define the robust counterparts under both polynomial and ellipsoidal uncertainty sets. Our computational analysis shows that robust optimization is critical in highly constrained inventory settings, however the price of robustness drastically grows as a result of pricing strategies if the level of conservatism is too high.

1.2 Thesis Contributions

In Chapter 2 we develop a new modeling framework that captures personalization at an individual level, as well as the trade-off between myopic profit maximization and long-run profitability through inventory management. This approach guards against inventory-related losses resulting from both salvaging excess inventory and stock outs as a result of consumer choices. However, this is a challenging dynamic programming problem that involves personalization, bi-level pricing, future demand forecasting, and inventory management, and is therefore not tractable for implementation in the online setting. Therefore, we construct two classes of approximation algorithms, multiplicative and additive, which provide efficient real-time recommendations. We present analytical guarantees on the performance of the multiplicative approach relative to a full-knowledge benchmark strategy that knows the entire consumer arrival sequence in advance. We also provide guarantees on the optimality gap between the additive approaches and our original dynamic programming formulation that is linear in the number of available products. We empirically assess the performance of these analytical guarantees by considering their behavior based on various instance parameters such as initial inventory levels and continuity in the demand functions. Based on data-driven studies, we find that our overall best heuristic obtains 90-98% of the expected revenue of the full-information benchmark across various problem instances.

Chapter 3 presents the detailed analysis of two case studies on actual data from the retail and airline travel industries that demonstrate significant improvement in expected revenue on the order of 3-7% over existing industry practices. We consider a two year period of transaction data from a large U.S. e-tailer and find that our algorithms provide output in real-time with average expected gains of up to 12% in profits over current pricing schemes for seasonal products. We also analyze a one month period of ticket transaction data from a premier global airline and find that average gains in predicted sales volume and revenue are as high as 8-9% in data scenarios when consumers are unaware of the existence of ancillary services such as in-flight wi-fi or priority boarding. We objectively analyze our model's ability to distinguish between changes in online context and changes in personalized features and found that changes in online context resulted in different bundle composition but similar discounting strategies, whereas changes in personalized features generated discounts that differed by as much as 5% on average. In aggregate, we found that the overall greatest gains in revenue resulted from *personalized pricing* targeted at consumers with lower price sensitivities, who had generally higher willingness-to-pay across all products and therefore resulted in significantly more conversions from discounted offers. However, the largest growth in predicted sales came from *relevant product recommendations* and differed between various consumer groups depending on the products. Finally, we present a detailed comparison of the algorithm classes and empirically demonstrate that multiplicative methods perform up to 7% worse than additive methods relative to the full-knowledge strategy, but are significantly easier to implement. Thus, the combined work in these two first chapters presents a solution for a challenging problem that is highly relevant to the existing literature, while providing both analytical guarantees and data-driven computational results that highlight the complexity of the problem structure and extract innovative business insights.

The main contributions in Chapter 4 lie in the analysis and development of (i) a new high-dimensional learning framework based on contextual non-linear multi-armed bandits, and, (ii) the robust optimization formulations for the personalized bundle recommendation problem. Both approaches capture individualized demand modeling in the online bundling setting and also capture the trade-off between profitability and inventory management. In the context of dynamic learning, many convex optimization methods place assumptions on the structure of the expected reward function in this problem. We construct a more generalized approach that requires only the assumption that the function be differentiable. Furthermore, we develop an algorithm based on upper confidence bounds and first-order Taylor series approximation in order to then implement our proposed approach and establish analytical guarantees on the asymptotic behavior of the regret relative to an oracle strategy. We also present empirical results that demonstrate that this algorithm converges to the true demand faster than existing benchmarks from the literature and it performs well over a range of settings with respect to various numbers of products, numbers of consumer features, and different demand functional forms. However, we find that demand estimation particularly in the case of model misspecification, may be subject to a significantly level of error. Therefore, we also the robust optimization setting in which we define the robust counterparts to the personalized bundle recommendation problem using both polynomial and ellipsoidal uncertainty sets. We conduct an extensive computational analysis to analyze the feasibility of optimal solutions in the full-information setting when demand is slightly perturbed and find that these are particularly sensitive to initial inventory settings. Furthermore, we find that in mildly conservative settings ellipsoidal uncertainty sets typically outperform polyhedral uncertainty sets by 4-7% with respect to the "optimal" objective value of the nominal problem, which is relatively marginal when considering the significant computational advantage of using polyhedral sets. Our numerical results demonstrate that the use of robust optimization to account for demand uncertainty is crucial in highly constrained inventory problems, in which the cost of robustness is only up to 10%, but optimal solutions to the nominal problem are infeasible in up to 50% of instances with perturbations of only 0.10% in demand. Furthermore, a comparison of the numerical analysis between the two demand uncertainty settings demonstrates the extent of this critical result. We find that on average, the demand estimation error associated with model misspecification ranges from 1.3% to 1.9%; in parallel, perturbations in demand by 1-2% in the robust computational study resulted in infeasibility in 30-70% of problem instances. Therefore, error in demand estimation can have a significant impact on the quality of the solutions made in the fundamental recommendation model. In a practical setting, this implies that even minor errors in demand estimation can result in infeasible and sub-optimal solutions for the bundle recommendation and pricing problem, and it is therefore absolutely necessary to also account for this effect when implementing our model in realistic business settings.

Chapter 2

Personalized Bundle Pricing and Recommendation

2.1 Introduction

The online market as a whole has grown enormously over the past decade. According to a forecast released by Forrester Research in 2016, U.S. e-commerce retail sales are expected to grow from \$373 billion in 2016 to more than \$500 billion in 2020, an increase of over 34%; furthermore, the online sector alone impacts over \$1.5 trillion of total retail sales in the United States. This surge in online shopping has led to an increased availability of data regarding consumer preferences, which can be leveraged by businesses across all industries in order to improve operations, revenue and consumer satisfaction. As a direct example of this, strategic recommendation services that utilize such data effectively are currently undergoing massive and rapid expansion. StitchFix, which offers personalized clothing styling for its customers and only became cash-flow positive in 2014, achieved \$250M in revenue in 2015 that nearly tripled to \$730M by the end of 2016. The travel industry has also undergone dramatic growth over the past decade due to the increasing availability of online services. As a result, travel products are becoming increasingly commoditized. Consumers are not willing to pay exorbitant fees for these generic services, resulting in a great deal of competition across industries such as airlines. However, travelers are both interested in and willing to pay for customized experiences. Therefore, businesses in the travel and hospitality industries have begun providing recommendations at the time of reservation by offering ancillary services that are un-bundled from ticket cost or room rate. In the case of airlines, supplementary products to improve the travelers' experience before, during and after their ticketed trip such as VIP lounge access, priority boarding, seat upgrades, in-flight wi-fi, and destination-relevant deals are now provided at their own prices and offered throughout the online purchase process, which was previously not the case. As a result of these industry trends in the online sector, the development of a more sophisticated product recommendation system can provide the necessary competitive edge for any online seller, making the difference on the order of millions in profits.

As demonstrated by these industry examples, the majority of businesses with an online component now utilize recommendation systems. However, these methods are often primarily based on historical purchase trends across segments of the online population when there is also a wealth of individualized consumer information. Motivated by this increasingly prevalent cross-selling problem and current industry practices, the goal of our work is the development of a personalized model that selects, prices, and recommends a bundle of related products to a consumer during their online session. Having dynamically received this offer while browsing a particular item or ticket itinerary, the consumer can then choose to accept this discounted offer, or purchase any combination of items at their full prices, or simply exit the online marketplace without making any purchase at all. This dynamic bundle offer is constructed using a new model that combines diverse recommendations with personalized discounts by leveraging consumer profiles and in-session context, while considering the trade-off between myopic current profit with long-term profitability under inventory constraints. Note that because we consider the possibility that consumers may choose to purchase products at their full prices, we must additionally incorporate the upper-level problem of determining the time-dependent trajectories of full prices over the course of the selling horizon. Thus, the novelty of this work consists of *simultaneously* incorporating personalization, bundle assortment selection, bi-level pricing, and inventory-balancing within this particular online bundle offer setting. These challenges have not yet been explored jointly in the existing literature.

In order to construct our bundle pricing and recommendation model we focus on incorporating all of the above components simultaneously. We aim to make relevant offers by solving a personalized bundle assortment selection and pricing problem that uses individualized propensity-to-buy models based on consumer profiles and online context. We integrate this personalized online offer setting within the goal of long-run profitability by additionally considering future demand through an inventory balancing function in our model, which improves expected profits by mitigating costs associated with overstocking and lost sales. Balancing all of these factors is novel to the analytical problem and practically crucial to sellers from an operational perspective, but also gives rise to several challenges with respect to both the analytical problem structure and its implementation. The combination of all of these components results in a complex dynamic programming problem that is highly intractable in an online setting. Furthermore, focusing on inventory-constrained products leads to the additional difficulty of incorporating upper-level dynamic pricing schemes that affect the full prices of products as the selling horizon progresses. Thus, our resulting model simultaneously addresses personalization, multiple levels of pricing, bundle assortment selection, demand forecasting, and inventory management. We develop approximation algorithms and provide analytical guarantees that improve in tightness as the problem becomes less inventory constrained. Furthermore, we analyze the performance of our algorithms through two case studies: (i) using point-of-sale transaction data from a major U.S. e-tailer that includes personalized features such as customer IDs and loyalty information, and, (ii) using ticket transaction data from a premier global airline that includes consumer-specific information such as tier level, miles balance and previous flight history at the time of ticket purchase. These case studies demonstrate that our approaches result in significant improvement in expected revenue over existing industry practices. In industries that operate on razor-thin margins, these gains can scale up to several millions of dollars in revenue.

2.1.1 Contributions

We analyze the problem of personalized online bundle recommendation, which lies at the intersection of several branches of revenue management literature. Our main contributions consist of,

1. **Two classes of approximation algorithms that provide real-time bundle recommendations and simultaneously incorporate personalization, inventory balancing and tractability.** We develop multiplicative and additive methods to implement our model in real- time in an online setting. These heuristics capture personalization as well as the trade-off between myopic profit maximization and long-run profitability under inventory constraints. We also coordinate the dynamic lower-level personalized bundle prices with an upper-level global pricing strategy that periodically determines the time-dependent trajectories of the full prices of all items.

2. **Analytical guarantee on the performance of the multiplicative algorithm and empirical comparisons of both classes.** We provide a bound on the ratio of the expected revenue of the multiplicative approach relative to a full-knowledge strategy that knows the entire consumer arrival sequence in advance. This becomes even tighter as the problem becomes less constrained by inventory and falls on average within 15% of the algorithm's actual empirical performance ratio on data. We further compare the empirical performance of both algorithm classes and show that on average, the overall best heuristic is an additive benchmark that obtains 90-98% of the expected revenue of the full-information benchmark across various initial inventory settings. Furthermore, we show that the multiplicative approach is **easier to implement compared to the additive methods** and on average obtains an expected revenue that is within 1-6% of that achieved by additive methods, relative to the full-knowledge strategy.

3. **Two detailed case studies on actual data from the retail and airline travel industries that demonstrate significant improvement in**

expected revenue on the order of 2-7% on average over existing practices depending on the setting. In the retail case, we observed that our algorithms provide output in real-time with predicted gains of up to 14% in revenue over current pricing schemes in the most unconstrained discounting settings. In the airline case, our model predicted improvements in sales volume and revenue as high as 7-8% over current strategies in settings when a fraction of the online population is unaware of the existence of ancillary services. The greatest gains in expected revenue were a result of *personalized pricing* targeted at consumers with lower price sensitivities, who are easily incentivized to make additional purchases through smaller personalized discounts. Conversely, the largest growth in predicted sales volume was dependent on product category and primarily a result of *relevant recommendations*, resulting in lifts on the order of up to 10% over current practices.

2.1.2 Literature Review

We consider two bodies of literature most closely related to our work: constrained assortment optimization and dynamic pricing. The first line of literature pertains to the assortment planning problem under capacity constraints. Initial works such as [21] consider a single-period stochastic model under which the retailer selects a profit-maximizing assortment of substitutable products and determines their initial stock prior to the selling period, under the assumption that consumers choose products according to a multinomial logit model, which was extended in [22] to incorporate dynamic substitution effects when a consumer's product of choice may be stocked out. [18] presents a summary of the initial works on the single-period assortment planning setting under inventory or budget constraints, which captures extensions to other consumer choice models and various dynamic substitution effects. Later works such as [2] provide provably efficient algorithms under stochastic demand and dynamic substitution and show that these approximations are order optimal, or near-optimal in the case of [13], under general random-utility choice models. [11] consider both assortment cardinality and display space constraints, showing that the assortment problem is efficiently solvable by linear programming, while the space problem is NP hard. Considering a setting in which each period has two phases of pricing, [15] explore the joint problem of inventory and markdown management with inter temporal demand substitution and perishable goods. Building on the one-period assortment planning problem, another body of works has recently evolved in the direction of the dynamic assortment optimization problem that is solved distinctly for each consumer arrival. [30] and [29] study this problem when the parameters of the consumer demand functions are unknown. In [3] they formulate this as a dynamic program to identify which optimal assortment of substitutable goods to offer each consumer and develop inventory threshold policies for determining this. By contrast, [12] propose an index-based inventory balancing approach for determining the optimal personalized assortments, which motivates our multiplicative algorithm that extends this setting by also incorporating pricing. [33] and [16] present more generalized approaches to inventory balancing in the dynamic assortment planning problem and demonstrate the value of duality-based approximations of DP formulations for online implementation in the context of the network revenue management problem and dynamic resource allocation problems, respectively. In [9] they develop an asymptotically optimal policy for this dynamic setting, and in [19] and [7] they consider further extensions under the d-nested logit choice model and the MNLD choice model (in which consumer segments have non-overlapping consideration sets), respectively. In this work, we consider an inventory-constrained assortment planning problem to dynamically determine the composition of personalized product bundles. However, the bundle recommendation system we propose sets this work apart from the existing literature in constrained assortment planning primarily because we extend the problem to incorporate dynamic pricing. Furthermore, we do not limit our analysis to any specific consumer choice model, nor do we assume that the assortment consists of only substitutable goods.

The second body of work is related to dynamic pricing and cross-selling. Dynamic pricing literature, which initially focused primarily on single products, is very well summarized in [4]. However, there is a vast body of more recent literature on this topic across a wide variety of consumer utility models and considering multiple products. In [26] and [27] they explore multi-period pricing strategies in which retailers determine optimal inventory and pricing decisions based on consumer behavior regularities. [20] consider consumer disappointment from stock outs and determine a policy for pricing and product rationing to leverage this strategic behavior to increase profits. The applied work in [5] presents a price optimization model for markdowns in fast-fashion that was implemented by Zara & Co. and connects retail operations with pricing methods, which is in line with our practical goals in this work. There is also a body of work on online pricing problems in which information is revealed dynamically upon arrival of a consumer or service request. In [1] and [31] the authors consider the online resource allocation problem under different settings and develop efficient algorithms for maximizing long-term system revenue based on dual price updates. By contrast, [17] develop an algorithm based on a scaled version of the partially known primal problem (as opposed to dual prices) to obtain and round fractional solutions and develop integral allocations for all requests. [24] study the problem of ordering and pricing products for consumers who view the items sequentially by using a Markov decision process to obtain the optimal policy for finding the product prices and purchase probabilities. Tying online pricing to learning with a multi-armed bandit framework, [8] develop an algorithm based on Thompson sampling for dynamically pricing multiple products under inventory constraint in order to maximize long-run profitability. In [23] they study randomized markdown strategies and demonstrate the benefit of these pricing methods for retailers by exploiting known consumer product monitoring behavior. By contrast to all of these works, dynamic cross-selling grew as its own field from economics and initially did not incorporate pricing. [6] provides an overview of the growth and expansion of this early literature. [25] is a key pivotal work that combined these fields by analyzing models with stochastic arrivals for the joint problem of cross-selling and pricing in which a consumer has one primary product of interest and is offered one additional complementary product at a discounted price for both; in [28] the authors extend this dynamic programming setting by proposing a rule-based approach to the joint bundle selection and pricing problem. New directions such as [34] consider request for quote (RFQ) models where consumers interactively participate in the pricing process. Recent work in [10] presents a dynamic approach to product pricing and framing to determine optimal product displays on webpages for consumers. We consider a setting in which the bundle offer is presented as an additional option that the consumer can choose not to purchase in favor of any other combination of non-discounted products. Thus, since all products are also available for purchase at their full prices, we must consider the upper-level problem of determining the time-dependent full product prices over the course of the selling horizon. Thus, our work addresses a two-level pricing problem and aligns: (i) the lower-level personalized bundle prices offered dynamically to each consumer, with, (ii) the upper-level full price trajectories for each product. To the best of our knowledge, this simultaneous bi-level pricing problem is not addressed in the cross-selling literature.

2.2 Problem Setting and Model Formulation

We consider a monopolist online seller that makes a dynamic bundle offer to each arriving consumer who may choose to accept the offer, purchase individual items separately at full price, or choose to purchase nothing at all, as shown in Figure 2-1 below. If the consumer chooses to purchase either the bundle or some other collection of items at their full prices, we assume that they only purchase one unit of each item. Let us consider a set of items $i = 1, ..., n$ denoted by $\hat{S}$. These items' prices may affect one another and they can be complementary, substitutable, or even independent as is often the case in the travel industry. Given a captive online consumer considering products within $\hat{S}$, or a specific ticket itinerary for which $\hat{S}$ is the set of ancillary goods, our model offers a relevant bundle of products from $\hat{S}$. We are interested in cases where $\hat{S}$ contains inventory-constrained products that we leverage to maximize expected long-run profitability by accounting for future demand. Therefore, we consider a finite selling horizon with a fixed number of periods T with no replenishments.

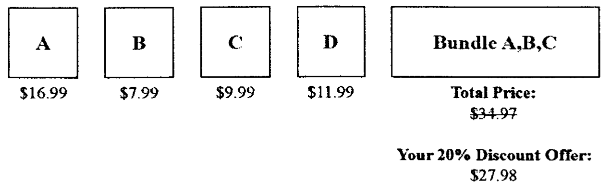

Figure 2-1: This is an example of a personalized bundle recommendation when a consumer is shopping online and is offered the set A, B and C at a 20% discount, otherwise they can choose from any combination of A, B, C and D at full price.

Each arriving consumer is uniquely described by a combination of categorical and continuous features related to preferences, demographics, purchase history, loyalty, and online shopping context. Thus, we do not consider a discrete set of consumer types as is traditionally done in segmentation and instead assume that there is an infinite set of continuous consumer types. Furthermore, since we address a bi-level pricing problem, we index consumers within a given period t by (k, t), where $k = 1, ..., K^t$ and the total number of arrivals K^t in each period can differ. We define the full price of item i in period t as $\bar{p}_i^t$; thus, the full price $\bar{p}_{S_{k,t}}$ of a bundle $S_{k,t}$ offered to consumer (k, t) is defined by,

$$\bar{p}_{S_{k,t}} = \sum\nolimits_{i \in S_{k,t}} \bar{p}_i^t \tag{2.1}$$

The full prices $\bar{p}_i^t$ are not necessarily fixed throughout the horizon and may follow some dynamic trajectory, summarized in each period by vector $\bar{\mathbf{p}}^t = [\bar{p}_1^t,\ \bar{p}_2^t,\ \dots\ ,\ \bar{p}_n^t]$. We thus define price vector,

$$\mathbf{p}_{S_{k,t}} = [\bar{\mathbf{p}}^t,\ \ p_{S_{k,t}}] = [\bar{p}_1^t,\ \ \bar{p}_2^t,\ \ \dots\ ,\ \ \bar{p}_n^t,\ \ p_{S_{k,t}}], \tag{2.2}$$

in which we append the discounted price of the personalized bundle for consumer (k, t) to the vector of full price settings for period t. It is common in business practice for sellers to consider discrete price ladders. Therefore, we make the assumption that we have a fixed set of price levels for every product i from which we can choose to construct bundle offers. We define the individual consumer propensity-to-buy $\xi_S^{k,t}(\mathbf{p}_{S_{k,t}})$ as the probability that consumer (k,t) will purchase the combination of products $S : S \neq S_{k,t}$ (and nothing else) at their *full prices* if their personalized bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$.We similarly define the probability that consumer (k,t) will purchase only the bundle $S_{k,t}$ (and no other products) when it is offered at the *discounted price* $p_{S_{k,t}}$ as $\xi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})$. We will refer to $\mathbf{e}_{S_{k,t}}$ as the bundle unit vector that takes the value 1 for all $i \in S_{k,t}$ and 0 otherwise. Finally, we define $\mathbf{I}^{k,t}$ as the vector of inventory levels of all $i \in \hat{S}$ at the time when consumer (k,t) arrives, written explicitly as $\mathbf{I}^{k,t} = [I_1^{k,t},\ I_2^{k,t},\ \dots,\ I_n^{k,t}]$. This leads to the following decision variables for any given consumer (k,t): the optimal bundle to recommend $S_{k,t} \in \hat{S}$, and, its personalized price $p_{S_{k,t}} \leq \bar{p}_{S_{k,t}}$. For convenience, we summarize the notation in Table A.1 in Appendix A.1.

2.2.1 Dynamic Programming Formulation

We formulate this personalized bundle offer problem ideally as a dynamic program, as is traditional in the revenue management literature. This results in a complex model that is difficult to solve, as we discuss below in Section 2.2.2. This DP approach leads us to the following formulation (2.3), defined by $\{\text{Dynamic}\}_{\forall(k,t)}$:

$$
\begin{aligned}
\underset{S_{k,t}\subset\hat{S},\, p_{S_{k,t}},\, \bar{p}_i^t}{\text{maximize}} \quad & V_{k,t}(\mathbf{I}^{k,t}) \\
\text{subject to} \quad & V_{k,t}(\mathbf{I}^{k,t}) = \sum\nolimits_{S\subset\hat{S}} \xi_S^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \Bigg(\Big(\sum\nolimits_{i\in S} \bar{p}_i^t \Big) \\
& \qquad + \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}}\right) \cdot \mathbb{1}_{\{S_{k,t}\subset S\}} + V_{k+1,t}(\mathbf{I}^{k,t} - \mathbf{e}_S) \Bigg) \\
& V_{K^t+1,t}(\mathbf{I}) = V_{1,t+1}(\mathbf{I}) \qquad \forall K^t,\ t = 1,\dots,T \\
& (1-\epsilon)\bar{p}_{S_{k,t}} \leq p_{S_{k,t}} \leq \bar{p}_{S_{k,t}} \qquad \forall (k,t),\, S_{k,t} \subset \hat{S}
\end{aligned}
\tag{2.3}
$$

We solve this problem for every consumer $k = 1,\dots,K^t$ who arrives within each period $t = 1,\dots,T$ and connect the periods t through the forward-looking inventory cost-to-go functions $V(\cdot)$. Note that in addition to the offer for each consumer, the full price trajectories $\bar{p}_i^t$ over all products in all periods are also decision variables in this model. However we note here that unlike the consumer-level bundle offer decisions, these prices are calculated periodically in an upper-level pricing problem at the conclusion of each period t, then held fixed for that period and updated with the most currently inventory for period $t+1$.

The objective function consists of several terms: the first term captures the probability $\xi_S^{k,t}(\mathbf{p}_{S_{k,t}})$ with which a consumer (k,t) purchases some set of products $S \subset \hat{S}$, , summed over all possible sets S (note that this captures the cases in which S is a superset that encompasses the personalized offer $S_{k,t}$); the second set of terms account for the expected revenue from consumers purchasing individual items at full price, as well as from accepting the bundle offer (at which point the discounted bundle at full price $\bar{p}_{S_{k,t}}$ is removed from the summation of $\bar{p}_i^t$). Note that we do not include the probability of a "no-buy" because this is innately captured in the set of collections of products $S \subset \hat{S}$ that also includes the null set $\emptyset$, corresponding to the consumer's decision to make no purchase. This is a complex dynamic programming problem because it relies on knowledge of future demand and inventory levels to utilize $V_{k+1,t}(\cdot)$ in making bundle offers. The first constraint accounts for the recursive transition of the inventory revenue-to-go function $V(\cdot)$ between periods and the second constraint limits the depth of the bundle discount $\bar{p}_{S_{k,t}}$ and ensures that the bundle offers remain attractive. This DP formulation is intractable for the online setting due to the forward-looking nature of the functions $V(\cdot)$ and the need to calculate the full prices $\bar{p}_i^t$ for all products in all periods.

If we were given the full price trajectories $\bar{p}_i^t$ for every i, along with the values for the functions $V(\cdot)$ at all possible inventory levels and bundle combinations, then solving formulation (2.3) would be an enumeration over all the discrete prices and bundle combinations. We remark that even if provided with all of these values, this enumeration problem could potentially suffer from the curse of dimensionality. However, appropriate limitations to ensure a small size of $\hat{S}$ would reasonably bound the number of bundle combinations and thus make this problem tractable. Therefore, if we were provided with the full price trajectories $\bar{p}_i^t$ and all the precise values of $V(\cdot)$ (which become 0 in the absence of inventory constraints), we could use the above model to make optimal individually tailored offers of bundles $S_{k,t}$ at prices $p_{S_{k,t}}$ for each consumer in a tractable manner.

2.2.2 Challenges

While we are interested in optimally solving $\{\text{Dynamic}\}_{\forall(k,t)}$ for every consumer in real-time, realistically the consumer arrival sequence, full price trajectories, and values of $V(\cdot)$ are unknown. This results in three fairly sizable challenges: (i) how to estimate a personalized propensity-to-buy, (ii) how to determine the upper-level prices $\bar{p}_i^t$ and align them with offers $(S_{k,t}, p_{S_{k,t}})$, and, (iii) how to estimate the values of $V(\cdot)$ while maintaining tractability in an online setting.

Developing a personalized bundle recommendation at an individually tailored price requires the most granular possible estimate of a consumer's propensity-to-buy. Traditional methods lack distinctive information that distinguishes a customer from others in their segment. Therefore, to address (i), we use machine learning methods to fit high-dimensional models that capture all of these features through covariates as described in detail in Chapter 3. Considering an inventory-constrained problem with a finite horizon, results in challenge (ii) of determining and incorporating an upper-level pricing strategy into our model that alters the full prices of individual products over time. Thus, we propose a method (described in Section 2.3.1) for determining these price trajectories within our problem framework as follows: at the beginning of each t we calculate the full prices $\bar{p}_i^t$ across all products i and fix them for the duration of that period, after which we update them at the beginning of the next period $t+1$ using current inventory levels after consumer demand is realized. This rolling approach coordinates the upper-level full price trajectories with the lower-level bundle offers (which are based on the values of $\bar{p}_i^t$) made to individual consumers within a given period. Finally, we address (iii) by developing various approximation approaches to the forward-looking inventory balancing functions $V(\cdot)$. A common linear programming approximation, in which we solved a series of LPs to estimate the values of $V(\cdot)$

at various inventory levels (as is commonly done in the revenue management literature, see [32]), runs far too slowly. Thus we propose two classes of approximation algorithms in Section 2.3, multiplicative and additive methods, that are practically tractable and therefore applicable in an online setting.

2.2.3 Clairvoyant Formulation

Before presenting any approximation algorithms, we first generalize the dynamic programming problem to a "full-knowledge" model to establish a benchmark against which we can compare any algorithm's performance. We assume that the entire consumer arrival sequence $\{k,t\}_{\forall k=1,\dots,K^t}^{t=1,\dots,T}$ is known in advance, as well as the full price trajectories $\bar{p}_i^t$ for all products i in all periods t, which we assume are provided to us by an oracle. In order to model this perfect information setting we propose the following formulation that we refer to as the {Clairvoyant} problem:

$$
\begin{aligned}
\underset{y^{k,t}_{S_{k,t}}}{\text{maximize}} \quad & \sum\nolimits_{t=1}^{T} \sum\nolimits_{k=1}^{K^t} \sum\nolimits_{S_{k,t}\subset \hat{S}} \left(\left[\sum\nolimits_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \right] \right. \\
& \qquad \left. + \phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) \cdot \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}} \right) \right) \cdot y^{k,t}_{S_{k,t}} \\
\text{subject to} \quad & \sum\nolimits_{t=1}^{T} \sum\nolimits_{k=1}^{K^t} \sum\nolimits_{S_{k,t}\subset \hat{S}} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \right] \cdot y^{k,t}_{S_{k,t}} \leq I_i^0 \quad \forall i \\
& \sum\nolimits_{S_{k,t}\subset \hat{S}} \; y^{k,t}_{S_{k,t}} = 1 \qquad \forall (k,t) \\
& y^{k,t}_{S_{k,t}} \geq 0 \qquad \forall (k,t),\; S_{k,t} \subset \hat{S}
\end{aligned} \tag{2.4}
$$

The decision variables $y^{k,t}_{S_{k,t},p_{S_{k,t}}}$ [1] correspond to the probability with which bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$ to consumer (k,t) when the full product prices are set at $\bar{p}_i^t$. The discrete price setting allows us to relax these initially binary decisions to continuous variables, resulting in the above linear programming formulation. For this formulation we define the individual consumer propensity-to-buy $\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})$ as the

[1]We define $y^{k,t}_{S_{k,t},p_{S_{k,t}}}$ as being dependent on both the bundle composition and price. However, under discrete pricing, we can enumerate the collection $\hat{S}$ of all bundles at all prices so each $S_{k,t}$ is inherently defined by its corresponding price. Thus for ease of notation we neglect the additional subscript of $p_{S_{k,t}}$ and write the summations over $\forall S_{k,t} \subset \hat{S}$.

probability that consumer (k,t) will purchase item i if their personalized bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$. Unlike the prior exclusive definitions of $\xi(\cdot)$ for formulation (2.3), these propensities-to-buy are defined as follows: $\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})$ captures all of the combinations in which product i can be purchased when only the bundle offer is discounted and all other products remain at full price $\bar{p}_i^t$. For example, if consumer (k,t) is offered bundle $S_{k,t}$ at price $p_{S_{k,t}}$, we can define the probability they purchase product i in some combination of other products S at full price as $\phi_S^{k,t}(\bar{p}_S^t, p_{S_{k,t}})$. Thus the complete probability of (k,t) purchasing i is given by,

$$\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) = \sum\nolimits_{S\subset\hat{S}:S\ni i} \phi_S^{k,t}(\bar{p}_S^t, p_{S_{k,t}}) \tag{2.5}$$

We similarly define the probability $\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})$ with which a personalized bundle $S_{k,t}$ is purchased as,

$$\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) = \sum\nolimits_{S\subset\hat{S}:S\supset S_{k,t}} \phi_S^{k,t}(\mathbf{p}_{S_{k,t}}) \tag{2.6}$$

Note that the above probability includes the scenarios in which the consumer purchases full-priced items in addition to the bundle $S_{k,t}$, meaning that their purchase set S contains $S_{k,t}$. Thus, formulation (2.4) is the offline version of formulation (2.3) in which the entire sample path of consumer arrivals and full price trajectories are both known before the start of the horizon. Therefore, the objective function is an expectation of the total revenue taken over the consumer purchase decisions for a specific known sample path $\{k,t\}_{\forall k=1,\ldots,K^t}^{t=1,\ldots,T}$, and is thus an upper bound on the expected revenue of any online algorithm. Notice that this problem is also subject to inventory constraints defined through initial stock levels I_i^0 for all products $i = 1,\ldots,n$, because it allocates bundle offers according to expected consumer behavior over the entire horizon. By definition of the Clairvoyant, this problem eliminates the need for forward-looking inventory functions because it identifies the optimal bundle $S_{k,t}$ for every consumer (k,t) utilizing its full knowledge of all future arrivals. This benchmark is not actually attainable because it relies on precise future knowledge that is never available to any practical online model. However, since this formulation provides an upper bound on the expected profit for our setting, we will use its objective value

{Clairvoyant} as an "optimal" best-case benchmark against which we measure the performance of all proposed algorithms.

2.3 Approximation Algorithms

The primary source of complexity in our model stems from the calculation of the expected future revenue as a function of the inventory levels. Therefore, our main goal in this section is to develop methods that approximate the $V(\cdot)$ terms in formulation (2.3). We also aim to address the second challenge that arises from the consideration of products with limited stock, which is the incorporation of inventory-based dynamic pricing strategies that optimize the full product prices $\bar{p}_i^t$ over the course of the selling horizon. Thus, we also aim to develop approximation algorithms within a framework that aligns: (i) the bundle offer pricing in our lower-level recommendation system for individual consumers within each period t, with, (ii) the global upper-level full price trajectories.

2.3.1 Multiplicative Approximation Algorithm

We first consider the following approach to our lower-level bundle recommendation problem, which incorporates the value of inventory through a multiplicative penalty on the bundle terms from the objective function of $\{\text{Dynamic}\}_{\forall(k,t)}$. This multiplicative penalty can be viewed as an approximation to the negative counterpart of the dual variables corresponding to the inventory constraints in formulation (2.4) of the Clairvoyant problem. Using this multiplicative approach allows us to maintain the previous trade-offs captured in the objective function of the DP problem in formulation (2.3), but include inventory balancing through a tractable calculation that does not require demand forecasting. In formulation (2.7) we present the general formulation for this multiplicative approximation algorithm, denoted by $\{\text{MultAlg}\}_{\forall(k,t)}$, which requires the full price trajectories $\bar{p}_i^t$ as inputs (the procedure for computing the values of $\bar{p}_i^t$ is below).

$$\begin{aligned}
\{\text{MultAlg}\}_{\forall(k,t)} = \underset{S_{k,t}\subset\hat{S},\ p_{S_{k,t}}}{\text{maximize}} \quad & \left[\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot \psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
& + \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot (p_{S_{k,t}} - \bar{p}_{S_{k,t}}) \cdot \min_{i\in S_{k,t}} \psi\left(\frac{I_i^{k,t}}{I_i^0}\right) \qquad (2.7) \\
\text{subject to} \quad & (1-\epsilon)\bar{p}_{S_{k,t}} \leq p_{S_{k,t}} \leq \bar{p}_{S_{k,t}} \qquad \forall(k,t), S_{k,t} \in \hat{\mathcal{S}}
\end{aligned}$$

The above formulation has a similar structure to the objective function of {Clairvoyant}. However, the key difference between this algorithm and formulation (2.4) lies in the use of the multiplicative inventory penalty $\psi(\cdot)$ to determine bundle composition and pricing using an approach that requires no estimation of future consumer behavior, as was previously the case in the use of the $V(\cdot)$ functions in formulation (2.3). This penalty $\psi(\cdot)$ is a twice-differentiable, monotone increasing, concave function on the interval $[0, 1]$ and takes as input the fraction of remaining inventory $I_i^{k,t}/I_i^0$ at the time of arrival of consumer (k, t). We consider several different forms for this function including linear, $\psi(x) = x$; polynomial, $\psi(x) = \sqrt{x}$; and exponential, $\psi(x) = (1 - e^{-x})$. In this work we consider the joint problem of bundle composition and pricing, and therefore introduce the minimization of $\psi(\cdot)$ over all i in bundle $S_{k,t}$, which penalizes bundles composed of items with low stock and instead promotes products with excess inventory. We also introduce a corresponding set of $\psi(\cdot)$ functions to the individual product purchases to approximately account for the corresponding $V(\cdot)$ functions that would influence those terms in the original dynamic programming formulation.[2] By avoiding any demand forecasting, this approach reduces the recommendation problem to an enumeration over all possible bundles and prices, which is typically small in scale given limitations on the size of $\hat{S}$. However, note that the choice of functional form for the multiplicative penalty is important to the implementation of this algorithm. We show in Section 3.4 that a choice of polynomial $\psi(\cdot)$, which is often

[2]If we directly replaced each $V(\cdot)$ term from the DP in formulation (2.3) with a minimization over $\psi(\cdot)$ we would get: $\sum_{S\subset\hat{S}} \xi_S^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot (\sum_{i\in S} \bar{p}_i^t) \cdot \min_{i\in S} \psi(I_i^{k,t}/I_i^0)$ instead of the first term in formulation (2.7). However, we instead consider a good upper bound on this term without the minimization, which provides us with this approximation algorithm.

a good approach when approximating a function whose true form is unknown, may result in lower empirical performance on the order of up to 6% when compared to the performance of a more sophisticated exponential $\psi(\cdot)$, relative to the full-knowledge Clairvoyant strategy. Nevertheless, we also find that our results are fairly robust to the choice of $\psi(\cdot)$ and still capture 81% of the expected profit of the full-knowledge approach, even in the most inventory-constrained cases.

We further improve on this algorithm by introducing a rolling extension that augments formulation (2.7) to use $\psi\left(I_i^{k,t}/\max\{I_i^t - 1, 1\}\right)$ instead of $\psi\left(I_i^{k,t}/I_i^0\right)$ for all consumers $k = 1, ..., K^t$ arriving during period t. This periodic approach increasingly emphasizes the difference between products with highly depleted stock and those with great excess as the horizon progresses. Note that we use $I_i^t - 1$ as opposed to I_i^t. Intuitively, if we set the denominator of $\psi(\cdot)$ to I_i^t at the start of each period t, the inventory levels of all available products initialize to 100% and become equivalent in terms of $\psi(\cdot)$ for first-arriving consumers $(k = 1, t)$. Therefore, subtracting one unit consistently differentiates the fractions $I_i^{k,t}/\max\{I_i^t - 1, 1\}$. This extension is equally tractable and has improved empirical performance over the approach in formulation (2.7), as shown in Section 3.4.

Calculating Upper-Level Full Price Trajectories

To address our second challenge of aligning upper and lower level pricing strategies, note that the multiplicative algorithm only requires individual full product prices $\bar{p}_i^t$ to make bundle offers. Due to lack of demand forecasting, formulation (2.7) cannot adequately calculate these full price trajectories. Thus, we propose an upper-level method for calculating the time-dependent full price trajectories, denoted by $\hat{p}_i^t$ as they are now estimated quantities, using following formulation:

$$\begin{array}{ll} \max\limits_{\hat{\mathbf{p}}^t,\, \forall t} & \sum\limits_{t=1}^{T} \sum\limits_{i=1}^{n} D_i^t(\hat{\mathbf{p}}^t) \cdot \hat{p}_i^t \\ \text{subject to} & \sum\limits_{t=1}^{T} D_i^t(\hat{\mathbf{p}}^t) \leq I_i^0 \ \ \forall i \end{array} \tag{2.8}$$

We define $D_i^t(\hat{\mathbf{p}}^t)$ as the expected demand for product i during period t based on all current product prices $\hat{\mathbf{p}}^t$, which can be calculated using expected future consumer arrival rates based on historical transactions. Based on the prior definition of $\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})$ from equation (2.5), $D_i^t(\hat{p}_i^t)$ is the expected aggregate demand for product i and incorporates all combinations S in which i is purchased along with other products at full price. Thus, formulation (2.8) is a tractable linear programming problem, as shown in [32]. As is commonly done in practice, we implement this using a rolling approach by periodically re-solving the above LP at the beginning of each period t using updated inventory levels. The output of this upper-level problem provides us with a set of full price trajectories $\hat{p}_i^t$ for all products i in all periods t. By holding these fixed for the duration of a given period t, we can now easily solve the lower-level bundle recommendation problem using the multiplicative algorithm.

Performance Ratios of the Multiplicative Approximation Algorithm

The strength of the multiplicative approximation algorithm lies in the fact that it only assumes broad conditions on the structure of $\psi(\cdot)$ and $\phi(\cdot)$. This eliminates the need for demand forecasting and is thus applicable to the majority of possible demand groups $\hat{S}$ and models $\phi(\cdot)$. However, note that the full price trajectories utilized by our algorithm from the upper-level problem may differ significantly from those selected by the full-knowledge Clairvoyant strategy. Therefore, we define:

$$\alpha_i^t = \frac{\hat{p}_i^t}{\bar{p}_i^0} \ \ \forall i,t, \quad \text{where } \hat{p}_i^t \text{ are full price trajectories chosen by formulation (2.8), and,}$$

$$\beta_i^t = \frac{\bar{p}_i^t}{\bar{p}_i^0} \ \ \forall i,t, \quad \text{where } \bar{p}_i^t \text{ are optimal price trajectories provided by an oracle to the Clairvoyant in formulation (2.4).}$$

In this multi-period setting, for a given sample path $\{k,t\}_{\forall k=1,\ldots,K^t}^{t=1,\ldots,T}$, we show the following result.

Theorem 1. *Given a fixed adversarial sequence of consumer arrivals (k,t) and time-dependent trajectories of full product prices $\hat{p}_i^t$ from formulation (2.8) defined through α_i^t for all products i and periods t, the worst-case competitive ratio of our multiplicative algorithm $\{MultAlg\}_{\forall(k,t)}$ relative to the full-knowledge strategy of $\{Clairvoyant\}$ when* *choosing personalized bundle composition and prices* **as well as** *the global full prices* $\bar{p}_i^t$ *is bounded by,*

$$
\begin{aligned}
1 &\geq \frac{\{MultAlg\}_{\forall(k,t)}}{\{Clairvoyant\}} \\
&\geq \min_{(I^0_{min},x):x\leq 1-\frac{1}{I^0_{min}}} \frac{\frac{l}{I^0_{max}}\sum_{t=1}^{T}\sum_{k=1}^{K^t} R^{k,t}_{min} - \frac{l}{I^0_{min}}\cdot\frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T} M^t}{\beta_{max}\cdot\frac{l}{I^0_{min}}\cdot\int_{x=1+\frac{1}{I^0_{min}}}^{1}\psi(y)dy + 1 - \psi(x) - \frac{l}{I^0_{max}}\cdot\frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T} M^t}
\end{aligned}
$$

The bound parameters are explicitly defined as follows: $I^0_{\boldsymbol{min}}$ *is the minimum initial inventory level across all products* $i \in \hat{S}$*, defined by* $I^0_{\boldsymbol{min}} = \min_{i=1,\ldots,n} I_i^0$*,* $I^0_{\boldsymbol{max}}$ *is similarly the maximum across all initial inventory levels,* $R^{k,t}_{min}$ *is the minimum of the product of propensity-to-buy* $\phi_i^{k,t}(\boldsymbol{p}_{S_{k,t}})$ *with the nominal price discount level* α_i^t *across all* $i \in \hat{S}$ *for consumer* (k,t)*, defined by* $R^{k,t}_{min} = \min_i \phi_i^{k,t}(\boldsymbol{p}_{S_{k,t}})\cdot\alpha_i^t$*,* M^t *is defined as the maximum revenue loss from bundle discounting in period* t *that is defined explicitly in Proposition 4 of Appendix A.2 as* $M^t = \max_{S_{k,t}\subset\hat{S},d_{S_{k,t}}} \sum_{k=1}^{K^t}\sum_{i\in S_{k,t}}\phi^{k,t}_{S_{k,t}}(\bar{\boldsymbol{p}}_{S_{k,t}})\cdot\bar{p}_i^t\cdot(1-d_{S_{k,t}})$*, where* $d_{S_{k,t}}$ *is the bundle discount price ratio* $p_{S_{k,t}}/\bar{p}_{S_{k,t}}$*.*

The proof of this theorem is provided in detail in Appendix A.2. Notice that if we were to remove pricing from formulation (2.7) the resulting problem would identify the most relevant bundle of products to offer, while still allowing consumers to purchase any other combination $S \neq S_{k,t}$ of products where all products (including the bundle $S_{k,t}$) are at full price. Similarly, if we were to remove bundling but continue pricing, the problem would reduce to identifying the best single-item discount offer from among products $i \in \hat{S}$, while allowing the consumer to purchase any other products at their full price. Thus in the combined setting in which both bundling and pricing are removed from the problem, the above bound reduces to the result in [12]. We empirically evaluate the performance of this ratio in realistic scenarios by using actual data from our case studies to generate the results in Figure 2-2 below.

Notice that this bound depends on the choice of the inventory penalty $\psi(\cdot)$ and will vary depending on its functional form. Furthermore, the initial stock levels I_i^0 dictate the extent to which the problem is constrained by inventory, and thus the extremity of the lower bound. Remark that the larger the gap between $I^0_{\min}$ and $I^0_{\max}$, the more conservative the lower bound, as shown in Figure 2-2. Intuitively, a less

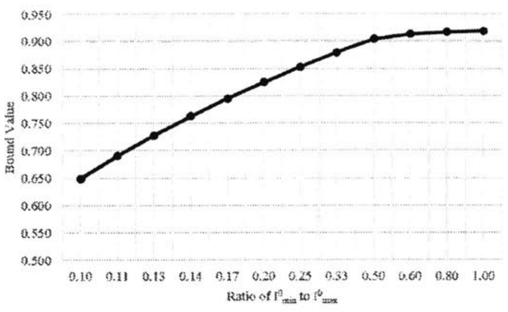

Figure 2-2: This plot shows the empirical value of the bound as a function of the ratio between lowest initial inventory I^0_{min} and the highest level setting I^0_{max} across all products, when all else is held constant.

inventory-constrained problem with higher initial inventory settings will result in a significantly tighter bound as there is inherently less error in the multiplicative approach relative to the full-knowledge strategy due to the fact that the consideration of future consumer behavior becomes less critical. Overall, we find that more inventory-constrained instances with limited discounting opportunity due to initially low full price settings generate the most extremely conservative values of the above bound. However, as we demonstrate in Table 3.8 in Section 3.4, the empirical performance of the bound on the multiplicative algorithm on actual data falls within 7% of its actual performance relative to a full-knowledge strategy across all possible inventory cases.

We conduct an in-depth computational analysis with various functional forms for this multiplicative penalty algorithm, the results of which are summarized in Table 3.5 in Section 2.3.2 and Table 3.6 in Section 3.4 using real industry data from our case studies. We find that the empirically best multiplicative algorithm utilizes an exponential form for the penalty function and obtains 70% of the expected revenue achieved by the full-knowledge strategy in highly constrained inventory cases, which improves to 97% in the less constrained cases. We will refer to this approach as the exponential multiplicative penalty algorithm (EMPA) for the remainder of the work. By implementing the rolling version of the EMPA, we can improve these results to 72% and 97.4%, respectively.

2.3.2 Additive Approximation Methods

The multiplicative algorithm provides a tractable approach with analytical guarantees that does not need to account for demand forecasting. Therefore, in this section we develop a second class of approximation algorithms to use as benchmarks in order to evaluate the empirical value of considering future consumer behavior. Based on well-known methods from the DP literature such as problem decomposition and Lagrangian relaxation, for example as presented in [14], we construct two additive approaches to approximating the inventory functions $V(\cdot)$ in the lower-level bundle pricing and selection problem. In Section 3.4 we show that on average the multiplicative algorithm empirically performs within 1-6% of these additive approximations when compared to the full-information benchmark. Thus, the multiplicative approach provides a much easier implementation method at a very marginal cost in terms of expected revenue.

Separable Item Additive Algorithm (SIAA)

We aim to estimate the functions $V(\cdot)$ efficiently by decomposing the expected future revenue function $V(\mathbf{I})$ of a given inventory state $\mathbf{I}$ into the sum of the expected future revenues $f_i(I_i)$ for each of the items in demand group $\hat{S}$. Thus, we propose the following separable-by-item approximation:

$$\begin{aligned} V_{k+1,t}(\mathbf{I}) &\approx \sum_{i \in \hat{S}} f_i^{k+1,t}(I_i), \text{ where,} \\ f_i^{k+1,t}(I_i) &= \sum_{\tau=t}^{T} \bar{p}_i^\tau \cdot \min\{\delta^\tau D_i^\tau(\bar{p}_i^\tau), C_i^\tau\}, \text{ and,} \\ C_i^\tau &= \{C_i^{\tau-1} - \delta^\tau D_i^\tau(\bar{p}_i^\tau)\}^+, \text{ initialized at } C_i^t = I_i. \end{aligned} \tag{2.9}$$

The values of $D_i^t(\hat{p}_i^t)$ and $\hat{p}_i^t$ in this expression are provided by output of the periodically re-solved upper-level LP problem defined by formulation (2.8) in Section 2.3.1. We define δ^t as the fraction of time remaining in the period t during which consumer (k, t) arrives, implying that $\delta^\tau = 1$ for all periods τ after the current one. Each of the terms in $f_i^{k+1,t}(I_i)$ from the second line of equation (2.9) considers the minimum between the expected demand for product i in that period and its expected available inventory. We capture this by defining inventory levels C_i^τ recursively for all periods τ after the current one and initializing the inventory level at I_i. Thus equation (2.9) provides us with an estimate of expected revenue $f_i^{k+1,t}(I_i)$ from product i over the remainder of the horizon given current inventory I_i. Leveraging this estimation for every product i in the approximation approach from the first line of equation (2.9), we develop a tractable algorithm that depends only on current inventory levels and fixed quantities that are known entirely in advance. We define this method as the SIAA, Separable-Item Additive Algorithm, which allows us to provide each individual customer with a personalized bundle offer in real-time at significantly less computational cost than the original dynamic programming formulation.

Additive Lagrangian Algorithm (ALA)

Based on the above framework we now propose a more sophisticated approach, which we refer to as the Additive Lagrangian Heuristic (ALA). If we consider the SIAA more closely, we observe that the separable-by-item decomposition omits any terms related to the expected revenue from bundle purchases at possible bundle discounts. Therefore, we construct a second additive algorithm that incorporates these additional terms. Recalling the approximation framework from equation (2.9), we propose the following extension to include bundle purchases:

$$
\begin{aligned}
V_{k+1,t}(\mathbf{I}) &\approx \sum_{S \subset \hat{S}} f_S^{k+1,t}(\mathbf{I}), \text{ where,} \\
f_S^{k+1,t}(\mathbf{I}) &= \sum_{\tau=t}^{T} \bar{p}_S^{\tau} \cdot \min\left\{ \delta^{\tau} D_S^{\tau}(\bar{p}_S^{\tau}),\ \min_{i \in S} \{C_i^{\tau}\} \right\}, \text{ and,} \\
C_i^{\tau} &= \left\{ C_i^{\tau-1} - \delta^{\tau} \left(\sum_{S \subset \hat{S} : S \ni i} D_S^{\tau}(\bar{p}_S^{\tau}) \right) \right\}^{+}
\end{aligned}
\tag{2.10}
$$

The extension here is the additional consideration of bundle terms at bundle prices, as captured by $f_S^{k+1,t}(\mathbf{I})$. Instead of solving the LP problem from formulation (2.8), which only outputs single-item trajectories and demand, we formulate an extension in which $\bar{p}_S^t$ are also decision variables. To improve tractability, we further relax this problem by introducing the inventory constraints into the objective function using Lagrange multipliers and ultimately obtain an LP formulation extension of (2.8) that provides the distinct sets of trajectories $\bar{p}_i^t$, for individual products i, and $\bar{p}_S^t$, for combinations S. (In this extended setting $D_S(\bar{p}_S^t)$ is the bundle demand exclusively for bundle S.) Thus for any consumer we can generate an estimate of the cost-to-go $V_{k+1,t}(\mathbf{I})$ using equation (2.10), which includes the expected revenue from both individual items *and* bundle purchases, accounting for the additional revenue not captured by the SIAA due to bundle discounts. This approximation is similarly tractable and in Section 3.4 we show that the ALA achieves empirical results on the order of up to 5% higher than the SIAA in average expected revenue relative to the full-knowledge strategy. Ultimately, these additive approximations provide empirically strong benchmarks against which we can measure the performance of the multiplicative algorithm as opposed to comparing only to the Clairvoyant strategy, which is unattainable in realistic business practice.

2.4 Conclusions

As demonstrated by leading market forecasts, the online channel stands to inherit a significant proportion of the retail market, and is also a rapidly growing avenue of business for many other industries offering personalized experiences such as travel and styling. Gaining the competitive edge this sector is of utmost importance to any participating firm's success. By leveraging personalized consumer information and making relevant offer recommendations, online sellers can develop a loyal customer following and drastically increase their future sales. Furthermore, personalized bundle pricing not only incentivizes consumers to purchase more, but also increases their overall satisfaction with the shopping experience.

We present a novel modeling approach for personalized bundle selection, pricing and recommendation in real-time in an online setting. By incorporating individualized estimates of consumer propensity-to-buy with a long-term profit maximizing framework that capture inventory management, we develop a new analytical problem structure that simultaneously addresses several relevant fields in the existing literature. We develop practical approaches for tractable implementation of our model through two classes of inventory balancing approximation algorithms, for which we provide analytical performance guarantees and present corresponding empirical results from two in-depth industry case studies in Chapter 3.

Bibliography

[1] Shipra Agrawal, Zizhuo Wang, and Yinyu Ye. A dynamic near-optimal algorithm for online linear programming. *Operations Research*, 62(4):876–890, 2014.

[2] Ali Aouad, Retsef Levi, and Danny Segev. Approximation algorithms for dynamic assortment optimization models. 2015.

[3] Fernando Bernstein, A Gürhan Kök, and Lei Xie. Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, 17(4):538–553, 2015.

[4] Gabriel Bitran and René Caldentey. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–229, 2003.

[5] Felipe Caro and Jérémie Gallien. Clearance pricing optimization for a fast-fashion retailer. *Operations Research*, 60(6):1404–1422, 2012.

[6] S Tamer Cavusgil and Shaoming Zou. Marketing strategy-performance relationship: an investigation of the empirical link in export market ventures. *The Journal of Marketing*, pages 1–21, 1994.

[7] Rui Chen and Hai Jiang. Capacitated assortment and price optimization for customers with disjoint consideration sets. *Operations Research Letters*, 45(2):170–174, 2017.

[8] Kris Johnson Ferreira, David Simchi-Levi, and He Wang. Online network revenue management using thompson sampling. 2016.

[9] Guillermo Gallego, Anran Li, Van-Anh Truong, and Xinshang Wang. Online resource allocation with customer choice. *arXiv preprint arXiv:1511.01837*, 2015.

[10] Guillermo Gallego, Anran Li, Van-Anh Truong, and Xinshang Wang. Approximation algorithms for product framing and pricing. Technical report, Working paper, available online as SSRN report, 2016.

[11] Guillermo Gallego and Huseyin Topaloglu. Constrained assortment optimization for the nested logit model. *Management Science*, 60(10):2583–2601, 2014.

[12] Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. Real-time optimization of personalized assortments. *Management Science*, 60(6):1532–1551, 2014.

[13] Vineet Goyal, Retsef Levi, and Danny Segev. Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand. *Operations Research*, 64(1):219–235, 2016.

[14] Jeffrey Thomas Hawkins. *A Langrangian decomposition approach to weakly coupled dynamic optimization problems and its applications.* PhD thesis, Massachusetts Institute of Technology, 2003.

[15] Peng Hu, Stephen Shum, and Man Yu. Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Operations Research*, 64(1):118–134, 2015.

[16] Patrick Jaillet and Xin Lu. Near-optimal online algorithms for dynamic resource allocation problems. *arXiv preprint arXiv:1208.2596*, 2012.

[17] Thomas Kesselheim, Andreas Tönnis, Klaus Radke, and Berthold Vöcking. Primal beats dual on online packing lps in the random-order model. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 303–312. ACM, 2014.

[18] A Gürhan Kök, Marshall L Fisher, and Ramnath Vaidyanathan. Assortment planning: Review of literature and industry practice. In *Retail supply chain management*, pages 99–153. Springer, 2009.

[19] Guang Li, Paat Rusmevichientong, and Huseyin Topaloglu. The d-level nested logit model: Assortment and price optimization problems. *Operations Research*, 63(2):325–342, 2015.

[20] Qian Liu and Stephen Shum. Pricing and capacity rationing with customer disappointment aversion. *Production and Operations Management*, 22(5):1269–1286, 2013.

[21] Siddharth Mahajan and Garrett van Ryzin. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, 45(11):1496–1509, 1999.

[22] Siddharth Mahajan and Garrett van Ryzin. Stocking retail assortments under dynamic consumer substitution. *Operations Research*, 49(3):334–351, 2001.

[23] Ken Moon, Kostas Bimpikis, and Haim Mendelson. Randomized markdowns and online monitoring. *Management Science*, 2017.

[24] Sajjad Najafi, Chi-Guhn Lee, Sami Najafi Asadolahi, and Steve Nahmias. Dynamic pricing under consumer's sequential search. 2016.

[25] Serguei Netessine, Sergei Savin, and Wenqiang Xiao. Revenue management through dynamic cross selling in e-commerce retailing. *Operations Research*, 54(5):893–913, 2006.

[26] Ö Özer and Yanchong Zheng. Behavioral issues in pricing management. *The Oxford handbook of pricing management*, pages 415–460, 2012.

[27] Özalp Özer and Yanchong Zheng. Markdown or everyday low price? the role of behavioral motives. *Management Science*, 62(2):326–346, 2015.

[28] Elli Rapti, Anthony Karageorgos, and Georgios Ntalos. Adaptive constraint and rule-based product bundling in enterprise networks. In *WETICE Conference (WETICE), 2014 IEEE 23rd International*, pages 15–20. IEEE, 2014.

[29] Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039, 2014.

[30] Paat Rusmevichientong and Huseyin Topaloglu. Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research*, 60(4):865–882, 2012.

[31] Davood Shamsi, Marius Holtan, Robert Luenberger, and Yinyu Ye. Online allocation rules in display advertising. *arXiv preprint arXiv:1407.5710*, 2014.

[32] Kalyan T Talluri and Garrett J van Ryzin. *The theory and practice of revenue management*, volume 68. Springer Science & Business Media, 2006.

[33] Huseyin Topaloglu. Using lagrangian relaxation to compute capacity-dependent bid prices in network revenue management. *Operations Research*, 57(3):637–649, 2009.

[34] Zhengliang Xue, Zizhuo Wang, and Markus Ettl. Pricing personalized bundles: A new approach and an empirical study. *Manufacturing & Service Operations Management*, 18(1):51–68, 2015.

Chapter 3

Industry Driven Case Studies

3.1 Introduction

To test the performance of our proposed algorithms we conducted two extensive case studies using data sets from large industry partners. In what follows, we present a summary of our computational results:

1. Airline Data Case Study: We analyzed online ticket transactions from a premier airline with a set of ancillary goods (examples may include in-flight wifi access, priority boarding or seat upgrades) offered in addition to the ticket itself. We found that our models generated an increase in revenue and sales on the order of 3-7% when compared to existing methods.

2. Retail Data Case Study: We analyzed data from a two year selling horizon of a large online e-tailer. By benchmarking against actual pricing strategies used by the e-tailer, we observed an improvement in revenue on the order of up to 10%.

Given that both of these industries operate on very tight margins, these results are very promising for practical business implementation. We also extract innovative business insights in both cases that could help sellers in both industries make strategic pricing and recommendation decisions. We conclude with an in-depth comparison of the two classes of approximation algorithms in Section 3.4 to demonstrate the trade-off between approximation accuracy and tractability.

3.2 Airline Case Study

The increasing trend in online shopping is not only limited to tangible consumer goods and products. The travel industry, which encompasses a vast variety of consumer services such as airlines, cruise lines and hospitality, has undergone massive expansion over the past ten years alone. According to the International Air Transport Association (IATA), the airline industry (when measured by revenue) has doubled over the past decade from $369B in 2004 to approximately $746B in 2014. As a direct result of this, travel products such as airline tickets are becoming commoditized. Travelers are unwilling to pay premiums for these generic services and therefore airline businesses are forced to price very competitively and operate on razor-thin margins. However, despite this commoditization, consumers are increasingly willing to pay for unique experiences, which airlines have begun providing in the form of un-bundled ancillary services that were not previously provided. These ancillary goods are offered to consumers in addition to their ticket and include products to customize their journey before, during and after the flight such as: seat selection and upgrades, VIP lounge access, priority boarding, in-flight wi-fi access, priority baggage handling and various destination-related services such as transportation or tickets to local attractions. Personalization solutions for offering these ancillary services to potential passengers can greatly increase traveler intimacy during their journey and improve their satisfaction with the airline. Therefore, in the context of the airline industry, our modeling goal translates to offering relevant personalized offers consisting of ancillary goods that complement the ticket itinerary the customer is considering. As airlines employ various revenue management methods for setting their ticket prices, we consider the ticket price to be fixed at a nominal value and aim to bundle and discount ancillary goods that will customize and improve the journey for the traveler.

3.2.1 Overview of Data

We analyzed a one-month period of ticket transactions from a premier international airline. This corresponds to approximately 640,000 historical ticket sales, in addition to ancillary item purchases that were linked to each transaction. Each of these ticket transactions corresponds to the arrival and purchase behavior of a specific consumer. As this is a particularly short window of time, we do not have any repeat consumers and thus no details regarding the contexts of previous flight itineraries for any given consumer. Every consumer transaction is described by a set of features that we categorized into the following two types: (i) personal consumer information including airline tier level, mileage balance, miles to next tier, time since joining rewards program, and the number of previous business and economy flights taken; and, (ii) contextual itinerary booking data, which includes the transaction date, fare paid (USD), connection time, time to departure (in weeks), day of travel (of the week), and the number of passengers in the booking. We also had the corresponding purchase history for some of the following ancillary goods: in-flight wi-fi access, premium on-board entertainment, priority security, priority boarding, priority baggage handling, seat upgrades, checked excess baggage, VIP lounge access, gourmet in-flight meals, and, offers of 2,000 and 4,000 bonus miles. The nominal prices at which these unbundled services were offered are summarized in Table 3.1 below. Note also that in this case study, the products are by definition independent of one another since they correspond to very different services which are not necessarily categorized as purely substitutable or complementary.

Given all of this personalized information that captures both individual consumer features as well as context features that describe their itinerary of interest, our goal is to implement our proposed model through the presented approximation algorithms to make a personalized bundle offer consisting of relevant ancillary services for every consumer in this historical arrival sequence. We achieve this by first analyzing the personalized consumer information to develop consumer profile clusters, which we then leverage to estimate individualized models of propensity-to-buy for each

Ancillary Good	Nominal Price
In-Flight Wi-Fi Access	$25
Premium Entertainment	$20
Priority Security	$20
Priority Boarding	$10
Priority Baggage Handling	$25
Comfort Seating Upgrade	$50
Excess Checked Baggage	$45
VIP Lounge Access	$50
In-Flight Gourmet Meals	$15
2,000 Bonus Miles	$100
4,000 Bonus Miles	$200

Table 3.1: This table summarizes the nominal prices at which each of the ancillary goods are typically offered. Remark that by definition these services are independent of one another and priced separately.

consumer. By dynamically using these estimates, we implement our approximation algorithms through simulations on the actual consumer arrival sequence and evaluate the empirical performance of our methods relative to the full-knowledge Clairvoyant recommendation strategy. We present results that demonstrate the effects of personalized pricing, product recommendation and inventory balancing, while also developing practical business insights.

3.2.2 Developing Consumer Profiles

In order to ultimately construct personalized models of propensity-to-buy for individual consumers, we first analyzed the features in the available data. We used the k-means clustering method to develop distinct categories of consumer profiles in order to map the transactions for our demand model estimations. We ultimately established 7 unique consumer profiles that ranged from premium business travelers to leisure individuals; the distribution of consumers across these clusters is shown in Figure 3-1 below. These clusters were constructed using a combination of personalized consumer features and flight itinerary context features. For example, the premium business traveler cluster was identified by features such as: single passenger, short time until departure, departure on a week day, higher tier level and miles balance, higher fare paid, and, a history of previous premium flights. By contrast, a leisure individual traveler profile corresponded to: single passenger, lower tier level, longer time before departure, and, a departure on a Thursday, Friday or Saturday. Notice that because this is a premier global airline, there is a reasonable fraction

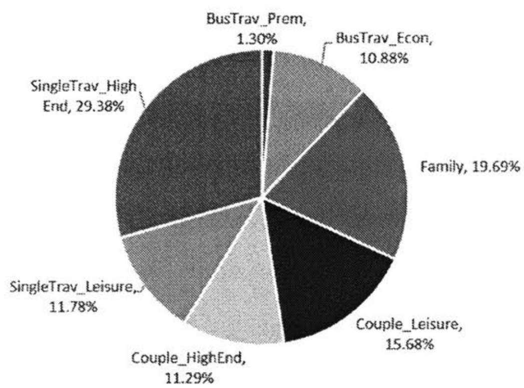

Figure 3-1: This chart shows the distribution of consumers across various persona profile clusters.

of consumers that constitute both business and high end leisure travelers. Having developed these clusters, we were able to automatically identify and map consumers upon arrival to one of these distinct categories. These clusters allow us to identify both consumer profile and relevant itinerary context, which are both necessary for estimating the desired individualized propensity-to-buy models.

3.2.3 Demand Estimation

After developing a consumer to persona profile mapping using the above clustering method, we then needed to estimate individualized models of consumer propensity-to-buy. We accomplished this by fitting a model for every (persona profile, product) pair. Note that the independence property between products is crucial in this case study as it allowed us to define the purchase probability of any bundle $\phi_S(\cdot)$ as the product of the purchase probabilities $\phi_i(\cdot)$ of all the products i in that bundle S; this greatly eased our estimation problem since we did not have to additionally capture and estimate a complicated relationship between the ancillary services.

Instead of constructing segment-level estimation models in which every consumer profile has the same *fixed* propensity-to-buy for any given product, we built a complete set of models for all (persona, product) pairs. Thus, when any consumer arrives, they are mapped to a particular distinct profile for which there is a set of personalized estimated demand models that are then populated with this consumer's unique feature vectors to produce their individual propensity-to-buy for any given ancillary product. This estimate of propensity-to-buy is therefore *unique for every consumer* as they consider any one of the ancillary services. These individual-level models for consumer propensity-to-buy were estimated using logistic regressions on (persona profile, product) pairs, which ultimately produced an exhaustive set of logistic MNL models. These estimated models had an out of sample weighted mean absolute percent error (WMAPE) of 0.12 on average across persona profiles.

3.2.4 Simulation Design

Having developed consumer profile mappings based on clustering and estimated personalized models of propensity-to-buy for all consumers, we design a simulation and corresponding benchmark methods in order to test our bundle recommendation model and observe the effects of personalized pricing, product recommendation and inventory management on expected revenues. To this end, we consider two simulation settings: (1) making bundle offers in the context of unconstrained inventory as is provided in the data set, and, (2) injecting inventory constraints by considering reasonable choices of ancillary services that would have restricted quantities. We now describe our simulation design in detail, which is primarily the same in implementation across both settings but varies regarding benchmark methods and the analysis of the results.

Our goal is to observe the effect of making personalized bundle recommendations to the consumers in this airline data set. Therefore, we take the known consumer arrival sequence as given and iterate over these arrivals, making a personalized offer to each traveler. More specifically, as we consider each consumer in the sequence, our offline clustering approach automatically maps this traveler to one of the 7 persona profiles. Having identified this consumer's profile, the recommendation model solving the lower-level bundle offer problem now uses the corresponding set of individualized demand models for all available ancillary products to decide which bundle is optimal. Note that because we are considered the unconstrained inventory setting, there is no need for any estimation of the functions $V(\cdot)$. Thus, our model reduces to the selection of the profit-maximizing bundle for this specific customer based on their personalized propensities-to-buy for all of the products (which can be easily combined due to product independence to generate the bundle purchase probabilities). Having identified the optimal bundle composition and price, our model proposes this recommendation to the traveler, who may choose to accept the bundle offer, purchase any combination of the ancillary products at their nominal prices, or make no purchase at all. We iterate this process over the entire arrival sequence thousands of times and ultimately analyze the average performance ratios of our recommendation relative to the Clairvoyant approach, which knows the entire arrival sequence in advance. Again we would like to remark that in the absence inventory constraints, our model reduces exactly to the Clairvoyant approach.

In the results under our first unconstrained setting we benchmarked our methods against the current business practice in which *no personalized pricing is offered.* This baseline benchmark is equivalent to a model that always offers all of the ancillary products at their full prices, as provided in Table 3.1. Thus, our results show the direct benefit of implementing personalized pricing techniques. We then extend these preliminary findings to also consider other benchmarks corresponding to lost sales, from which we develop the benefit of personalized recommendations strategies.

3.2.5 Results

The results are organized according to the two settings described above. Under (1) we analyze the effects of personalized bundle offers, then introduce the concept of lost sales to further enhance the impact of relevant product recommendations when consumers are unaware of the existence of ancillary products. We explain the design of (2) but discuss results in Section 3.4.

Value of Personalized Bundle Offers and Business Insights

In the first setting our initial goal was to analyze the effects of personalized pricing and the recommendation system. We considered a *baseline method that offers all of the ancillary products at their fixed full prices* from Table 3.1 in Appendix ??. We implemented our model and observed the average expected gain in revenue over the baseline for all (persona, product) pairs, as summarized in Table 3.2 below. Furthermore, in this simulation scenario our model produced bundle offer outputs in 2.5ms on average. Note that under unconstrained inventory there is no need for the $V(\cdot)$ functions so our model reduces to the unconstrained Clairvoyant problem and selects a personalized myopic profit-maximizing bundle for each customer. Thus, the predicted improvements over the baseline are a direct result of *personalized bundle offers.*

From Table 3.2 we observe that on average the predicted gains in revenue over the baseline varied from 2% to 7% depending on the persona or ancillary product. The overall largest predicted relative improvements in revenues are generated by consumers with low price elasticities such as premium business and high end leisure travelers. Small discounts targeted at these frequent high loyalty consumers result in significantly more conversions and therefore the most expected revenue. By analyzing the corresponding counterpart table of gains in sales volume across all (persona, product) pairs relative to the baseline, we find product-dependent effects. High elasticity personas such as families and lower end leisure travelers see the greatest expected gains in sales volumes across cheaper travel convenience products such as priority security, boarding, and baggage handling, as well as in-flight meals. Conversely, higher-end personas have the highest predicted sales volume gains for more luxe products relevant to frequent travelers, such as VIP lounge access and bonus miles. Intuitively, cheaper products such as in-flight wi-fi have a relatively low elasticity and grow unanimously in predicted revenue across all persona types, particularly among personas

Average Expected Lift in Revenue Over No-Pricing Baseline

	Wi-fi Access	Prem. Enter-tain.	Prior. Secur.	Prior. Board	Prior. Bag Handl.	Seat Up-grade	Ex. Check Bag.	VIP Lounge Access	In-Flight Meals	2,000 Bonus Miles	4,000 Bonus Miles	Total Avg
BusTrav PREM	10.1%	3.2%	—	—	—	—	—	—	—	8.6%	3.5%	2.3%
BusTrav ECON	5.3%	2.3%	1.4%	2.6%	4.4%	8.7%	2.1%	9.7%	3.6%	4.3%	2.5%	4.3%
Family Group	4.2%	2.5%	3.7%	4.0%	1.9%	3.7%	1.9%	5.2%	6.4%	3.9%	2.7%	3.6%
LastMin Group	2.2%	5.0%	3.9%	2.5%	2.1%	3.1%	2.4%	4.3%	5.4%	4.5%	2.5%	3.4%
Couple Normal	7.1%	4.2%	3.3%	2.1%	3.4%	2.5%	2.6%	3.3%	3.6%	4.5%	3.0%	3.6%
Couple HighEnd	8.8%	7.3%	3.8%	2.1%	4.0%	5.3%	3.4%	2.9%	3.8%	7.4%	4.1%	4.8%
Leis Normal	6.5%	2.6%	2.4%	1.8%	2.6%	4.6%	3.6%	3.2%	3.8%	4.4%	2.8%	3.5%
Leis HighEnd	9.8%	8.0%	2.8%	1.9%	3.8%	9.1%	4.3%	6.5%	3.7%	8.8%	5.3%	5.8%
Total Avg	6.7%	4.4%	2.7%	2.1%	2.8%	4.6%	2.5%	4.4%	3.8%	5.8%	3.3%	3.9%

Table 3.2: This table shows the lift in revenue from implementing our personalized pricing and recommendation model over the baseline benchmark in which all products are offered to all arriving consumers at full prices. (Note: some products are not offered to premium business travelers because they are included in their tier level benefits.)

with less price sensitivity that are easily converted with slightly discounted offers. These insights provide potential marketing and pricing strategies that could improve revenues and sales volume if used in the right combinations for (persona, product) pairs.

Impact of Context on Personalization

We also assess our model's ability to distinguish between changes in personalized consumer features and itinerary contexts by analyzing the differences in the average offers made in the following two scenarios: (i) considering the same customer booking two different itineraries, and, (ii) considering two different customers interested in the same itinerary.

In scenario (i), due to lack of purchase history, we generate repeat consumers with similar personal features but different ticket itinerary contexts. The resulting pair of vectors has relatively constant personal features such as tier level and miles balance, but itinerary features such as ticket fare, day of departure and time to departure vary. We find that our model recommends bundles with different compositions but similar discounts. For example in one such simulation, it categorizes the first context as a business trip and offers in-flight wi-fi and lounge access at a 5.2% discount, while it recognizes the second trip as leisure and offers in-flight entertainment and seat upgrades at a 6.1% discount. Since the personalized consumer features were held constant, the consumer's price elasticities stayed relatively constant over time and therefore the discount remained similar across this scenario. Thus, the primary benefit in this setting comes from the model's ability to *identify the significance of itinerary context* in the absence of major changes in personal features.

Under scenario (ii) we discover the converse effect. For example, we can compare one customer of high tier level with historical premium flights to another passenger with lower tier level traveling in a group. The first customer is offered in-flight wi-fi and 2,000 bonus miles at a 1.8% discount, whereas the second customer is offered in-flight meals and priority boarding at a 6.7% discount. The first customer has low price elasticities across all products (on average below -1) and is recommended business-related products, whereas the family traveler has much higher price elasticities (on average between -2 and -3) and receives a greater discount on products convenient for travel with a group. This demonstrates that the model not only identifies context-relevant items but also maximizes expected profit through personalized pricing.

Value of Relevant Product Recommendations

We also objectively analyze the enhanced effect of product recommendation by introducing a parameter α, which is defined as the proportion of consumers who are unaware of the existence of ancillary products and hence do not consider them at all. This is quite common in the travel industry, such as in cruise lines, where there is often an abundance of products that are not explicitly offered to consumers during their online browsing process resulting in loss of potentially interested consumers. Notice that $\alpha = 0$ corresponds to the setting in which all consumers are aware of all ancillary products and there are no lost sales, which is precisely the previous setting from Table 3.2. Thus, the results for any fixed setting across varying levels of α explicitly quantify the expected improvement from *relevant product recommendations.*

We consider the *same baseline method as before* and implement our myopic profit maximizing recommendation model without inventory constraints. We summarize our predicted lifts in revenue over the baseline in Table 4.6 and note that in this expanded setting with the lost sales included, our model still produced output in under 3ms.

Average Expected Lift in Revenue Over No-Pricing Baseline Method

α-level	No Constraints	20% Max	15% Max	10% Max	By Product	"Bundle-Only"
$\alpha = 0$	8.04%	6.11%	4.21%	3.51%	4.83%	4.15%
$\alpha = 0.05$	9.48%	6.09%	5.27%	3.49%	6.18%	5.74%
$\alpha = 0.10$	13.23%	6.23%	5.30%	3.80%	5.66%	6.05%
$\alpha = 0.15$	14.92%	7.74%	5.89%	3.66%	6.89%	6.24%
$\alpha = 0.20$	18.83%	7.24%	6.64%	5.49%	7.94%	7.61%

Table 3.3: This table summarizes the lifts in revenue over the no-pricing benchmark in various scenarios of lost sales (α) ranging from 0 to 20%. When we compare across varying α levels we see the benefit of product recommendation, and as we compare across a fixed α we see the expected improvement from personalized pricing.

Each of the columns in Table 4.6 corresponds to a simulation setting in which we have imposed limitations on our recommendation model. For example, the column "20% Max" corresponds to comparing our recommendation model to the baseline in the case where no product in the bundle offer is discounted by more than 20%; this definition similarly extends to the columns "No Constraints", "15% Max", and "10% Max". In these scenarios we reasonably limit discounts for all products and observe that the predicted improvement in revenue over the baseline is on the order of 3-8% across all possible cases of lost sales, captured by the varying values of α. We found a similar trend in the corresponding results for expected lifts in sales volume on the order of 2-3%. The column "Discount by Item" imposes limitations depending on the full prices of the products; for example, cheaper products are only discounted up to 10%, but more expensive ones are discounted up to 15-20%. The "Bundle-Only" column is the case where the consumer is offered an optimal bundle by our model, but they can only purchase *any other subset of the bundle at full price.* This corresponds to the realistic setting where there is a vast number of ancillary products and the consumer only considers those displayed to them at checkout. Interestingly, the expected improvement in this case is comparable to the "15% Max" scenario because a consumer's propensity-to-buy is typically highest for the set of products selected by our model; therefore, disregarding the products outside this relevant bundle does not heavily impact overall expected sales volume or revenue. Thus, we conclude that in reasonable discount-limiting scenarios the expected gain in revenue from our personalized pricing strategy is on the order of 5-6%. By definition, higher α values indicate that a greater proportion of the population is unaware of ancillary products. The results in Table 4.6 are robust to changes in α and by analyzing the symmetric results for sales volume we observe that these trends are consistent across both metrics. Therefore, the predicted improvements on the order of 2-3% over the baseline from the lowest $\alpha = 0$ level to the highest $\alpha = 0.2$ level are a direct result of exposing consumers to products of which they are otherwise unaware through *personalized and relevant product recommendations.*

Value of Inventory

We lastly consider setting (2) to assess the validity of the approximation algorithms presented in Section 2.3. While inventory is not inherent to this data set, we consider a subset of ancillary products that would reasonably be limited such as VIP lounge access, on-board wi-fi, gourmet meals, excess checked baggage and seat upgrades. We introduce initial inventory levels at quantities that are proportional to the length of the consumer arrival sequence. The simulation is identical to setting (1), except that we consider an inventory-constrained problem across this smaller set of ancillary products. Instead of solving a myopic personalized profit-maximizing problem, we now solve our original DP problem using both classes of algorithms and observe the average percentage of Clairvoyant revenue they obtain. However, we still do not solve the upper-level problem of time-dependent full price trajectories due to the nature of the data and current industry practice. We present our results and a detailed discussion comparing the algorithms in Table 3.6 in Section 3.4.

3.3 Retail Case Study

In this second case study we analyze data from a major U.S. e-tailer over the two year sales period from July 2011 to September 2013. We were provided with point-of-sale transaction data for electronic fulfillment orders across 312 departments which totaled approximately 13M customers and over 34M transactions. This data consisted of order information defined through customer IDs, transaction IDs, SKU numbers, fulfillment center codes, prices at time of purchase, costs at time of purchase, transaction types, and dates and times of purchases. In parallel we also analyzed the corresponding inventory data for the same period across the electronic fulfillment centers (EFCs) responsible for these online orders. The key to this case study lies in the fact that the products are no longer inherently independent as in the airline case, and inventory plays a significant role as we focus on considering seasonal goods with impending periods of steep markdowns. We believe that the resulting computational study clearly demonstrates the continuously robust performance of our various algorithms.

3.3.1 Overview of Data

We briefly summarize the available data and our modeling approaches regarding personalization demand estimation. While this point-of-sale data set provided a wealth of time-series information in the form of consumer purchase histories, it severely lacked in personalized features outside of customer IDs, which linked consecutive transactions. In comparison to the airline case study data set, we did not have access to any individualized consumer information outside of historical product purchase history. Our ultimate goal, similar to the previous case study, was to simulate the performance of our various algorithms against benchmark methods and observe the effects of personalization, pricing and recommendations on the e-tailer's expected revenue.

3.3.2 Developing Personalization Metrics

In order to adapt and test our algorithms in the setting of this case study, we first needed to construct personalized consumer features that were not present innately in the data. Given that our only distinctly individual identifying factors were customer ID and transaction ID, we analyzed the data set and recorded each consumer's cumulative number of visits at the time of each of their transactions, along with their corresponding total cumulative expenditure up until that time. By considering the behavior of these metrics across all consumers over the two year period, we developed a time-dependent consumer loyalty mapping consisting of low, medium and high frequency categories; this mapping allowed us to categorize a consumer at the time of any of their transactions based on these two metrics. At the time of any consumer's arrival, they would be mapped into one of the three resulting loyalty groups as follows: (i) low frequency consumers (loyalty group 1) had no previous purchase history in this two-year selling period and accounted for 77% of all of the transactions in the data set; (ii) medium frequency customers had made at least one previous purchase, but had currently spent below the mean cumulative expenditure (across the whole consumer population) at the time of their arrival and accounted for 17.5% of all transactions; finally, (iii) high frequency customers also had at least one historical transaction, but had spent over the mean cumulative expenditure amount at their time of arrival and accounted for 5.4% of all transactions.

Note that consumers could change loyalty groups over time as they returned for more purchases; every transaction mapped to a given consumer ID was assigned a corresponding loyalty category flag, which was time-dependent and varied depending on that consumer's progression in visits and expenditure. Thus, a returning consumer in the data set was automatically mapped to the medium frequency loyalty group in their second transaction since a prior purchase indicated that this was their second shopping visit. In Figures 3-2 and 3-3 we see the time-series behavior of the middle and high frequency consumer loyalty groups with respect to their cumulative expenditure over the two year selling horizon.

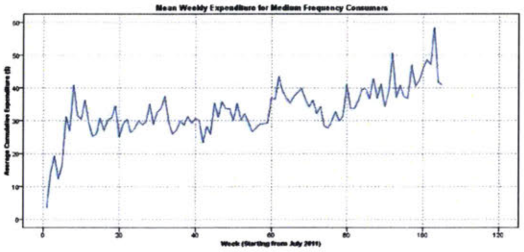

Figure 3-2: This plot shows the mean expenditure for medium loyalty repeat customers.

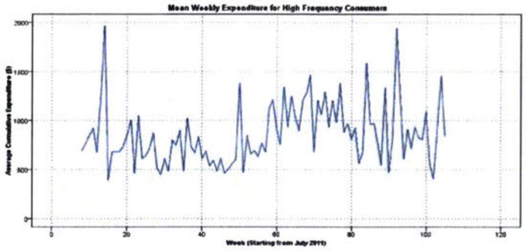

Figure 3-3: This plot shows the mean expenditure for high loyalty repeat customers.

We mapped these loyalty groups to every transaction in this time series fashion, as opposed to simply assigning every consumer just one general loyalty category, so that our personalized demand estimation could learn only from past purchases as it also would in a realistic setting. For example, the entirety of the data set may indicate that a particular consumer ultimately reaches high frequency by the end of the two year period; however, any dynamic recommendation model would not know this during their first or second visit and must therefore only use an estimate of their personalized propensity-to-buy that is based on the historical information known to the model at that time, which would sort them into a low or medium category for earlier transactions.

3.3.3 Choosing Demand Groups

Having established metrics for personalization, we next had to narrow our focus from 312 potential retail departments to select relevant and interesting demand groups to analyze. We ultimately chose to consider seasonal home decor products because their actual price trajectories, as determined by the e-tailer, included steep clearance periods at the end of their respective selling seasons.

Even after selecting this particular department, we were faced with transactions ranging over thousands of SKUs to choose from and initially shrank our consideration set to only the top 500 products based on historical purchase frequency over our two year period. We were then able to extract meaningful combinations of related products by utilizing association rule learning. This particular branch of machine learning is a subset of collaborative filtering that is primarily used for finding groups of items that are frequently purchased together and constructing probabilistic implications, known as association rules, based on these historical transactions. This methodology was originally applied in market basket analysis for advertising and shelf-space optimization, particularly with point-of-sale data in supermarkets. Growing interest in this field led to the development of many efficient algorithms, such as Apriori ([1]) and FP-Growth ([4]), that allow for easy computation of frequent item sets. By leveraging these computational approaches we were able to narrow our scope to a set of approximately 25 demand groups, each of which consisted of a group of products united by a common holiday or seasonal theme such as Valentine's Day, St. Patrick's Day, Halloween, Patriotic, Autumnal, Western or Coastal. However, a more close analysis of these demand groups demonstrated that many of them were interconnected in historical purchases across this department; for example, in Figure 3-4, we see that demand group (winter holiday themed) and demand group 2 (snow themed) are fairly related. Therefore, in order to separate these products into more manageable sized

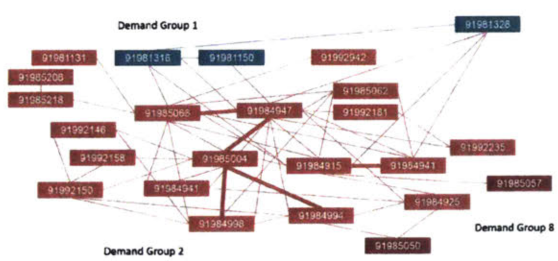

Figure 3-4: This is a visualization of the joint historical transactions between multiple demand groups.

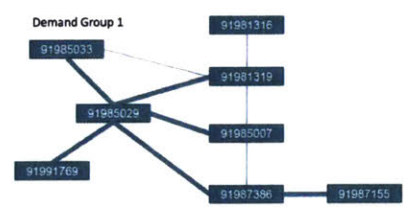

Figure 3-5: This visualization focuses on the most frequent interactions between products in demand group 1.

groups for analysis, we employed association rule learning algorithms again to assess the strength of the connections between products. In both Figures 3-4 and 3-5 we see that certain products (indicated by a darker connecting line) were purchased together significantly more frequently than others. Thus, by leveraging the outputs of these machine learning algorithms we ultimately selected five distinct demand groups, based on these frequency of purchase thresholding policies, for our algorithm testing.

3.3.4 Demand Estimation

After extracting relevant and meaningful demand groups for analysis and preliminarily preparing the data by developing personalized consumer features, we then proceeded to estimate out individualized propensity-to-buy models. However, unlike in the airline case study where we had a core and ancillary product structure, we did not have any information regarding lost sales. Our data set consisted only of historical consumer purchases, and therefore we did not have any records of consumers arriving and choosing not to purchase specific items. Furthermore, the long-term selling horizon we considered required consideration of time-dependent information which we included in the form of loyalty values, holiday and seasonality flags, and nominal price variations as decided by the e-tailer. For each consumer loyalty group we then fit a set of demand models corresponding to each of the product demand groups under consideration using the approach in [3]. The out of sample WMAPE in this case study was approximately 0.40 when averaged across all loyalty and demand group pairs. While this is significantly higher than in the airline case, we feel that it is a reasonable result due to the lack of personalized consumer features for which we developed our own metrics, as well as the lack of historical lost sales. Furthermore, the estimated coefficients across all of the models intuitively corresponded to realistic consumer choice behaviors. For example, the coefficients related to promotion periods and seasonality were all strongly positive, reinforcing the idea that peak selling periods and popularity drive consumers to have a higher propensity-to-buy for products. Conversely, steep clearance periods resulted in negative coefficients; this result corresponds to the scenarios in which the prime life span of the seasonal good has expired (for example the holiday has passed) and therefore consumer willingness-to-pay drastically decreases. As in the airline case, this estimation method provided us with a model for every consumer (loyalty group, demand group) pair, which we could leverage dynamically in our simulation to make truly personalized bundle offers to every individual consumer using their particular loyalty features, as well as the time-dependent context features such as holiday weeks or promotion periods.

3.3.5 Simulation Design

We designed a similar simulation to the airline case to test the efficacy of our proposed algorithms relative to several benchmark pricing methods. For every demand group under consideration, we simulated the historical arrival sequence of all consumers who purchased products in that group during our two-year consideration period. For each historical transaction we aimed to use the personalized propensity-to-buy models that we estimated along with the historical inventory levels at the time of that consumer's arrival in order to make them a personalized bundle offer. Having been presented with this discounted offer the consumer could choose to accept it, choose to purchase any combination of the products in the demand group at full price, or choose to purchase nothing at all. By averaging over this consumer arrival sequence for each demand group thousands of times, we ultimately measured two performance metrics: (i) the average percentage of full-knowledge Clairvoyant profit achieved by each algorithm, and (ii) the average conversion rate (percentage of offers that resulted in a purchase) of each algorithm. Note that these averages were taken both over the arrival sequences for each demand group, as well as over all five demand groups under consideration.

3.3.6 Results

Our empirical results are divided into two discussions: (1) the effects of personalization and dynamic pricing, and, (2) the value of inventory balancing.

Value of Personalization and Dynamic Pricing & Business Insights

We initially considered the bundle recommendation problem in the unconstrained inventory setting in order to objectively measure and *emphasize the impact of personalization* and dynamic pricing on expected revenue, while implementing the upper-level problem framework for determining the full prices of all products in all periods. To develop these results we introduce three relevant benchmarks: (i) the "actual" pricing strategy that parallels the airline case and offers every product in the demand group at its historical full price at the time of each consumer's arrival; (ii) a rolling LP method that periodically re-optimizes the full prices for all products based on segment-level expected future demand, then offers all products at these optimized fixed prices in a given period t to all arriving consumers; and, (iii) an un-personalized version of our myopic recommendation model that uses segment-level consumer features to make bundle offers to each arrival. The results are presented as the average expected percentage of Clairvoyant revenue attained by each pricing strategy and are summarized in Table 3.4.

Effect of Personalization on Empirical Performance

Model	Personalization	Percent of Clairvoyant Revenue
Actual Historical Prices	X	88.2%
Rolling LP Model	X	95.5%
Segment-Level Dynamic Model	X	96.8%
Personalized Dynamic Model	✓	98.5%

Table 3.4: This table summarizes the empirical performance of all the benchmarks for the unconstrained setting in the retail case as a percentage of the expected Clairvoyant revenue, averaged across all demand and loyalty groups.

Note that our model does not achieve 100% of the Clairvoyant profit due to discrepancies in the pricing of the upper-level problem. The output of personalized bundle offers was produced on average in approximately 3ms. We find that on average employing a dynamic pricing strategy over a static approach, such as a rolling LP benchmark, improves expected revenue by 1.3%. Furthermore, leveraging *personalized models* of propensity-to-buy to make bundle offers *increases the expected revenue by an additional 1.7%* over a generic segment-level strategy. The overall improvement over the current pricing strategy is on average on the order of 10%, which is very substantial in a thin-margin setting such as retail.

Our loyalty analysis shows that majority of online consumers in a product category are one-stop shoppers ($\approx$40% within our selected demand groups). While these can be potentially converted using recommendation systems, the real focus of online etailers should be on *improving* conversions of higher frequency customers. The results from Table 3.4 present the objective benefit in expected revenue from personalized dynamic pricing strategies. We found that higher loyalty consumers provided the greatest expected lifts in revenue over the "actual" pricing benchmark within each demand group. Furthermore, as shown in Figures 3-2 and 3-3 in Appendix ??, these consumers spend substantially more than other customers and have a larger source of historical data from which our model can develop more tailored bundle offers. Furthermore, these consumers have lower price elasticities and by definition spend significantly more across the selling horizon. Thus we conclude a similar result to the airline case: on average, higher frequency consumers respond the most effectively to personalized discounted prices, and thus should be the primary target audience for recommendation systems aiming to raise expected revenues and conversions.

Value of Inventory

We now expand our results to the inventory-constrained setting inherent in the data set in order to analyze the practicality and performance of our approximation algorithms from Section 2.3 relative to benchmark methods and the Clairvoyant strategy. The "actual" and rolling LP benchmarks remain the same in this setting. We additionally introduce the myopic heuristic benchmark, which offers the personalized myopically profit-maximizing bundle to each consumer as in the unconstrained case. We consider two metrics of performance for each method: (i) the average expected percentage of Clairvoyant revenue achieved across all loyalty and demand groups, and, (ii) the average conversion rate. The resulting empirical performance ratios are summarized in Table 3.5 below.

In this setting where we implemented our approximation algorithms that depended on inventory levels, the output of personalized bundle recommendations was produced on average in 15ms, which is still very efficient, as is necessary for implementation in an online setting. These results show that on average all of the methods perform relatively well: the ALA obtains 97% of the expected revenue achieved by the full-knowledge strategy, the SIAA reaches 93% and the multiplicative algorithm reaches 91%. The 4% performance gap between the additive methods is precisely the esti-

Constrained Inventory Results Across All Algorithms

Model	Percent of Clairvoyant Profit	Conversion Rate
Actual Historical Prices	83.2%	1.5%
Rolling LP Model	84.6%	3.1%
Myopic Heuristic	87.9%	4.2%
Exponential Multiplicative Algorithm	91.5%	6.0%
Separable-Item Algorithm (SIAA)	93.4%	6.6%
Lagrangian Algorithm (ALA)	97.5%	7.8%
Clairvoyant Model	100%	8.6%

Table 3.5: This table summarizes the empirical performance in expected revenue of all the proposed algorithms for the retail case study as a percentage of the full-information Clairvoyant revenue, averaged across all demand and loyalty groups.

mation difference accounted for by the additional bundle terms included in the ALA at discounted prices. We also observe an average expected gain of 5.5% in revenue over the myopic approach by accounting for inventory balancing and future demand in the SIAA. Furthermore, the overall average improvement in expected revenue from our best algorithm compared to the "actual" historical pricing strategy from the data set is 14% across these demand groups. Note that the while the best multiplicative algorithm (EMPA) is slightly outperformed by the additive methods in this highly inventory-constrained setting with steep markdown periods, it still performs within 9% of the full-knowledge strategy and within 6% of the best additive approach. Furthermore, the EMPA is significantly easier to implement and maintains a very close empirical performance relative to the ALA even in this challenging setting. Note that Table 3.5 illustrates the fact that the SIAA and ALA are empirically effective benchmarks that perform well relative to the full-knowledge strategy, but are much more reasonable for comparison to the EMPA because the Clairvoyant is not actually attainable in any practical setting. Furthermore, as shown by the range of inventory constrained cases above, these results demonstrate the significant benefit of inventory management through personalized recommendations by bundling items at a lesser discount ahead of the markdown period in order to preserve already narrow margins.

3.4 Comparisons

We conclude by presenting comparisons between the relative performances of our algorithms, as well as the empirical behavior of their analytical guarantees under various inventory settings.

Comparison of Approximation Algorithms

The first set of comparisons, summarized using airline case data in Table 3.6 below, show the expected percentage of Clairvoyant revenue achieved on average by each algorithm. As described in Section 3.2.5, we introduce inventory constraints in the airline data on ancillary products for which this is realistic: VIP lounge access, in-flight wi-fi, gourmet meals, excess checked baggage and seat upgrades. Furthermore, we implement the bi-level framework and also determine the full price trajectories of all ancillary products, which are initialized at the values in Table 3.1. Each column in Table 3.6 corresponds to the initial inventory level of the ancillary products as a function of the total number of consumer arrivals in the data set as described in Section 3.2; furthermore, we define the column "unlimited" as having a higher initial stock of each product than there are consumer arrivals, meaning that none of the products can ever be consumed entirely.

The "actual" prices benchmark corresponds to the prior baseline that offers all the available inventory-constrained ancillary products at their full prices. We include two additional sets of hybrid benchmarks based on (i) a threshold policy, and, (ii) an automated procedure. In approach (i) the hybrid algorithm makes all recommendations based on the EMPA until one of the products' inventories is depleted by 20%, after which all recommendations are made using the SIAA. In hybrid approach (ii), which we consider with two parameter settings, we employ a variant of the hybrid method in [2] where γ is a multiplicative weighting factor applied to the objective function of the offer chosen by the SIAA when compared to the offer selected by the EMPA. At greater values of γ, the SIAA recommendation is made more frequently. We discuss the implications of these hybrid resuls in more detail below in conjunction with Table 3.7. For robustness, we conducted this set of simulations on the retail

Algorithm Comparison on Airline Data as Percentage of Clairvoyant Revenue

	Initial Inventory Level					
Algorithm	**Unlimited**	**100%**	**90%**	**80%**	**75%**	**50%**
"Actual" Prices	85.1%	82.6%	78.4%	76.4%	73.1%	60.7%
Re-Optimized Rolling LP	90.4%	88.4%	85.7%	83.8%	81.5%	65.2%
Linear Mult. Penalty	94.2%	92.3%	87.6%	84.1%	78.7%	63.8%
SIAA (Separable-Item Additive)	**96.9%**	**96.1%**	**95.3%**	**92.2%**	**88.6%**	**76.6%**
Polynomial Mult. Penalty	95.1%	92.3%	89.3%	84.8%	79.9%	64.7%
EMPA (Exp. Mult. Penalty)	97.0%	94.8%	92.4%	88.2%	84.2%	69.2%
Rolling Exp. Mult. Penalty	**97.4%**	**95.6%**	**94.2%**	**91.1%**	**87.3%**	**72.8%**
ALA (Additive Lagrangian)	**97.9%**	**97.4%**	**96.1%**	**93.5%**	**89.8%**	**79.5%**
Threshold Hybrid	97.3%	95.7%	94.4%	91.3%	88.9%	74.1%
Automated Hybrid: $\gamma = 1.5$	97.3%	95.7%	94.5%	91.7%	88.1%	73.7%
Automated Hybrid: $\gamma = 2$	**97.4%**	**95.8%**	**94.6%**	**91.9%**	**88.4%**	**75.3%**

Table 3.6: This table summarizes the performance gaps of the proposed additive and multiplicative algorithms, as well as some hybrid algorithms, in the airline case study in percent of expected revenue attained relative to the full-knowledge Clairvoyant strategy.

data and observed symmetrical results.

From Table 3.6 we can conclude that the rolling implementation of the EMPA performs within 3-6% of the full knowledge strategy and within 1-2% of the ALA in less constrained inventory settings; furthermore, it outperforms the EMPA that uses initial inventory levels I_i^0 by 1-3% across all inventory cases. We observe that the SIAA begins to slightly outperform the rolling EMPA in increasingly more inventory-constrained settings, which is fairly intuitive: as it becomes more important to avoid inventory-related costs, the difference in approximation accuracy between the additive approaches and the multiplicative penalty becomes increasingly greater. However, the multiplicative method requires no re-optimization and is easy to implement compared to the ALA, while only under-performing by a margin of up to 5.6% in worst cases in while still achieving on average at least 87% of the expected Clairvoyant revenue in reasonable inventory settings, and at least 73% in the most constrained case. This marginal trade-off in empirical performance is largely offset by the practicality of the multiplicative approach, as well as the fact that in less constrained settings it performs within 1% of the best additive method. Finally, Table 3.6 also demonstrates the value of choosing the correct functional form of the inventory penalty function $\psi(\cdot)$ depending on the data. For example, in this airline case, the polynomial form of the multiplicative penalty under performs on average by approximately 3% compared to the multiplicative algorithm using the exponential penalty function across all inventory scenarios. Furthermore, these gaps in performance grow from 1.9% to up to 4.6% as the problem becomes more inventory constrained; thus, using an increasingly sophisticated form of $\psi(\cdot)$ results in a multiplicative algorithm with stronger empirical performance.

The results in Table 3.6 provide an empirical foundation for understanding the performance differences between the algorithms relative to the Clairvoyant "optimal", but it is also important to gain an insight into which settings each algorithm is best suited for. For this purpose, we introduced the hybrid methods in these results, which alternate between recommendations from both the EMPA and the SIAA. It is clear from Table 3.6 that under mildly constrained inventory settings, all of the hybrids behave essentially the same way as the EMPA. It is in the more constrained inventory cases that we begin to see a greater gap between the performance of the hybrid methods and the EMPA, due to the incorporation of demand forecasting captured in the additive approach. To gain a better understanding of the magnitude of this effect, we provide a second set of hybrid results in Table 3.7 below.

Percentage of Offers Driven by Additive Algorithm (SIAA)

	Initial Inventory Level					
Algorithm	**50%**	**20%**	**15%**	**10%**	**5%**	**2%**
Threshold Hybrid (SIAA after 20%)	16.8%	52.6%	75.7%	87.1%	95.4%	98.3%
Automated Hybrid: $\gamma = 1.5$	31.6%	63.7%	81.9%	90.2%	94.8%	97.5%
Automated Hybrid: $\gamma = 2$	38.5%	68.8%	84.2%	92.8%	96.4%	98.6%

Table 3.7: This table summarizes the percentage of personalized bundle offers made by each of the hybrid algorithms that are selected by the SIAA (as opposed to the EMPA) in the airline case study.

The focus of Table 3.7 is to observe the percentage of personalized bundle offers that are made by the SIAA (as opposed to the EMPA) in the hybrid algorithms under increasingly more constrained inventory settings. Notice that in the 50% that column Tables 3.6 and 3.7 share in common, the SIAA outperforms the EMPA by 4% and accounts for up to 40% of the recommendations in the hybrid method with parameter $\gamma = 2$. Furthermore, the results of Table 3.6 demonstrate that the EMPA becomes less accurate than the SIAA by a growing margin as the problem becomes increasingly more inventory constrained. Supplementing this effect with the computations from Table 3.7, we find that in the most tightly constrained scenarios, the SIAA recommendations represent 75-99% of the offers made in the hybrid approaches. The joint results of these two table indicates that the solution quality of the EMPA deteriorates at a greater rate than the offers selected by the SIAA, and that in the most tightly constrained inventory scenarios, all hybrid methods ultimately resolve to implementing the SIAA. We can conclude that this indicates that the SIAA provide significantly better quality results under highly constrained inventory settings, whereas the multiplicative methods are more efficient and virtually indistinguishable in performance relative to additive methods when initial inventory levels are high.

In addition to algorithm performance and suitability, we also assess the difference between the actual composition and pricing of bundle offers made by the ALA and EMPA, by analyzing the most common bundle offers on average for each persona. We found that the multiplicative approach typically recommends more expensive products at a 1-2% steeper discount than the ALA. Consider the following case of a business traveler flying in economy: the ALA recommends seat upgrades (\$50) and VIP lounge access (\$50) at an average of a 2.4% discount. By contrast, the EMPA recommends VIP lounge access (\$50) and 2,000 bonus miles (\$100) at an average discount of 4.8%, which is an increase of 2.4% in average discount amount and \$45.20 in total offer amount over the ALA. This indicates that rougher inventory estimates in the EMPA generate bundles with a greater emphasis on myopic profit maximization that offset more expensive product offers with higher discounts to increase consumer propensity-to-buy. However in expectation, these two methods generate similar expected revenues that are on average within 10-13% of the full-knowledge approach in less inventory constrained cases.

Empirical Comparison of Analytical Guarantees

Finally, we discuss the difference between the empirical analytical guarantees and the practical algorithm performance in the multiplicative approach. Consider the following airline case results for the EMPA in Table 3.8 in which we now also implement the upper-level full pricing problem (similar tables for polynomial and linear multiplicative penalties echo the trends observed here). The first row shows the average expected percentage of Clairvoyant revenue achieved by the EMPA from empirical data simulations. Note that these ratios slightly decrease from the results in Table 3.6 due to discrepancies between the upper-level solutions of formulation (2.8) and the Clairvoyant. The second row represents the corresponding data-driven values of a path-dependent lower bound on the performance ratio between $\{\text{MultAlg}\}_{\forall k,t}$ and $\{\text{Clairvoyant}\}$, which is derived in the proof of Theorem 1 in Appendix A.2 and depends on the trajectory of inventory levels I_i^t for a fixed consumer arrival sequence $\{k,t\}_{k=1,\ldots,K^t}^{t=1,\ldots,T}$.

Empirical Performance of the EMPA

	Initial Inventory Level				
Exp. Penalty Function	**Unlimited**	**100%**	**90%**	**80%**	**75%**
Performance Ratio from Data	91.6%	88.0%	84.3%	82.0%	76.6%
Posterior Bound on Ratio	84.8%	79.4%	75.8%	71.3%	65.6%
Prior Bound on Ratio	80.4%	74.7%	70.9%	65.6%	58.7%

Table 3.8: This table presents the average expected percentage of Clairvoyant revenue attained by the EMPA as well as the empirical values of its analytical guarantees that are path-dependent (posterior bound) and path-independent (prior bound).

This intermediate bound is less conservative than the worst-case result of Theorem 1, which we refer to as the prior bound in the third row. The results in Table 3.8 demonstrate that the relative gaps between the algorithm's empirical performance ratios (first row) and the posterior bound (second row) grow increasingly as the problem becomes more inventory-constrained; we observe the same effect between the posterior bounds (second row) and the prior bounds (third row). These effects are also visualized in Figure 3.8 above. Note that the empirical performance ratios relative to the Clairvoyant also decrease as a function of initial inventory levels. Furthermore, as shown in Figure 3-6, the gaps between the bound and empirical performance also

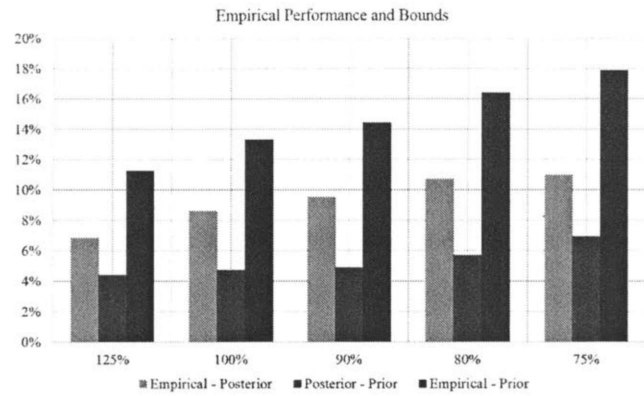

Figure 3-6: This shows the empirical gap between the analytical bounds of the EMPA and its practical performance.

grows with decreased inventory levels. These effects reinforce the insight that the more limited the inventory, the greater the performance lag in all of the online algorithms relative to the full-knowledge offline strategy. Through further simulation we found that this result was consistent across all scenarios in both case studies and all forms of penalty functions. However, the empirical performance of the algorithm on the actual data is significantly better than the worst-case analytical guarantees provided by the prior bound from Theorem 1. In the most inventory-constrained scenarios, the gap between the empirical ratio and the prior bound reaches up to 18%, while the actual performance of the algorithm is within at least 14% of the expected Clairvoyant revenue in practice in reasonably inventory-constrained scenarios, improving to within 9% on average in the least constrained cases.

3.5 Conclusions

We show that our methods are computationally efficient and that the resulting personalized offers produce an expected marginal gain of up to 12% in profits over existing pricing practices in retail, while performing achieving on average 98% of the expected profit of a full-knowledge Clairvoyant strategy. In the airline study we similarly observe revenue lifts on the order of 3-7% over existing industry baseline methods that do not use personalized pricing or recommendation methods. We conduct empirical comparisons of our proposed algorithms and establish that in constrained inventory cases, additive methods outperform multiplicative approaches by as much as 10% relative to a Clairvoyant strategy. However, in most scenarios, the relative improvement by using an additive approach is marginal compared to the significantly greater computational ease of using a multiplicative method. We also find that multiplicative methods place a greater emphasis on profit maximization by offering more expensive products on average, which are coupled with larger discounts. We develop innovative business insights within each case study that help online sellers determine target audiences for various recommendation schemes and general approaches to pricing and personalization strategies. Thus, our work provides contributions through novel analytical problem structure and performance guarantees, as well as through the development of tractable methods that are suitable for practical application and result in a business edge potentially worth millions in profits to an online seller.

Bibliography

[1] Rakesh Agrawal, Ramakrishnan Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499, 1994.

[2] Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. Real-time optimization of personalized assortments. *Management Science*, 60(6):1532–1551, 2014.

[3] Pavithra Harsha and Shiva Subramanian. Hierarchical demand modeling in the presence of partially or completely unobserved lost sales. In *INFORMS Revenue Management and Pricing Conference (New York, NY, United States, June 2016)*, 2016.

[4] Florian Verhein. Frequent pattern growth (fp-growth) algorithm. *The University of Sydne. p1-10*, 2008.

Chapter 4

Personalized Bundle Recommendations under Demand Uncertainty

4.1 Introduction

With the growing popularity of smart phones and increasing rates of digital connectivity through social networks across the world, the online market has undergone enormous and rapid growth. Based on a study by eMarketer, a market research company focused on digital trends, media and commerce, "Women and men ages 18 to 34 are more likely than 35 to 64 year olds to engage in nearly every online shopping activity, with 40% of males and 33% of females in the younger age group saying they would buy everything online if they could." This swiftly expanding behavior across more generations generates a very large body of available data for online sellers to leverage in strategic personalized product recommendation and pricing strategies. Consider the industry example of StitchFix, a personalized styling startup that only became cash-flow positive in 2014, then achieved $250M in revenue in 2015 that was doubled in 2016. In addition to smart business practices, strategic inventory management and sophisticated product recommendation models, the real innovation behind

StitchFix's success is its massive collection of consumer feedback data. On any given item sent to a consumer in a personalized styling box, they collect 100-150 data points ranging from product features (such as color, style, size) to consumer features (such as weight, height, location). Due to the nature of their interaction with their consumers, they are able to capture data about lost sales that is traditionally not available to brick-and-mortar retail channels. This business example clearly demonstrates that the true key to successful personalization and cross-selling in this expanding online channel is accurate knowledge of demand and consumer preferences.

As demonstrated by industry examples such as that of StitchFix, gaining a competitive edge through personalized assortment or pricing strategies in the online setting relies fundamentally on the ability to accurately estimate or rely on a given estimate of consumer demand and preferences. However, many of the following settings have become increasingly common across all industries in their business practices: (1) no (or few) product bundles have ever been offered historically, thus there is an absence of transaction data from which demand estimations can be constructed; (2) new products are introduced into the market with no previous purchase history known to the seller, hindering reasonable demand estimates; and finally, (3) first time shoppers frequently appear in the market without any known purchase history or preference information. We wish to address the problem of making personalized bundle offers under all of these challenging circumstances by incorporating robustness into our modeling framework for personalized bundle recommendations. More concretely, we are interested in the problem of offering personalized product bundles to consumers during their shopping session, which they can choose to accept, or to purchase individual items at their full prices, or to simply leave the online market without making any purchase at all. We want to maintain the trade-off between myopic profit maximization with long-term profitability under inventory constraints while also offering bundles that are robust in the realizations of the unknown consumer demand functions.

We address this problem of capturing uncertainty in consumer demand by considering two possible approaches: (i) dynamic learning, and, (ii) robust optimization.

Under setting (i) we consider the multi-armed bandit problem, which models the trade-off between exploration and exploitation in a dynamic setting in which the agent playing tries to learn information about the system by trying new "arms", while maximizing his payoff over the selling period by playing "arms" whose rewards he has already learned about, are high. In the context of our problem, this agent is the online seller who is interested in maximizing long-term expected revenue over the entire selling horizon by offering personalized bundles, represented by the arms, to arriving consumers whose demand function is unknown. Depending on their choice to purchase items or not when presented with the bundle offer, the seller observes a reward, generated from their decision, in the form of revenue. Our specific modeling framework is a contextual multi-armed bandit in which the contexts correspond to high-dimensional vectors of personalized consumer features, which are realized when the consumer arrives in the online market for a transaction. Due to the potentially complicated functional forms of the true demand, the objective function of our optimization problem for personalized bundle offers becomes highly non-linear. Furthermore, we still incorporate the profitability trade-off of our prior modeling approach in order to reduce inventory-related costs throughout the selling period. Thus, we develop a fairly challenging dynamic learning model that couples the joint problems of constrained assortment planning and dynamic pricing under the personalized setting of contextual bandits. We then demonstrate computationally that our modeling framework is stable and performs well against existing benchmarks from the literature across changes in demand function, number of available products and number of consumer features. However, we also find that in the cases where the true demand model is misspecificed, there is average estimation error of 1% to 2%, which may cause severe infeasibility or suboptimality of the recommended offers in the dynamic learning setting. Therefore we mitigate this in setting (ii), by considering the robust counterparts to our personalized bundle recommendation problem under various potential uncertainty sets. We investigate the full-knowledge Clairvoyant setting in which the entire consumer arrival sequence is known in advance and the problem is inventory constrained. Under this setting, we observe that the effects of minor perturbations in the demand on the feasibility of the optimal solution can be as large as 70% of inventory constraints violated in highly constrained inventory cases. Therefore, we conclude that robust optimization is critical to protecting against error in demand learning and preserving margin in the pricing problem. We also find that ellipsoidal uncertainty sets outperform polyhedral uncertainty sets in terms of the price of robustness and also find that robust optimization is crucial in highly inventory constrained problems.

We are interested in implementing this complex model practically in both of our desired uncertainty settings. In the case of the multi-armed bandit problem with dynamic learning, we consider constructing an algorithm based on the coupling of Taylor series approximations and upper confidence bound (indexing) policies. We do not assume that the expected reward function has any particular structural conditions other than differentiability, and consider its first-order Taylor series approximation. By making some assumptions on the bounded nature of the higher-order terms, we are able to establish analytical guarantees on the performance of this algorithm. More specifically, the performance of any algorithm in the context of multi-armed bandit problems is modeled by regret, which measures the difference in long-run expected reward between a given policy and an oracle strategy that knows the true demand distribution in advance of the selling period and thus always chooses the optimal arm (in this case bundle offer for each consumer). Our Taylor series-based algorithm allows us to establish near-optimal bounds on the asymptotic behavior of this regret, showing that it is on the order of $\tilde{O}(\sqrt{(T)})$ and is independent of the number of arms. Given our potentially combinatorially expansive state space of possible product bundles, this result is crucial for the tractability of our approximation algorithm. The robust optimization setting presents computational challenges depending on the uncertainty set choice. In the case of a polyhedral uncertainty set, we are able to formulate our robust counterpart as a linear optimization problem, which is efficiently solvable and only marginally less optimal, on the order of 3-5% in mildly conservative settings, in performance relative to the ellipsoidal uncertainty set that results in a non-linear problem. However, for smaller instances of demand group size, ellipsoidal uncertainty sets are also reasonably implementable and offer performance within 10% of the optimal profit, which proves to be infeasible in 50% of the instances when demand is perturbed by as little as 0.01%.

4.1.1 Contributions

As described above, our primary goal is the incorporation of demand uncertainty into our personalized bundle recommendation problem. In the setting of dynamic learning, our aim is to develop a generalization of current approximation approaches for solving the contextual multi-armed bandit problem in order to obtain near-optimal bounds on the asymptotic regret. In the robust optimization setting, we seek to establish the importance of capturing demand uncertainty and observe its impact in highly constrained inventory settings. We summarize our main contributions in this work as follows:

1. **Development of a high-dimensional learning framework based on contextual multi-armed bandits that incorporates personalized demand modeling while considering inventory levels and future demand.** We adapt our personalized bundle pricing and recommendation model from Chapter 2 to the online setting where demand cannot necessarily be estimated from historical transactions. We model this new problem as a contextual non-linear multi-armed bandit problem with the goal of learning personalized high-dimensional consumer demand functions while jointly solving the dynamic bundle assortment and pricing problem to maximize expected revenue over the course of the entire selling horizon. We also capture our previous trade-off between myopic profit maximization and long-run expected rewards by considering inventory management in the optimization problem for selecting and pricing the bundle offer.

2. **Construction of a generalized approximation algorithm for solving and implementing this model with minimal assumptions on the structure of the demand.** In order to implement and solve the above dynamic

learning model, we develop an approximation algorithm based on the first-order Taylor series expansion of the expected reward function in combination with the UCB approach. We achieve this by making no assumptions on the properties of the true demand function other than differentiability. This approximation approach allows us to iteratively estimate the coefficients for a linear demand function that approximates the true underlying demand distribution, which may be parametric and non-linear. Thus, we generalize the approaches of many existing works in the literature, which assume specific structural conditions on the true demand or the expected reward function. Given any demand function, regardless of its form, we are able to model the above problem and solve it using this approximation algorithm.

3. **Establishment of analytical guarantees on the asymptotic behavior of the regret relative to an oracle strategy.** We are able to establish regret bounds that are similar to existing works with linear expected reward functions for this Taylor series approximation algorithm. With some assumptions on the bounded nature of the higher-order terms in the Taylor series expansion of the reward and their role in developing adequate upper confidence bounds, we obtain near-optimal asymptotic regret on the order of $\tilde{O}(\sqrt{(T)})$.

4. **Analysis of empirical results based on simulations that show our algorithm outperforms baseline methods such as the myopic ϵ-greedy approach.** We test the efficacy of our approximation algorithm on a range of problem instances with varying numbers of consumer arrivals and numbers of products, different demand functional forms, and various lengths of cold starts relative to an ϵ-greedy myopic strategy. We find that across all cases our approach ultimately outperforms this baseline and converges more efficiently with errors decreasing in time.

5. **Analysis of the effect of demand perturbations on the optimality of the Clairvoyant solution.** We consider the inventory constrained Clairvoyant problem, which considers the full-knowledge setting where the entire consumer

arrival sequence is known in advance. This benchmark is considered the "optimal" solution against which our prior approximation algorithms were tested. We find that minor perturbations on the order of 0.10% to 0.50% in the values of the personalized consumer willingness-to-pay results in average infeasibility of approximately 50% of the optimal Clairvoyant solution. We also find that certain products with higher consumer popularity, defined through a generally higher propensity-to-buy across all consumer types, are more susceptible to this effect.

6. **Formulation and computational studies of robust counterparts to the personalized bundle recommendation problem under polyhedral and ellipsoidal uncertainty sets.** In the case of polyhedral uncertainty sets, the robust counterpart to the Clairvoyant problem is a linear optimization problem, which is highly tractable. We find that while the ellipsoidal uncertainty sets result in a non-linear problem, they outperform polyhedral uncertainty sets with respect to percentage of optimal objective obtained and are reasonably tractable for small demand groups $\hat{S}$ of up to 10 products. We demonstrate that we can achieve within 10%-20% of the optimal objective function value while maintaining feasibility with very high probability. Furthermore, we show that robust optimization is invaluable under highly inventory constrained settings.

4.1.2 Literature Review

The growing availability of personalized consumer data in the online sector has sparked a great deal of work on the problems of (i) dynamically learning consumer preferences through strategic pricing and recommendation systems, and, (ii) developing pricing and assortment optimization strategies through robust optimization when facing demand uncertainty. We summarize the primary bodies of literature related to our work within the field of research on the multi-armed bandit problem, as well as within robust optimization.

The multi-armed bandit problem models the trade-off between exploration and exploitation in a dynamic learning setting where the agent playing the "arms" in each round tries to acquire information about the system by playing new "arms", while maximizing his long-term payoff by playing "arms" whose rewards he knows to be high. The performance of any algorithm in the context of this problem is modeled by regret, which measures the difference in expected revenue between a given policy and an oracle strategy that knows the true demand distribution in advance of the selling period and thus always chooses the optimal arm. This model was pioneered by the work in [34], which develops a converging stochastic approximation method for estimating the true mean value of each arm. The seminal work by [25], since published as a textbook, established the Gittins index, which is an optimal policy for maximizing the long-term expected reward over the entire horizon of trials in the setting with independent of arms, infinite horizon, one arm pulled in each round, and a discount factor that is stricly less than 1; extensions to this work analyze this indexing policy under variations of these setting parameters. In [29], they develop asymptotically efficient policies for this problem based on upper confidence bounds, which function as an indexing policy, for each arm. The survey work in [40] provides a general overview and evaluation of these various multi-armed bandit algorithms. Later works, such as [30], explore the role of dependencies between arms in the context of this problem and provide optimal policies under different scenarios of discounted and non-discounted rewards, showing that their methods outperform those made for bandit problems with independent arms.

Many algorithms have been developed for solving this general multi-armed bandit problem. One initial class of methods is based on the concept of upper confidence bounds, which provide optimal strategies for maximizing long-term expected rewards by essentially behaving as an indexing policy for each arm that is based in Bayesian updates of the posterior distribution over the expected reward of each arm. The strong analytical performance of this Bayesian-UCB approach is established in [27]. However, there is also a growing body of work on the randomized Bayesian algorithm known as the Thompson sampling algorithm (or the probability matching algorithm). By contrast, this approach randomly samples the posterior distributions across all of the arms and chooses the one with the highest sampled reward. In works such as [7] and [20], they evaluate the performance of this approach relative to the UCB algorithm and show similar theoretical guarantees as well as better empirical performance, respectively. More recently in [37], they convert UCB regret bounds into Bayesian regret for posterior sampling and show that their bounds from Thompson sampling are more generalized and in some cases stronger.

Having developed the necessary framework for modeling and solving this problem through tractable algorithms to varying degrees of optimality, there are generally two classes of problems to which this model is applied: (i) dynamic assortment planning, and, (ii) dynamic pricing. We have summarized the static approaches to these problems, in which the demand in considered known, in our modeling and literature review from Chapter 2. In the context of multi-armed bandits, both the problems of assortment planning and pricing becoming dynamic due to the need for demand learning as the problem horizon progresses. In the case of dynamic assortment planning, the existing literature analyzes both constrained and unconstrained cases with respect to inventory. In [3], the authors consider unknown consumer demand defined by a multinominal logit (MNL) choice model and provides an approach that simultaneously explores and exploits (as opposed to explore-then-exploit methods) that achieves a near-optimal worst-case regret bound of $\tilde{O}(\sqrt{(NT)})$ that is independent of instance parameters; they further extend this to also consider cardinality constraints on the size of the assortment. The work in [17] uses Lagrangian relaxation of weakly coupled dynamic programs to obtain closed-form indexing policies with corresponding near-optimal bounds and show that their approach outperforms greedy methods that chooses arms myopically to optimize the single-period reward in each round. Building on the prior approach, the authors of [38] propose a stochastic dynamic programming model that incorporates learning with Bayesian updates and demonstrates that it is profitable for an agent (online seller) to use discounting strategies early on in the horizon in order to accelerate demand learning.

In addition to dynamic assortment optimization, there is a great deal of ongoing work that applies the multi-armed bandit framework to dynamic pricing (in some cases in tandem with assortment planning and inventory constraints). In [9] they present a framework based on partially observed Markov decision processes for dynamically pricing fashion-like products. In the joint works of [18], and [19], they analyze "one-step ahead" pricing based on a Taylor expansion of the expected reward function for the next period; while they do not provide any analytical guarantees, their numerical results on various classes of demand functions indicate that their policies perform well. The work in [14] analyzes the problem of jointly learning the demand and dynamically pricing products through dynamic programming methods under unknown linear demand and analyze the performance of their approaches in various settings with competition. In [28], they also consider dynamic pricing under a setting in which the unknown demand function is linear and achieve lower bounds on the regret on the order of $\tilde{O}(\sqrt{(T)})$ and $\tilde{O}(log(T))$ for scenarios where the seller has no information or limited information about demand under incumbent prices, respectively. As opposed to using dynamic programming, the authors in [16] consider an adaptive optimization approach based on data-driven uncertainty sets that are dynamically updated and align the reward optimization with the demand estimation. Finally, other recent works such as [12], [23], and [37], consider machine learning methods as opposed to optimization-based strategies for dynamic pricing bandit problems. In [12], a LASSO-based approach is presented with near-optimal performance, whereas in [23], they employ an algorithm based on Thompson sampling in an inventory-constrained setting. While Thompson sampling is easy to implement and does not require the use of confidence bounds, it requires the specification of an underlying probabilistic model, which is an assumption we do not make for our more generalized learning method.

While many of the multi-armed bandit problems are applicable to a wide range of demand functions and settings, two particular directions of interest to us in this work are that of contextual bandits and multi-armed bandits with linear expected reward functions. In contextual bandit problems, each consumer is summarized by some unique vector of contextual features which are realized to the online seller upon arrival. The seller must strategically make pricing or joint assortment decisions in this setting with unknown demand, which is described by a high-dimensional feature vector. In the seminal work of [8] on contextual linear bandits, they present algorithms for two settings in the best of which they achieve a regret on the order of $\tilde{O}(\sqrt{(T)})$, where the regret is dependent on the number of arms. By contrast in [1], they also consider a linear reward function structure, but improve the regret bound by a logarithmic factor under the assumption that the observed noise is conditionally R-sub-Gaussian, which allows them to construct smaller confidence intervals when using the UCB approach. In the more recent work in [4], they consider a linear reward function under knapsack constraints and establish a regret bound that is independent of the number of arms. There is also a body of research under multi-armed bandits that relaxes the linearity assumption on the structure of the expected reward function. The works in [10], and [11], in this nonlinear setting with knapsack constraints where in the first work there are no contexts and the regret is characterized by the number of arms, and in the second work arbitrary contexts are considered and the regret is independent of the number of arms. In [5] and [6], they analyze non-linear reward functions under convexity and knapsack constraints, respectively. In particular, the authors of [5] establish regret that is independent of the number of arms but contexts are known in advance of the learning horizon; whereas in [6], they provide an extension to a concave reward function and constraints, in which they assume that rewards and consumption are drawn IID from a joint distribution. Furthermore, the work in [33] also considers this non-linear case in which unlike all the previous works, each round consists of a subset of arm pulls as opposed to only a single one; their regret is characterized by a dependence on the number of arms.

The literature on robust optimization has been rapidly expanding over the past decade, during which a great deal of fascinating research has been conducted across a great deal of applications and theoretical fields. The seminal work in [13] characterized the construction of robust counterparts to linear optimization problems where the uncertainty set is defined through a general norm. This is a building block for the work done in the robust section of this chapter as we consider the first and second norms, corresponding to polyhedral and ellipsoidal uncertainty sets, respectively. In the following work in [15], the authors propose relaxed robust counterparts for general conic optimization problems that preserve the tractability and underlying structure of the nominal problems. Instead of the previously established result that linear programming problems had robust counterparts that were second order conic problems, and second order conic problems resulted in semidefinite programming problems, this work proposed a method for preserving problem structure in robust counterparts (i.e. linear programming robust counterparts are also linear programming problems). This work also directly influences the computational results obtained in this chapter.

The work across the field of robust optimization that is relevant to the model proposed in this thesis is related to dynamic pricing and assortment optimization. With regards to pricing, the work in [39] considers a robust formulation for a single-product pricing problem with capacity constraints when demand is uncertain but assumed to be a linear function of price. In an extension to this topic, [32] consider the dynamic pricing of a single product in a setting with firm competition in which each seller faces their own demand uncertainty. The authors consider another direction in the dynamic pricing and inventory control problem in [2], in which there are multiple products, and demonstrate how to use the deterministic solution of the original problem in the robust approach. In [22], the authors consider dynamic congestion pricing under demand uncertainty in which the flows correspond to user equilibrium on a network of interest and show that robust dynamic solutions outperform static ones. While they do not consider pricing under uncertainty in [21], they instead analyze demand response management with price interval uncertainty and demonstrate how to formulate these problems using mixed integer linear programming for this stochastic problem.

In addition to robust optimization applications to pricing problems, there is also a growing body of work related to robust applications in assortment optimization and revenue management problems. The work in [36] and [35], study the multi-period assortment planning problem that is dynamically solved for each consumer arrival under the assumption that the consumer choice model is a multinomial logit whose parameters are unknown and represented through an uncertainty set. In the revenue management literature, the authors of [31] develop a robust formulation for the capacity allocation problem using polyhedral uncertainty sets and demonstrate empirically that this approach outperforms well-known heuristics in the literature while maintaining scalability. A thorough general summary of the work done in the field of robust optimization is provided in [24], which covers many different fields of applications as well as settings of theoretical development such as across static and dynamic problems. To the best of our knowledge, none of the existing literature has considered applying robust optimization methods to the joint assortment and pricing problem addressed in a personalized bundle recommendation model.

Relating more closely to the above literature, in the first setting in this work we are primarily interested in the intersection of this last body of works between contextual bandit problems with expected linear reward functions. A summary of the contributions of the above works relative to our approach is provided in Table C.1 in Appendix C. We develop a non-linear contextual bandit framework that we reduce to a linear setting through Taylor series expansion. However, there is inherent error in dynamic demand learning, which may affect the optimality of pricing and assortment decisions. Therefore, in the second setting, we are interested in applying the methodologies from the robust optimization literature to the personalized bundle recommendation problem and analyzing the effects of such demand uncertainty on the feasibility of optimal offers in the full-knowledge Clairvoyant problem, which serves as the benchmark method for the approximation approaches presented in Chapter 2.

4.2 Dynamic Learning Approach

In order to address potential uncertainty in consumer demand, we first consider the approach of dynamic learning by developing a modeling framework for the personalized bundle recommendation problem based on multi-armed bandit theory. In particular, we focus on contextual multi-armed bandits and how to incorporate bundling into this problem setting. We then establish an approximation algorithm to make this approach tractable for the online setting, and establish analytical guarantees for the asymptotic behavior of the regret, which we show to be on the order of $O(\sqrt{T})$. Finally, we conclude Section 4.2 with computational analysis showing the performance of our proposed algorithm relative to existing benchmarks from the literature, and show that it is relatively robust across changes in numbers of consumer features and products, as well as across various lengths of cold starts and demand functional forms.

4.2.1 Problem Setting and Model Formulation

We consider the problem of offering personalized bundle recommendations for individual consumers as they arrival sequentially in an online setting. We have a monopolistic seller who offers products $j = 1, ..., n$ from a given demand group $\hat{S}$ of related items. The resulting set of all possible bundle composition and price combinations is indexed by $i = 1, ..., N$. In order to incorporate dynamic demand learning, we now model this problem using contextual multi-armed bandits. Under this setting, we consider a finite horizon of length T in which any given period $t = 1, ..., T$ has a single consumer arrival. We define $x_i(t)$ as the feature vector associated with arm i at time t, which corresponds to a given bundle at a given price. Note that this feature vector definition is easily extended to also incorporate personalized consumer context features. We then denote $i(t)$ as the arm (bundle offer) chosen by the online seller at time t. Symmetrically, we let $i^*(t)$ be the optimal arm for the consumer arriving at time t, which is known to the oracle strategy that has full knowledge of the true demand function parameters. Let $y_{i(t)}(t)$ be the reward received by the seller at time t as a result of choosing to offer bundle $i(t)$ to the consumer. Similarly we denote $y_{i^*(t)}(t)$ as the optimal reward received by the oracle strategy at time t as a result of offering bundle $i^*(t)$. Finally, let us define $f(\theta, x)$ as the unknown non-linear expected reward function that the seller is attempting to accurately learn through dynamic personalized bundle offers.

In order to evaluate the performance of our proposed algorithm, we define the following fundamental performance metric of regret for all multi-armed bandit problems. The regret is defined as the total difference in reward between a given policy π and the oracle strategy, which always chooses optimal arm $i^*(t)$ and receives reward

$y_{i^*(t)}(t)$. This total regret $Regret(\pi, T)$, which is a function of π and the length of the horizon T, is explicitly defined as,

$$Regret(\pi, T) = \sum_{t=1}^{t=T} y_{i^*(t)} - y_{i^\pi(t)}(t),$$

where the optimal choice of arm in period t is explicitly defined as the reward-maximizing option for each consumer, given by,

$$i^*(t) = \arg\max_i y_i(t).$$

We also define the first-order Taylor series transformation of the expected reward, which is a function of the historical matrix of observed feature vectors $x(t)$ up until time t and current expected reward function estimates θ, as $Z_i(t)$ below,

$$Z_i(t) = T_i(x(t), \theta) = f(x_i(t), 0) + (\theta)^T f'(X_i(t), 0)$$

Given these definitions, the general methodology behind the proposed dynamic learning approach based on this Taylor series approximation is as follows:

1. Upon their arrival to the online market, the feature variables of a given consumer in period t are realized to the seller and therefore our algorithm.

2. Using the most recent estimates of the transformed expected reward function parameters, we write the expected reward function as a sum of independent random variables (in our case these random variables are the first-order Taylor series transformation of incoming consumer context features).

3. By the Azuma-Hoeffding inequality we can then find an upper confidence bound over the actual reward of each potential bundle offer (arm), and thus obtain an upper confidence bound over the estimated reward of each arm from this Taylor series approximation.

4. Considering the upper confidence bounds on all of the available offers (arms),

our algorithm chooses the one with the highest UCB value and offers this to the consumer at time t.

5. After observing the purchase decision of consumer t, we regress their new feature vector and the corresponding observed demand against the matrix of historical features and respective demands in order to update our estimates of the transformed parameters of the expected reward function. Note that this step is equivalent to fitting a linear regression whose dependent variable is the expected reward and the independent variables are the transformed incoming random context variables (which we have transformed to the first-order derivative space).

6. Using these updated reward function estimates, we repeat this procedure for the next arriving consumer in period $t+1$.

This method requires the coordination of two separate algorithms, one of which is nested within the other; we must use and update the upper confidence bounds (UCBs) within the larger modeling framework in order to make dynamic personalized bundle offers, while also updating the global estimates of the demand. We now explicitly detail this below in the following Sections 4.2.3 and 4.2.2.

4.2.2 Lower-Level Algorithm for Upper Confidence Bounds

In order to construct the upper confidence bounds (UCBs) to select the best bundle offer for each arriving consumer and thus implement the larger global algorithm for sequential learning, we consider an extension of the UCB approach in [8] in which we replace the historical data matrix of past feature vectors and observed rewards with the first order transformation matrix based on $Z_i(t)$ of the observed features and expected reward function. We define the algorithm in detail below.

Note that in order to simplify implementation, you can also replace the $a_i(t)$ coefficient estimates in Step (6) above directly with the regression coefficients and observe the resulting performance, as we do in Section 4.2.5. In practice, the assumption of independence between the values of $a_i(t)$ and $x(t)$ is not necessarily crucial and circumventing this tedious construction process could greatly ease practical application. The intuitive reasoning for using the transformed matrix $Z(m)$ in Step (1) is due to our necessity to obtain confidence bounds; we need to upper bound the quantity $||a_i(t)||^2$, which can be done only if $Z(t) \cdot Z(t)$ is sufficiently regular in the sense that all eigenvalues are sufficiently large. If some of the eigenvalues are small, we have to deal with them separately. Therefore, we instead use the transformation below. Given a new consumer arrival, this algorithm gives us a method for estimating the upper confidence bounds across all arms (potential bundle offers) $i = 1, ..., K$ in each period t. By selecting and making the offer of the bundle with the highest upper confidence bound, we develop a recursive method for learning the consumer demand dynamically by using these UCBs in Step (2a) of Algorithm 4.2.2. Thus, we have developed a tractable approach for dynamic learning based on Taylor series approximation and UCBs with virtually no assumptions on the functional form of the demand. We now consider the analytical implications of using this methodology in Section 4.2.4 below.

Algorithm 4.2.1: TUCB($\delta \in [0,1]$, number of Trials T) for determining upper confidence bounds (UCBs) on each potential bundle offer (arm) $i = 1, ..., K$ in order to select the best option for the consumer arriving at time t.

Input : History of previous transformed features $Z_i(1...t-1)$, current incoming feature $X_i(t)$ and its transformation $Z_i(t)$.

Output: Upper confidence bounds (UCBs) of all potential bundle offers (arms) $i = 1, ..., K$, resulting in the algorithm's decision to offer arm $i(t)$ that maximizes the upper confidence bound $ucb_i(t)$ to the consumer at time t.

Let $Z(m)$ be the matrix of all previous transformed feature vectors up until time t;

Let $Y(m)$ be the vector of all previous observed rewards up until time t;

Calculate the eigenvalue decomposition:
$Z(t) \cdot Z(t)' = U(t)'\Delta(\lambda_1(t), ..., \lambda_d(t))U(t)$, where $\lambda_1(t), ...\lambda_k(t) \geq 1$ and the rest are less than 1. Also $U(t) \cdot U(t)' = \Delta(1, ..1)$;

Now for each feature vector $z_i(t)$, let $\tilde{z}_i(t) = U(t) \cdot z_i(t)$, and,
$\tilde{u}_i(t) = (z_{i,1}(t), z_{i,2}(t), ..z_{i,k}(t), 0, 0, 0...)'$, and,
$\tilde{v}_i(t) = (0, 0, 0, z_{i,k+1}(t), z_{i,k+1}(t), ..z_{i,d}(t))'$;

Calculate, $a_i(t) = \tilde{u}_i(t)'.\Delta(\frac{1}{\lambda_1(t)}, \frac{1}{\lambda_2(t)} .. \frac{1}{\lambda_k(t), 0, ..., 0}) \cdot U(t) \cdot Z(t)$;

Next, calculate the upper confidence bounds for all arms $i = 1, ..., K$ using,
$width_i(t) = ||a_i(t)||(\sqrt{(}ln(2TK/\delta) + ||\tilde{v}_i(t)||$, and,
$ucb_i(t) = z(t) \cdot a_i(t)' + width_i(t)$;

Choose that alternative $i(t)$ which maximizes the upper confidence bound $ucb_i(t)$.

4.2.3 Global Algorithm for Personalized Bundle Offers

Having established the necessary algorithm for acquiring the UCBs, we now develop a π policy based on this Taylor series approximation approach for selecting bundle offers (arms) from $i = 1, ..., K$ for every given arriving consumer t. We develop a global algorithm based on Taylor series expansion for solving the problem of personalized assortment planning with online demand learning, which is defined as follows:

Algorithm 4.2.2: Global($\delta \in [0, 1]$, number of Trials T) for utilizing upper confidence bounds (UCBs) on each potential bundle offer (arm) $i = 1, ..., K$ in order to select the best option for the consumer arriving at time t, and periodically updating the demand estimates in order to achieve asymptotic regret on the order of $\tilde{O}(\sqrt{T})$.

Input : The features $x_i(t)$ of a given consumer in period τ.

Output: Estimates $a_i(\tau)$ of transformed $f(\theta^*, x_i(\tau))$ of the expected reward function through the first-order Taylor series expansion.

1 Using the most recent estimates $a_i(\tau)$ of the transformed expected reward function parameters, write the expected reward X_τ as a sum of independent random variables, $X_\tau = Y_\tau \cdot a_i(\tau)$ (these random variables Y_τ are the first-order Taylor series transformation of context features).;

2 By the Azuma-Hoeffding inequality (Lemma 1) calculate an upper confidence bound $ucb_i(\tau)$ over the actual reward of each arm $i = 1, ..., K$, and thus obtain a UCB over the estimated reward of each arm from the Taylor series approximation using Algorithm 4.2.1 defined in Section 4.2.4 below;

3 Choose the arm i that minimizes $ucb_i(\tau)$ over all available offers $i = 1, ..., K$ and offer this bundle at time τ;

4 Having observed the decision of τ, regress the new feature vector $x_i(\tau)$ and corresponding observed demand against the matrix of historical features $x(\tau)$ and respective demand estimates θ. Update the estimates of the transformed parameters of the expected reward function to $a_i(\tau + 1)$. Note that this step is equivalent to fitting a linear regression whose dependent variable is the expected reward and the independent variables are the transformed incoming random context variables (which we have transformed to the first-order derivative space);

5 Return to Start for consumer $\tau + 1$.

We establish this framework based on the contextual linear approach from [8], which we extend by replacing the historical data matrix of past feature vectors and observed rewards with the first order Taylor series expansion matrix based on $Z_i(t)$ of the observed features and expected reward function. The above approach uses the indexing policy based on the UCBs developed in Algorithm 4.2.1 in order to make consecutive dynamic offering decisions to consumers. As this approach requires only the Taylor series expansion of the expected reward function, it makes no assumptions on the functional form of the true consumer demand model outside of differentiability. The above method is therefore very generalized compared to the majority of settings in the current literature, and as shown below in Section 4.2.4 achieves asymptotic regret on the order of $\tilde{O}(\sqrt{T})$.

4.2.4 Analytical Guarantees on Asymptotic Regret

In order to establish an analytical result regarding the asymptotic behavior of the regret of our global learning Algorithm 4.2.2 from Section 4.2.3, we must first establish guarantees with respect to the upper confidence bounds used in the lower-level Algorithm 4.2.1 for selecting the optimal arm with the highest index, as described in Section 4.2.2. Let us begin by recalling the following well-known result from probability theory regarding the values of martingales with bounded differences:

Lemma 1. ***Azuma-Hoeffding Inequality*** *Let $X_1, ...X_m$ be random variables with $|X_{tau} \leq a_{tau}|$ for some $a_1....a_m \geq 0$. Then,*

$$P\left\{ \sum X_\tau - \sum E[X_\tau | X_1, ...X_{\tau-1}] \geq B \right\} \leq 2exp\left\{ -\frac{B^2}{2\sum a_\tau^2} \right\} \tag{4.1}$$

We can utilize the above result from Lemma 1 in order to show the following inequality in Lemma 2 utilized in Step (6) of the lower-level Algorithm 4.2.1 in order to determine the values of the UCBs for all arms $i = 1, ..., K$. This result is necessary in order to ensure that Algorithm 4.2.2 in turn will have asymptotic regret on the order of $\tilde{O}(\sqrt{T}))$.

Lemma 2. *Let $\phi(t)$ be constructed in such a way that for fixed transformed feature vectors $A(\tau)$, $\tau \in \phi(t)$, the rewards $y(\tau)$, $\tau \in \phi(t)$, are independent random variables* *with mean reward* $E[y_{i(\tau)}(\tau)] = f(\theta^*, x_{i(\tau)}(\tau))$. *Then with probability 1-*δ/T *we have that for all* $i \in 1, ..., K$,

$$|Y(t) \cdot a_i(t)' - f(\theta^*, x_i(t))| \leq ||a_i(t)||(\sqrt{2ln(2TK/\delta)}) + ||\bar{v}_i(t)|| \qquad (4.2)$$

We detail the proof of Lemma 2 in the joint work in [26]. Establishing the above result in Lemma 2 relies primarily on applying the Taylor series expansion to the consumer feature vectors and making an assumption regarding the behavior of its higher order terms, coupled with the correct application of the Azuma-Hoeffding Inequality from Lemma 1 above. If we assume that the Hessian matrix in the region of evaluation is bounded, we are able to prove the desired inequality. Thus, having shown the result in Lemma 2, we now want to establish the larger result on the regret of Algorithm 4.2.2 and show that asymptotically it behaves on the order of $\tilde{O}(\sqrt{T})$. We consider Algorithm 4.2.2 in which we replace the historical input matrices with their Taylor series transformations as defined in Algorithm 4.2.1. We first prove the following lemma, which bounds the number of entries in each set of independent trials ϕ.

Lemma 3. *The number of trials for which an alternative offer is chosen at stage* s *during which a dynamic choice is made is bounded by the following quantity. For all stages* s,

$$|\psi^s(T+1)| \leq 5 \cdot 2^s(C + K + 2ln(2TK/\delta))\sqrt{d|\psi^s(T+1)|}. \qquad (4.3)$$

The proof of the above Lemma 3, also detailed in [26], relies on the definition of the widths of the UCBs utilized in Step (2) of Algorithm 4.2.2, which are defined in Step (6) of Algorithm 4.2.1. Lemma 3, in combination with Lemma 2, allows us to establish the following result regarding the asymptotic behavior of the regret.

Theorem 2. *When the Global Algorithm, as defined in Section 4.2.3, is run with parameter $\delta/(1 + ln(T))$ then with probability $1 - \delta$ the regret of the algorithm is bounded by:*

$$B(T) \leq 44 \cdot (1 + ln(2KTlnT))^{3/2} \cdot \sqrt{Td} + 2\sqrt{T}$$

The proof of the above result is also specified in [26] and demonstrates that the behavior of the regret of the global algorithm for our dynamic learning approach is on the order of $\tilde{O}(\sqrt{T})$. Thus, this result ultimately demonstrates that we are able to achieve the best possible asymptotic theoretical behavior of regret, on the order of that achieved by linear models, within a non-linear generalized dynamic learning framework. While this is a promising analytical result, we next test and analyze the practical performance of our approach in the following Section 4.2.5 in order to see its behavior across various settings.

4.2.5 Computational Results

Having established the desired analytical result regarding the asymptotic regret of the global Algorithm 4.2.2 for online learning in our bundle recommendation problem setting, we next want to practically assess the empirical performance of our algorithm. We do so by conducting studies on various problem instances using synthetic data and simulate the learning process averaged over thousands of iterations of sample arrival sequences. We partition this section based on various scenarios with respect to changes in functional forms of demand, number of products offered, number of consumer features, length of cold starts and settings with misspecification of demand. We compare our approach to existing benchmarks from the literature and demonstrate our algorithm's general improvement in performance over these approaches in a dynamic setting. Ultimately, our results indicate that our approach is relatively robust to changes across all of these factors and that this approximation method consistently outperforms the commonly studied ϵ-greedy policies.

Simulation Design

We consider simulated data in order to analyze the performance of our proposed algorithm compared to benchmarks from the literature, in the form of various greedy approaches. In each of the following scenarios, we simulate the true demand parameters from a uniform distribution and iterate over a given setting 1,000 times in order to obtain average values for cumulative regret or estimation error. The features of each consumer are also drawn randomly from a uniform distribution upon arrival. In each particular demand parameter setting, we cycle through the arrival sequence until the regret or error no longer significantly changes. We then repeat the parameter selection step 100 times, for each of which we run these 1,000 iterations over an arrival sequence. Thus, all of the displayed results in the following sections are averaged across all of these higher level true demand parameter selections, as well as over many realizations of the arrival sequences within each of these settings. We also introduce two verions of an ϵ-greedy policy for benchmarking, in which: (i) we consider a fixed policy in which ϵ remains the same throughout the entire arrival sequence, and, (ii) we vary the behavior of the ϵ-greedy policy from having a fixed value of ϵ, to time-dependent decreasing trajectories for values of ϵ over the progression of the selling horizon.

Variations in Demand Functional Form

In the first set of experiments our goal is to evaluate the performance of the Taylor series approximation algorithm relative to these ϵ-greedy policies over various functional forms of true demand such as: non-linear logarithmic ($log(\theta^2 x + C)$) and non-linear exponential ($e^{\theta^2 x}$). In Figures 4-1 and 4-2 below, we present the results in which we consider logarithmic non-linear demand functions under varying numbers of products (n=2 or n=3).

The results above demonstrate that our method performs noticeably better in long-term convergence than either the constant or time-dependent decreasing versions of the ϵ-greedy policy. Intuitively, the transition from Figure 4-1 to Figure 4-2 shows

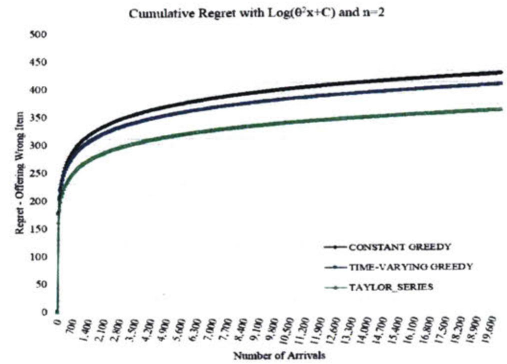

Figure 4-1: This plot shows the cumulative regret for a non-linear logarithmic demand function for $n = 2$ items under various ϵ greedy policies.

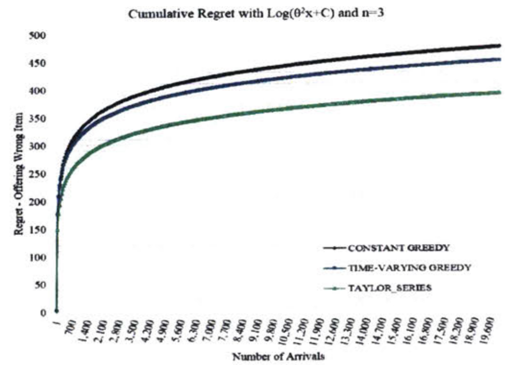

Figure 4-2: This plot shows the cumulative regret for a non-linear logarithmic demand function for $n = 3$ items under various ϵ greedy policies.

that the cumulative regret objectively grows as we study the case with more potential products. We present a more in-depth analysis of the effect on cumulative regret from increasing the number of products N in the next section below. We also observe this effect when we consider the non-linear exponential demand function, given by $e^{\theta^2 x}$, and shown in Figures 4-3 and 4-4 below.

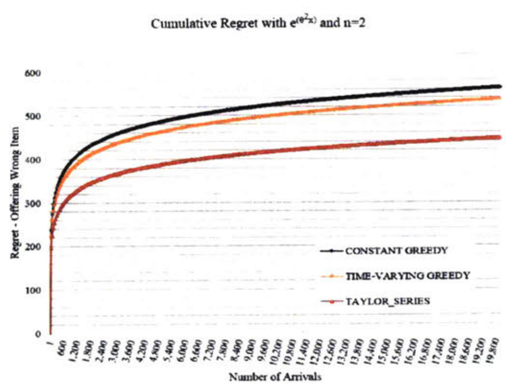

Figure 4-3: This plot shows the cumulative regret for a non-linear logarithmic demand function for $n = 2$ items under various ϵ greedy policies.

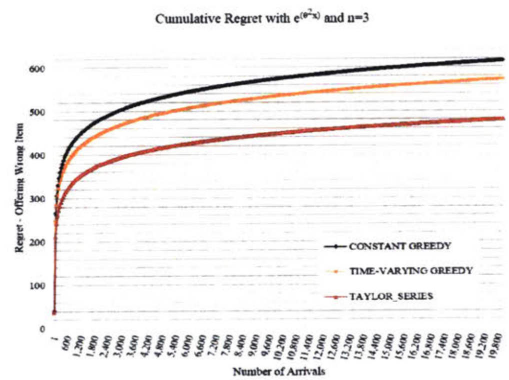

Figure 4-4: This plot shows the cumulative regret for a non-linear logarithmic demand function for $n = 3$ items under various ϵ greedy policies.

We note that across both demand functional forms, the gaps between the performance of the methods (in terms of cumulative regret) grow with an increase in number of products; this effect is also present in the total cumulative regret in each method with larger values of n. We observe that in the setting with a greater number of products, the greedy policies on average can stop exploring and in some cases never discover the optimal policy based on the correctly estimated reward parameters, resulting in an empirically much higher overall regret. Furthermore, we consider a fixed n and analyze the behavior of the regret in the various methods in order to quantify this exploration effect. The corresponding box plots of the distribution of cumulative regret for these simulated results are shown below in Figures 4-5 and 4-6 for the logarithmic demand function, and in Figures 4-7 and 4-8. Both sets of results demonstrate that across demand functions, the error in our method's estimation approach decreases as the selling horizon progresses, in contrast to the greedy policies which continue to accumulate regret through randomized trials throughout the length of the selling period.

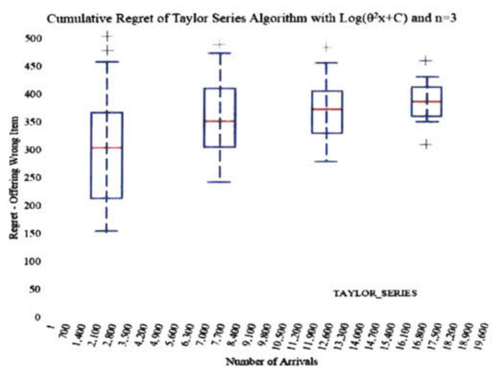

Figure 4-5: This plot shows the distribution of the cumulative regret of the our Taylor series algorithm at different points in the arrival sequence over many simulation iterations under logarithmic demand ($log(\theta^2 x + C)$) with $n = 3$ products being offered.

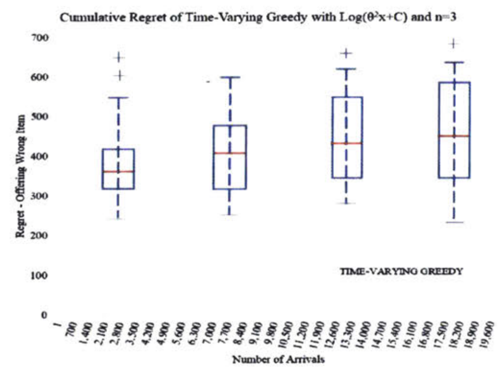

Figure 4-6: This plot shows the distribution of the cumulative regret of the time-varying ϵ-greedy policy at different points in the arrival sequence over many simulation iterations under logarithmic demand ($log(\theta^2 x + C)$) with $n = 3$ products being offered.

Performance Across Various Numbers of Products and Consumer Features

We also want to consider various number of available products n from which we develop bundles $i = 1, ..., K$, as well as more complicated demand functions in which there is a varying number of consumer features d. The effect on the cumulative regret

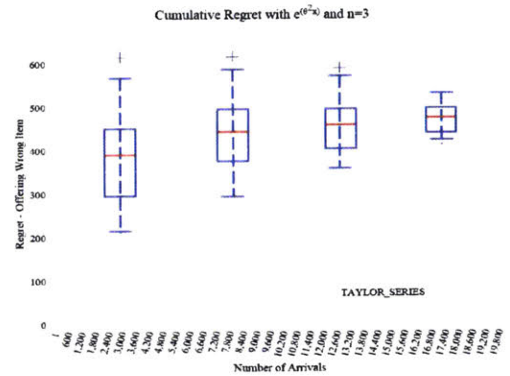

Figure 4-7: This plot shows the distribution of the cumulative regret of the our Taylor series algorithm at different points in the arrival sequence over many simulation iterations under exponential demand ($e^{\theta^2 x}$) with $n = 3$ products being offered.

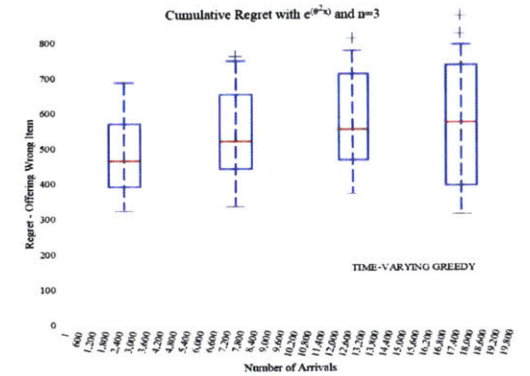

Figure 4-8: This plot shows the distribution of the cumulative regret of the time-varying ϵ-greedy policy at different points in the arrival sequence over many simulation iterations under exponential demand ($e^{\theta^2 x}$) with $n = 3$ products being offered.

of our proposed Taylor series approximation algorithm, as a function of these changes, is demonstrated below in Figures 4-9 and 4-10; we obtained symmetric results under exponential demand as well.

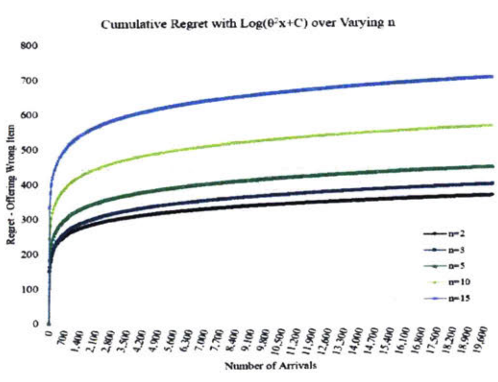

Figure 4-9: This plot shows the cumulative regret under our algorithm with logarithmic demand ($log(\theta^2 x + C)$) for a varying number of n items under various policies.

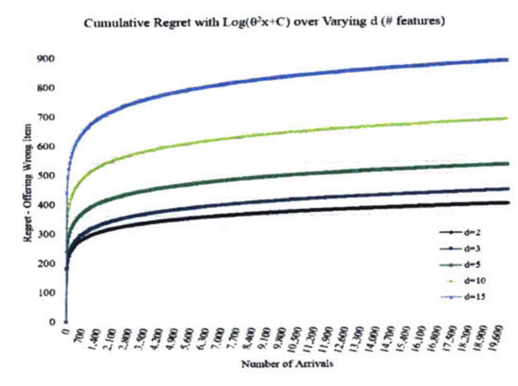

Figure 4-10: This plot shows the cumulative regret under our algorithm with logarithmic demand ($log(\theta^2 x + C)$) for a varying number of d consumer features.

We primarily observe two effects: (i) it becomes more difficult for the algorithm to identify the "optimal" bundle offer as the number of consumer features d or total number of products n grows; and, (ii) this effect is more greatly pronounced in the case of more consumer features d, as this indicates a more challenging learning problem for the algorithm due to the larger number of demand covariates. Notice that the scale on the two plots is not identical, and furthermore, the case with the largest number of consumer features $d = 15$ converges more slowly than in the case with the most products $n = 15$. The effect is due to the fact that it is easier for the model to learn the best few products for a fixed demand function with few covariates as opposed to a function with many covariates, each of which has a great deal of estimation error early in the learning process. We generalize the above results and summarize the cross effects in Table 4.1 below.

Arrivals for Convergence of Taylor Series Algorithm

Features (d) x Products (n)	$n = 1$	$n = 2$	$n = 3$	$n = 5$	$n = 10$	$n = 15$
$d = 1$	3,289	5,010	8,717	11,520	15,560	19,590
$d = 2$	6,086	8,950	10,890	13,770	18,640	22,360
$d = 3$	7,478	9,970	12,310	16,180	20,750	26,940
$d = 5$	13,198	16,920	17,990	20,330	25,420	29,170
$d = 10$	16,350	22,130	28,940	35,670	40,810	56,680
$d = 15$	29,640	31,870	36,640	47,660	55,210	71,320

Table 4.1: This table summarizes the time to convergence (in length of arrival sequence) of the Taylor series approximation algorithm, when we jointly vary the number of products offered n and the number of consumer features d.

Notice that the initial conclusions from the example in Figures 4-9 and 4-10 hold across the variations in Table 4.1. The time to convergence grows signicantly faster as we increase the number of features d when compared to the effect of increasing the number of products. Furthermore, we can also analyze this effect by observing the regret as a percentage of the number of arrivals needed for the algorithm's convergence, as summarized in Table 4.2. Here we observe that the growth is more marginal across both numbers of features and numbers of products in the ratio of cumulative regret to time to convergence. The effect of increasing d versus n is less noticeable, and we observe that in the most complicated joint learning scenarios, the ratio of regret to length of convergence remains at a reasonable proportion. This demonstrates the robustness of our approach to increases in the size and difficulty of the learning problem.

Regret to Convergence Arrival Ratio for Taylor Series Algorithm

Features (d) x Products (n)	$n=1$	$n=2$	$n=3$	$n=5$	$n=10$	$n=15$
$d=1$	9.2%	9.9%	13.7%	18.2%	23.2%	28.7%
$d=2$	11.1%	12.4%	14.7%	20.2%	25.3%	29.2%
$d=3$	14.0%	15.5%	16.3%	22.1%	28.7%	30.7%
$d=5$	18.7%	20.5%	22.8%	24.1%	30.1%	33.2%
$d=10$	23.5%	24.4%	26.9%	28.0%	33.2%	35.6%
$d=15$	29.2%	30.2%	31.3%	34.4%	35.7%	36.8%

Table 4.2: This table summarizes the ratio of the cumulative regret to the length of the arrival sequence needed for the convergence of the Taylor series approximation algorithm, when we jointly vary the number of products offered n and the number of consumer features d.

Performance under Misspecification of Demand

In addition to the effects of variations in demand functional forms and problem complexity (through number of available products and number of consumer demand function covariates), we are also interested in observing the effect of misspecification of demand on the performance of our proposed approach. More specifically, we test cases in which the true demand function is of some parametric form which we assume incorrectly, and observe the resulting error in the ultimate demand estimation of our methodology. This is a particular important setting due to the fact that true demand is often unknown and in many bundling-related scenarios there may not be enough adequate historical transaction information for initial assumptions regarding the form of the consumer demand. Proceeding along the same lines as in prior experiments, we test two particular such cases: (i) when the true demand is logarithmic ($log(\theta^2 x+C)$), and, (ii) when the true demand is exponential ($e^{\theta^2 x}$). We summarize the results below in Table 4.3 below.

Average Error in Demand Estimation Due to Misspecification of Demand

True Demand = Logarithmic			**True Demand = Exponential**		
Exponential	Multinomial Logit	Linear	Logarithmic	Multinomial Logit	Linear
2.1%	1.7%	3.6%	1.5%	1.1%	3.3%

Table 4.3: This table summarizes the performance gaps of the proposed additive and multiplicative algorithms, as well as some hybrid algorithms, in the airline case study in percent of expected revenue attained relative to the full-knowledge Clairvoyant strategy.

The above results are based on averages across scenarios in which the true demand is misspecified and after convergence, we record the average of the error between the true value of the demand for a given consumer and its estimated value. We find that the largest error result from linear misspecifications as these are over-simplifications of the underlying modeling structure. Disregarding these cases, the average error is between 1.3% and 1.9%. While this error does not seem particularly large, as a proportion of the value of an average buy probability (which ranges from 5-12%), this is a significant proportion.

Thus having analyzed the effects on our algorithm's performance when influenced by changes in demand function, number of products, and number of consumer features, we can conclude that our approach efficiently converges to a reasonable estimate of the true distribution, as approximated by the first-order Taylor series expansion. However, we also observe that under misspecification of demand functional form, there may be error up to the order of 2% across all possible cases, which can potentially be problematic when using these estimates for personalized bundle and pricing recommendations as is done in Chapters 2 and 3. Thus, we aim to further broaden this approach by incorporating some sort of protection against demand estimation error in the following work in Section 4.3.

4.3 Robust Optimization Approach

The second method of incorporating demand uncertainty into the personalized bundle recommendation problem relies on techniques developed in robust optimization. Note the above results in Section 4.2 indicate that we have constructed a dynamic learning approach that is relatively stable with growing problem complexity through changes in demand functions, number of available products and number of consumer features. However, the results of the study from Table 4.3 demonstrate that we can expect between 1-2% error in demand function value estimation when we misspecify the true functional form of the demand. In order to better understand the magnitude of this effect and establish a method of mitigating it, as a second step we propose to incorporate a robust optimization approach to the personalized bundle recommendation problem.

4.3.1 Problem Setting and Model Formulation

As in our initial setting, we consider a monopolist online seller that makes a dynamic bundle offer to each arriving consumer who may choose to accept the offer, purchase individual items separately at full price, or choose to purchase nothing at all. If the consumer chooses to purchase either the bundle or some other collection of items at their full prices, we assume that they only purchase one unit of each item. Let us consider a set of items $i = 1, ..., n$ denoted by $\hat{S}$. These items' prices may affect one another and they can be complementary, substitutable, or even independent as is often the case in the travel industry. Given a captive online consumer considering products within $\hat{S}$, or a specific ticket itinerary for which $\hat{S}$ is the set of ancillary goods, our model offers a relevant bundle of products from $\hat{S}$. We are interested in cases where $\hat{S}$ contains inventory-constrained products that we leverage to maximize expected long-run profitability by accounting for future demand. Therefore, we consider a finite selling horizon with a fixed number of periods T with no replenishments.

Each arriving consumer is uniquely described by a combination of categorical and continuous features related to preferences, demographics, purchase history, loyalty, and online shopping context. Thus, we do not consider a discrete set of consumer types as is traditionally done in segmentation and instead assume that there is an infinite set of continuous consumer types. Furthermore, since we address a bi-level pricing problem, we index consumers within a given period t by (k, t), where $k = 1, ..., K^t$ and the total number of arrivals K^t in each period can differ. We define the full price of item i in period t as $\bar{p}_i^t$; thus, the full price $\bar{p}_{S_{k,t}}$ of a bundle $S_{k,t}$ offered to consumer (k, t) is defined by,

$$\bar{p}_{S_{k,t}} = \sum\nolimits_{i \in S_{k,t}} \bar{p}_i^t \tag{4.4}$$

The full prices $\bar{p}_i^t$ are not necessarily fixed throughout the horizon and may follow some dynamic trajectory, summarized in each period by vector $\bar{\mathbf{p}}^t = [\bar{p}_1^t,\ \bar{p}_2^t,\ \dots,\ \bar{p}_n^t]$. We thus define price vector,

$$\mathbf{p}_{S_{k,t}} = [\bar{\mathbf{p}}^t,\ p_{S_{k,t}}] = [\bar{p}_1^t,\ \bar{p}_2^t,\ \dots,\ \bar{p}_n^t,\ p_{S_{k,t}}], \tag{4.5}$$

in which we append the discounted price of the personalized bundle for consumer (k,t) to the vector of full price settings for period t. It is common in business practice for sellers to consider discrete price ladders. Therefore, we make the assumption that we have a fixed set of price levels for every product i from which we can choose to construct bundle offers. We define the individual consumer propensity-to-buy $\xi_S^{k,t}(\mathbf{p}_{S_{k,t}})$ as the probability that consumer (k,t) will purchase the combination of products $S : S \neq S_{k,t}$ (and nothing else) at their *full prices* if their personalized bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$. We similarly define the probability that consumer (k,t) will purchase only the bundle $S_{k,t}$ (and no other products) when it is offered at the *discounted price* $p_{S_{k,t}}$ as $\xi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})$. We will refer to $\mathbf{e}_{S_{k,t}}$ as the bundle unit vector that takes the value 1 for all $i \in S_{k,t}$ and 0 otherwise. Finally, we define $\mathbf{I}^{k,t}$ as the vector of inventory levels of all $i \in \hat{S}$ at the time when consumer (k,t) arrives, written explicitly as $\mathbf{I}^{k,t} = [I_1^{k,t},\ I_2^{k,t},\ \dots,\ I_n^{k,t}]$. This leads to the following decision variables for any given consumer (k,t): the optimal bundle to recommend $S_{k,t} \in \hat{S}$, and, its personalized price $p_{S_{k,t}} \leq \bar{p}_{S_{k,t}}$.

We now consider the optimal benchmarking baseline for this setting, given by the Clairvoyant problem of making a personalized bundle recommendations given a known consumer sequence $\{k,t\}_{\forall k=1,\dots,K^t}^{t=1,\dots,T}$, which is formulated as follows:

$$\begin{aligned}
\underset{y_{S_{k,t}}^{k,t}}{\text{maximize}} \quad & \sum\nolimits_{t=1}^{T} \sum\nolimits_{k=1}^{K^t} \sum\nolimits_{S_{k,t} \subset \hat{S}} \left(\left[\sum\nolimits_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \right] \right. \\
& \qquad \left. + \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}} \right) \right) \cdot y_{S_{k,t}}^{k,t} \\
\text{subject to} \quad & \sum\nolimits_{t=1}^{T} \sum\nolimits_{k=1}^{K^t} \sum\nolimits_{S_{k,t} \subset \hat{S}} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \right] \cdot y_{S_{k,t}}^{k,t} \leq I_i^0 \quad \forall i \\
& \sum\nolimits_{S_{k,t} \subset \hat{S}} y_{S_{k,t}}^{k,t} = 1 \qquad \forall (k,t) \\
& y_{S_{k,t}}^{k,t} \geq 0 \qquad \forall (k,t),\ S_{k,t} \subset \hat{S}
\end{aligned}$$

$$\text{where } \phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) = \sum\nolimits_{S\subset\hat{S}:S\supset S_{k,t}} \phi^{k,t}_{S}(\mathbf{p}_{S_{k,t}}),$$
$$\text{and, } \phi^{k,t}_{i}(\mathbf{p}_{S_{k,t}}) = \sum\nolimits_{S\subset\hat{S}:S\ni i} \phi^{k,t}_{S}(\bar{p}^{t}_{S}, p_{S_{k,t}}).$$

The decision variables $y^{k,t}_{S_{k,t},p_{S_{k,t}}}$ correspond to the probability with which bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$ to consumer (k,t) when the full product prices are set at $\bar{p}^{t}_{i}$. In this perfect information setting, assume that the entire consumer arrival sequence $\{k,t\}^{t=1,\dots,T}_{\forall k=1,\dots,K^t}$ is known in advance, as well as the full price trajectories $\bar{p}^{t}_{i}$ for all products i in all periods t, which we assume are provided by an oracle. The recommendations, denoted by $y^{k,t}_{S_{k,t},p_{S_{k,t}}}$, that are made by this model are based on the knowledge of the individual consumer propensity-to-buy $\phi^{k,t}_{i}(\mathbf{p}_{S_{k,t}})$, defined as the probability that consumer (k,t) will purchase item i if their personalized bundle $S_{k,t}$ is offered at price $p_{S_{k,t}}$. Furthermore, all algorithms benchmarked against this perfect information setting make decisions based on this same knowledge upon the arrival of each (k,t). We now want to robustify this Clairvoyant problem and analyze its performance under various uncertainty settings.

4.3.2 Analyzing Uncertainty in Demand

We begin by observing that the Clairvoyant problem is an linear optimization problem, so we approach this using robust linear optimization (RLO) methods. We first consider the general infeasibility of this offline full-knowledge problem when we marginally perturb the demand and observe the effect on the corresponding inventory constraints. We then introduce uncertainty sets and formulate the RLO problem, and analyze the performance (with respect to the price of robustness) relative to the nominal problem objective value.

Summary of Data

We consider the airline case study data in order to practically assess the performance of our various robust methodologies. We analyzed the one-month period of approximately 640,000 ticket transactions from a premier international airline. There are no repeat consumers in this short time frame and thus no details from previously purchased flight itineraries. Every transaction is described by a set of features categorized into two types: (i) personal consumer information including tier level, mileage balance, time since joining rewards, and number of previous business and economy flights taken; and, (ii) contextual itinerary booking data that includes transaction date, fare paid (USD), connection time, time to departure, day of travel, and number of passengers. For these computations we focused on the products with inventory limitations such as priority security, priority boarding, priority baggage handling, seat upgrades, checked excess baggage, VIP lounge access, and gourmet in-flight meals. We were provided with the corresponding prices for these products, which varied historically across flight itineraries and weeks. Note that in this data set the products are independent by definition since they correspond to distinct unrelated products that are neither substitutable nor complementary and are priced separately. We used k-means clustering to analyze the personalized features in the data and develop distinct consumer profiles that we used to map the historical transactions for demand model estimations. We constructed 7 unique consumer personas and estimated all of the personalized pairwise demand models for each (persona, product) combination and treat this estimation as our nominal value.

4.3.3 Perturbations in Nominal Demand

Our initial analysis is divided into three sets of cases, in each of which we assess the level of infeasibility provided by the nominal problem solution (at the nominal values of the demand function values as provided by the demand estimation).

Analyzing Feasibility Across Personas and Products

We first consider a classic exercise in the benefit of robust optimization in which we perturb the estimated nominal values that are uncertain and observe the effect of the feasibility of the nominal optimal solution. In this case we perturb the demand models by different margins and observe the results over 10,000 iterations when averaged over all products and personas. Notice that as the problem becomes increasingly inventory constrained, we intuitively see a significantly more drastic level of potential infeasibility when demand estimations are inaccurate. Furthermore, we consider these more constrained instances in order to extract the value of the demand estimation in the feasibility of the optimal solution. Were we to also consider less constrained problem instances with higher levels of initial stock, we would rarely observe infeasibility at all due to the lack of importance of inventory in such settings.

Percentage of Violated Constraints

Perturbation	Inventory 25%	Inventory 10%	Inventory 5%
0.01%	1.12%	8.99%	40.52%
0.10%	1.98%	11.34%	49.67%
0.50%	2.79%	15.27%	56.98%
1.00%	3.55%	20.86%	64.33%
1.50%	5.04%	26.19%	70.09%
2.00%	7.11%	32.05%	77.21%

Table 4.4: This table summarizes the percentage of total constraints that are violated by the nominal solution when the demand is marginally perturbed, averaged over 10,000 across all personas and products.

Now let us recall the dynamic learning results from Table 4.3 in Section 4.2. On average, we expect 1.3%-1.9% error in demand estimation due to misspecification of the true underlying demand model. The above study in Table 4.4 clearly demonstrates that in highly constrained inventory problems, this could lead to infeasibility in the range of 20-70% of problem instances. This demonstrates the absolute necessity for incorporating robustness explicitly into our modeling approach, as dynamic learning captures enough error to greatly de-stabilize and invalidate "optimal" offer solutions in particularly inventory constrained settings.

Analyzing Feasibility for Specific Products

We conduct a similar exercise but here we consider the inventory constraints as isolated by product, as shown in Table 4.5 below.

We perturb the estimated nominal demand by 0.05% and see what percentage of instances had violated that product's specific inventory constraint. We again average

Percentage of Violated Constraints

Product	Inventory 25%	Inventory 10%	Inventory 5%
Priority Security	3.84%	22.31%	56.44%
Priority Boarding	3.95%	8.32%	46.87%
Priority Baggage Handling	1.13%	6.34%	22.16%
Seat Upgrades	3.00%	20.02%	70.86%
Excess Checked Baggage	2.30%	11.23%	50.34%
VIP Lounge Access	4.05%	25.17%	87.13%
Gourmet In-Flight Meals	1.36%	17.25%	68.12%

Table 4.5: This table summarizes the percentage of total constraints that are violated by the nominal solution when the demand is perturbed by 0.05%, averaged over 10,000 across all personas. It is particularly interesting to see the effect over products as we can see which are in higher demand and how this perturbation effects their particular constraint.

over 10,000 iterations of perturbations. Note that now we see the effects of product popularity and consumer price elasticities for certain goods. Goods that have generally higher demand such as lounge access and seat upgrades are subject to much greater potential infeasibility due to demand estimation error as the setting becomes more highly constrained.

Percentage of Constraint Violations by Scale for Fixed Perturbations

Finally, we consider the percentage of instances in which we have infeasibility, and also measure the extent to which they are violated by using the metric,

$$X_i = \max\left[\frac{I_i^0 - \text{Demand}}{|I_i^0|}, 0\right],$$

for each constraint corresponding to product i, which we then average across all products (note that Demand is defined through $\hat{\phi}$, based on our nominal covariate values that come from the demand estimation models). In this case we considered highly constrained inventory stock levels of 5%.

We note again here that the extent of constraint violation is quite high, ranging from 64% to 77% for demand value perturbations of 1% to 2%. Given that this is the range of estimation error (on average from 1.3% to 1.9%) established in the results

Percentage of Violated Constraints

Perturbation	% Violated Constraints	% Violated by over 50%	% Violated by over 100%
0.01%	40.52%	50.13%	17.46%
0.10%	49.67%	53.62%	19.12%
0.50%	56.98%	57.94%	21.55%
1.00%	64.33%	61.53%	23.67%
1.50%	70.09%	64.72%	26.02%
2.00%	77.21%	69.33%	28.44%

Table 4.6: This table summarizes the percentage of total constraints that are violated by the nominal solution when the demand is perturbed by a marginal percent, averaged over 10,000 across all personas and all products. Here we narrow to a lower inventory level of 5% and consider the scale of the constraint violations by the nominal solutions on average.

in Section 4.2, we can conclude that it is critical to incorporate robust optimization into a modeling framework in which the recommendation and pricing system is based on nominal demand estimates.

4.3.4 Comparison of Uncertainty Sets

Let us revisit our initial problem, given by Formulation (2.4) before. Note that this is a linear optimization problem with row-wise uncertainty in the demand. By incorporating this uncertainty, we define the following robust version of our Clairvoyant problem:

$$
\begin{aligned}
\max_{y^{k,t}_{S_{k,t}}} \quad & \left[\min_{\phi \in \mathcal{U}} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\left[\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \right] + \phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) \cdot (p_{S_{k,t}} - \bar{p}_{S_{k,t}}) \right) \cdot y^{k,t}_{S_{k,t}} \right] \\
\text{s.t.} \quad & \max_{\phi \in \mathcal{U}'} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \right] \cdot y^{k,t}_{S_{k,t}} \leq I_i^0 \quad \forall i \\
& \sum_{S_{k,t} \subset \hat{S}} y^{k,t}_{S_{k,t}} = 1 \qquad \forall (k,t) \\
& y^{k,t}_{S_{k,t}} \geq 0 \qquad \forall (k,t),\ S_{k,t} \subset \hat{S}
\end{aligned}
\tag{4.6}
$$

Note that while this appears to look like a MIO problem, the "binary" variables $y^{k,t}_{S_{k,t}}$ in our original Clairvoyant model are relaxed, thus ultimately providing us with a linear optimization problem. We consider two types of uncertainty sets, for each of which we define the subproblems of interest and their respective contributions to the robust counterparts:

1. Polyhedral – $\mathcal{U}_i = \{a_i : a_i = \bar{a}_i + \Delta_i' u_i, ||u_i||_1 \le \rho\}$

2. Ellipsoidal – $\mathcal{U}_i = \{a_i : a_i = \bar{a}_i + \Delta_i' u_i, ||u_i||_2 \le \rho\}$

Polyhedral Uncertainty Sets

We first consider the polyhedral uncertainty set $\mathcal{U}$. The first subproblem arises from the objective value under a fixed value of $\mathbf{y}$; note that this is very similar to the second subproblem that results from the inventory constraint, but does not contain the prices from the objective function. Considering first the initial subproblem, a fixed solution $\mathbf{y}$ corresponds to a fixed discounted bundle offer and price path corresponding to known arrival sequence (k, t) over the entire selling horizon. We assume in this case that the value of each consumer (k, t)'s willingness-to-pay falls into the polyhedral uncertainty set:

$$\mathcal{U}_{k,t} = \{\phi^{k,t} : \bar{\phi}^{k,t} - \delta_{k,t}\gamma_{k,t} \le \phi^{k,t} \le \bar{\phi}^{k,t} + \delta_{k,t}\gamma_{k,t}, ||\gamma||_1 \le \rho\} \tag{4.7}$$

Using definition (4.7) of our desired uncertainty set, we establish the following as our subproblem for the personalized willingness-to-pay functions ϕ:

$$\begin{aligned}
\min_{\phi \in \mathcal{U}_{k,t}} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\left[\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \right] + \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}}\right) \right) \cdot y_{S_{k,t}}^{k,t} \\
\text{subject to} \quad & \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}\gamma_{k,t} \le \bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) && \forall i, (k,t) \\
& - \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}\gamma_{k,t} \le -\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) && \forall i, (k,t) \\
& \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}\gamma_{k,t} \le \bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) && \forall S_{k,t}, (k,t) \\
& - \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}\gamma_{k,t} \le -\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) && \forall S_{k,t}, (k,t) \\
& \sum_{t=1}^{T} \sum_{k=1}^{K^t} \gamma_{k,t} \le \rho \\
& \gamma_{k,t} \ge 0 \qquad \forall (k,t)
\end{aligned} \tag{4.8}$$

Based on this, we can now formulate this subproblem's dual in order to develop the robust counterpart to the original Clairvoyant problem:

$$
\begin{aligned}
\max_{v,u,w} \quad & \sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{S_{k,t}\subset\hat{S}} v_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - u_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) + v_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \\
& \qquad - u_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) + \rho w \\
\text{subject to} \quad & v_{i,k,t} - u_{i,k,t} = \bar{p}_i && \forall i,(k,t) \\
& v_{S,k,t} - u_{S,k,t} = p_S - \bar{p}_{S_{k,t}} && \forall S_{k,t},(k,t) \\
& -\delta_{k,t}v_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}u_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}v_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \\
& \qquad - \delta_{k,t}u_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) + w \geq 0 && \forall i, S_{k,t},(k,t) \\
& v_{i,k,t}, v_{S,k,t}, u_{i,k,t}, u_{S,k,t}, w \geq 0 && \forall (i,S,(k,t)
\end{aligned}
\tag{4.9}
$$

Note that the second subproblem that corresponds to the demand uncertainty in the inventory constraints is trivial and a simplified case of the first subproblem, with two less decision variables. Combining the above dual from (4.9) with the constraint subproblem's dual, when substituted back into the original problem to replace the demand uncertainty, ultimately provides us with the following robust counterpart with a polyhedral uncertainty set:

$$
\begin{aligned}
\max_{y_{S_{k,t}}^{k,t}} \quad & \Bigg\{ \max_{v,v',u,u',w,w'} \sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{S_{k,t}\subset\hat{S}} v_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - u_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) + v_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \\
& \qquad - u_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) + \rho w \Bigg\} \\
\text{subject to} \quad & \textstyle\sum_{S_{k,t}\subset\hat{S}} y_{S_{k,t}}^{k,t} = 1 && \forall (k,t) \\
& y_{S_{k,t}}^{k,t} \geq 0 && \forall (k,t),\ S_{k,t}\subset\hat{S} \\
& v_{i,k,t} - u_{i,k,t} = \bar{p}_i && \forall i,(k,t) \\
& v_{S,k,t} - u_{S,k,t} = p_S - \bar{p}_{S_{k,t}} && \forall S_{k,t},(k,t) \\
& -\delta_{k,t}v_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}u_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}v_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \\
& \qquad - \delta_{k,t}u_{S,k,t}\bar{\phi}_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) + w \geq 0 && \forall i, S_{k,t},(k,t) \\
& \sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{S_{k,t}\subset\hat{S}} v_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - u_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) + \rho w \leq I_i^0 && \forall i \\
& v'_{i,k,t} - u'_{i,k,t} = 1 && \forall i,(k,t) \\
& -\delta_{k,t}v'_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \delta_{k,t}u'_{i,k,t}\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) + w' \geq 0 && \forall i, S_{k,t},(k,t) \\
& v_{i,k,t}, v_{S,k,t}, v'_{i,k,t}, u_{i,k,t}, u_{S,k,t}, u_{i,k,t}, w, w' \geq 0 && \forall i, S,(k,t)
\end{aligned}
\tag{4.10}
$$

The resulting robust counterpart in (4.10) is also a linear program under polyhedral uncertainty, and thus also possible to solve using the same methods as the original Clairvoyant problem. We demonstrate the empirical performance of this approach, compared to the ellipsoidal uncertainty set counterpart that we derive below, on the airline case study data from the previous chapter. The result are summarized in Section 4.3.4 below.

Ellipsoidal Uncertainty Sets

We similarly derive a robust counterpart for the Clairvoyant problem using an ellipsoidal uncertainty set $\mathcal{U}$, based on the approach above but using the L2 norm. This gives us the following uncertainty set:

$$\mathcal{U}_{k,t} = \{\phi^{k,t} : \bar{\phi}^{k,t} - \delta_{k,t}\gamma_{k,t} \leq \phi^{k,t} \leq \bar{\phi}^{k,t} + \delta'_{k,t}\gamma_{k,t}, ||\gamma||_2 \leq \rho\} \tag{4.11}$$

We then consider the vector notation of the uncertainty set definition, as defined below, so that we can apply a result from robust optimization theory in constructing robust counterparts using ellipsoidal uncertainty sets.

$$\begin{aligned}
\mathcal{U}_{k,t} &= \{\phi^{k,t} : \bar{\phi}^{k,t} - \delta_{k,t}\gamma_{k,t} \leq \phi^{k,t} \leq \bar{\phi}^{k,t} + \delta_{k,t}\gamma_{k,t},\ ||\gamma||_2 \leq \rho\} \\
&= \{\vec{\phi} : \phi^{k,t} = \bar{\phi}^{k,t} + \delta_{k,t}\gamma_{k,t},\ ||\gamma||_2 \leq \rho\} \\
&= \{\vec{\phi} : \vec{\phi} = \vec{\bar{\phi}} + \vec{\Delta}'\vec{\gamma},\ ||\gamma||_2 \leq \rho\}
\end{aligned}$$

We make use of a well-established result, which allows us to re-write the desired subproblem for the inventory constraint:

$$\begin{aligned}
&\max_{\phi \in \mathcal{U}_{k,t}} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\right) \cdot y_{S_{k,t}}^{k,t} \\
&= \max_{\phi \in \mathcal{U}_{k,t}} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\vec{\phi}\,'\mathbf{y}\right) \\
&= \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left((\vec{\bar{\phi}} + \vec{\Delta}'\vec{\gamma})\,'\mathbf{y} + \rho||\Delta\mathbf{y}||_2\right) \\
&= \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot y_{S_{k,t}}^{k,t} + \rho\sqrt{\sum_{t=1}^{T}\sum_{k=1}^{K^t} \delta_{k,t}^2 (y_{S_{k,t}}^{k,t})^2}\right)
\end{aligned} \tag{4.12}$$

Using an identical process to that above, we can formulate the subproblem for the objective function as well, and thus obtain the following robust counterpart to the Clairvoyant problem using ellipsoidal uncertainty sets:

$$
\begin{aligned}
\max_{y^{k,t}_{S_{k,t}}} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\left[\sum_{i=1}^{n} \bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot y^{k,t}_{S_{k,t}} \right] \right. \\
& \left. + \bar{\phi}^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) \cdot (p_{S_{k,t}} - \bar{p}_{S_{k,t}}) \cdot y^{k,t}_{S_{k,t}} + \rho \sqrt{\sum_{t=1}^{T} \sum_{k=1}^{K^t} \delta^2_{k,t} (y^{k,t}_{S_{k,t}})^2} \right) \\
\text{s.t.} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\bar{\phi}_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot y^{k,t}_{S_{k,t}} + \rho \sqrt{\sum_{t=1}^{T} \sum_{k=1}^{K^t} \delta^2_{k,t} (y^{k,t}_{S_{k,t}})^2} \right) \leq I_i^0 \qquad \forall i \\
& \sum_{S_{k,t} \subset \hat{S}} y^{k,t}_{S_{k,t}} = 1 \qquad \forall (k,t) \\
& y^{k,t}_{S_{k,t}} \geq 0 \qquad \forall (k,t),\ S_{k,t} \subset \hat{S}
\end{aligned}
\tag{4.13}
$$

Comparison

We now compare these formulations across different levels of inventory constraints and observe the price of robustness that comes from decreasing the objective function relative to the nominal optimal solution, over various levels of robustness ρ. Specifically, we use the well-known theoretical that the probability of infeasibility is bounded by ϵ when $\rho = \sqrt{2 * (ln(\frac{1}{\epsilon}))}$. This provides us with increasing probability of feasibility as ρ increases. We first consider the case where we have an initially constraining inventory level of 10%. We obtain the following results over 10,000 iterations. This is further visualized in Figure 4-11 below.

We observe that the ellipsoidal uncertainty set achieves better results in terms of price of robustness overall and dominates the polyhedral approach. We also see that as we drastically increase robustness and implement an overly conservative approach under highly constrained inventory, we suffer more in objective value because the model prices bundles much lower and lose more revenue in margin when attempting to incentivize consumers to buy.

We achieve a similar result from the more highly inventory-constrained setting in

Price of Robustness at 10% Inventory Level

ρ	$\mathbb{P}$ of Feasibility	Polyhedral $\mathcal{U}$ % of Obj Value	Polyhedral $\mathcal{U}$ % Actual Feas.	Ellipsoidal $\mathcal{U}$ % of Obj Value	Ellipsoidal $\mathcal{U}$ % Actual Feas.
0	66%	100%	75.4%	100%	76.8%
1.67	75%	92.3%	80.1%	95.6%	82.2%
1.79	80%	90.4%	85.5%	91.1%	87.0%
2.15	90%	81.3%	90.2%	85.8%	92.1%
3.03	99%	76.6%	95.6%	81.6%	96.7%
3.26	99.5%	75.4%	97.3%	79.1%	98.4%
3.72	99.9%	71.8%	100%	75.4%	100%
4.29	99.99%	70.9%	100%	72.3%	100%
4.80	99.999%	62.7%	100%	68.1%	100%

Table 4.7: This table summarizes the percentage of total constraints that are violated by the nominal solution when the demand is marginally perturbed, averaged over 10,000 across all personas and products.

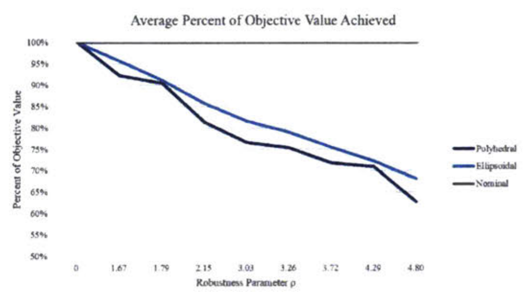

Figure 4-11: This shows the tradeoff in objective value with increasingly conservative robust parameter ρ when inventory is initialized at 10%.

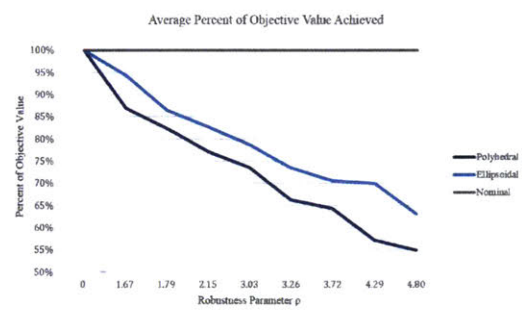

Figure 4-12: This shows the tradeoff in objective value with increasingly conservative robust parameter ρ when inventory is initialized at 5%.

which the stock is initialized to 5%. Notice that in Figure 4-12, the price becomes even steeper. This is primarily motivated by the pricing aspect of the model, which adjusts for worst-case realizations of demand and does not price accordingly, resulting in significantly lower revenues because consumers are not being used to full paying potential.

4.4 Conclusions

As is increasingly the case in the majority of online retail settings, the introduction of new products and arrival of new consumers to the market results in a lack of historical data from which consumer preferences and demand can be adequately estimated. We adapt our model of personalized bundle pricing and recommendation in two settings in order to, (i) incorporate dynamic demand learning within the objective of long-term profitability over the selling horizon and (ii) incorporate robust optimization methods. Under setting (i) we construct a generalized modeling framework for this problem using contextual non-linear multi-armed bandits. To adapt this model to practical implementation, we develop an approximation method based the coupling of the first-order Taylor series of the expected reward function with upper confidence bound approaches. Our method requires no assumptions on the functional form of the demand other than it being differentiable and generalizes many of the convex optimization methods in the existing literature. We provide analytical guarantees on the asymptotic behavior of our algorithm's regret relative to an oracle strategy that knows the true demand distribution in advance and show that the regret is on the order of $\tilde{O}(\sqrt{T})$, which is independent of the number of products and bundles. We also present empirical results that show the robust performance of this algorithm over relevant benchmarks from the existing literature across various functional forms of demand and numbers of available products, as well as over different lengths consumer arrival sequences. However, we also find that in the cases of demand function misspecification, the learning model may produce estimation errors that on average fall into the range of 1-2% in nominal value. Therefore, it is necessary to capture the uncertainty in the estimated demand values in order to develop a generally robust modeling approach to our recommendation problem. In setting (ii) we address this by constructing and analyzing the robust counterparts to our personalized bundle recommendation model under both polyhedral and ellipsoidal uncertainty sets. Our results demonstrate that under highly constrained inventory setting it is crucial to implement robust optimization in the Clairvoyant problem in order to account for even very minor errors in demand estimation. More specifically, perturbations in nominal demand values on the order of the 1% to 2% error resulting from demand learning errors can result in infeasibility of the recommended offers in up to 70% of problem instances when inventory is highly constrained. We also gain interesting insights into the price of robustness as related to the underlying recommendation problem structure. We find that the incorporation of a pricing problem also shows that overly conservative methods steeply increase the price of robustness because the bundle offers are made with unnecessarily low discounts, sacrificing margin that could potentially be extracted from consumers.

Bibliography

[1] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, pages 2312–2320, 2011.

[2] Elodie Adida and Georgia Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming*, 107(1):97–129, 2006.

[3] Shipra Agrawal, Vashist Avadhanula, Vineet Goyal, and Assaf Zeevi. A near-optimal exploration-exploitation approach for assortment selection under a multinomial logit model.

[4] Shipra Agrawal and Nikhil Devanur. Linear contextual bandits with knapsacks. In *Advances in Neural Information Processing Systems*, pages 3450–3458, 2016.

[5] Shipra Agrawal and Nikhil R Devanur. Bandits with concave rewards and convex knapsacks. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 989–1006. ACM, 2014.

[6] Shipra Agrawal, Nikhil R Devanur, and Lihong Li. An efficient algorithm for

contextual bandits with knapsacks, and an extension to concave objectives. In *Conference on Learning Theory*, pages 4–18, 2016.

[7] Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135, 2013.

[8] Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

[9] Yossi Aviv and Amit Pazgal. A partially observed markov decision process for dynamic pricing. *Management Science*, 51(9):1400–1416, 2005.

[10] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. Bandits with knapsacks. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 207–216. IEEE, 2013.

[11] Ashwinkumar Badanidiyuru, John Langford, and Aleksandrs Slivkins. Resourceful contextual bandits. In *Conference on Learning Theory*, pages 1109–1134, 2014.

[12] Hamsa Bastani and Mohsen Bayati. Online decision-making with high-dimensional covariates. 2015.

[13] Dimitris Bertsimas, Dessislava Pachamanova, and Melvyn Sim. Robust linear optimization under general norms. *Operations Research Letters*, 32(6):510–516, 2004.

[14] Dimitris Bertsimas and Georgia Perakis. Dynamic pricing: A learning approach. *Mathematical and computational models for congestion charging*, pages 45–79, 2006.

[15] Dimitris Bertsimas and Melvyn Sim. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107(1-2):5–36, 2006.

[16] Dimitris Bertsimas and Phebe Vayanos. Data-driven learning in dynamic pricing using adaptive optimization. *Preprint*, 2015.

[17] Felipe Caro and Jérémie Gallien. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science*, 53(2):276–292, 2007.

[18] Alexandre X Carvalho and Martin L Puterman. Learning and pricing in an internet environment with binomial demands. *Journal of Revenue and Pricing Management*, 3(4):320–336, 2005.

[19] Alexandre X Carvalho and ML Puterman. Dynamic pricing and reinforcement learning. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, volume 4, pages 2916–2921. IEEE, 2003.

[20] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In *Advances in neural information processing systems*, pages 2249–2257, 2011.

[21] Zhi Chen, Lei Wu, and Yong Fu. Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. *IEEE Transactions on Smart Grid*, 3(4):1822–1831, 2012.

[22] Byung Do Chung, Tao Yao, Terry L Friesz, and Hongcheng Liu. Dynamic congestion pricing with demand uncertainty: a robust optimization approach. *Transportation Research Part B: Methodological*, 46(10):1504–1518, 2012.

[23] Kris Johnson Ferreira, David Simchi-Levi, and He Wang. Online network revenue management using thompson sampling. 2016.

[24] Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European journal of operational research*, 235(3):471–483, 2014.

[25] John Gittins, Kevin Glazebrook, and Richard Weber. *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.

[26] Pavithra Harsha, Anna Papush, Georgia Perakis, and Divya Singhvi. Personalized bundle recommendation under demand uncertainty. 2018.

[27] Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On bayesian upper confidence bounds for bandit problems. In *Artificial Intelligence and Statistics*, pages 592–600, 2012.

[28] N Bora Keskin and Assaf Zeevi. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62(5):1142–1167, 2014.

[29] Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

[30] Sandeep Pandey, Deepayan Chakrabarti, and Deepak Agarwal. Multi-armed bandit problems with dependent arms. In *Proceedings of the 24th international conference on Machine learning*, pages 721–728. ACM, 2007.

[31] Georgia Perakis and Guillaume Roels. Robust controls for network revenue management. *Manufacturing & Service Operations Management*, 12(1):56–76, 2010.

[32] Georgia Perakis and Anshul Sood. Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming*, 107(1-2):295–335, 2006.

[33] Lijing Qin, Shouyuan Chen, and Xiaoyan Zhu. Contextual combinatorial bandit and its application on diversified online recommendation. In *Proceedings of the 2014 SIAM International Conference on Data Mining*, pages 461–469. SIAM, 2014.

[34] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

[35] Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039, 2014.

[36] Paat Rusmevichientong and Huseyin Topaloglu. Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research*, 60(4):865–882, 2012.

[37] Daniel Russo and Benjamin Van Roy. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.

[38] Masoud Talebian, Natashia Boland, and Martin Savelsbergh. Pricing to accelerate demand learning in dynamic assortment planning for perishable products. *European Journal of Operational Research*, 237(2):555–565, 2014.

[39] Aurélie Thiele. Single-product pricing via robust optimization. *submitted to Operations Research*, 2006.

[40] Joannes Vermorel and Mehryar Mohri. Multi-armed bandit algorithms and empirical evaluation. In *ECML*, volume 3720, pages 437–448. Springer, 2005.

Chapter 5

Conclusions

The work in this thesis spans several prevalent fields of research within the growing revenue management literature, including personalized assortment optimization, assortment planning under inventory constraints, dynamic pricing, cross-selling and online demand learning. Specifically, we study the intersection of these methodologies within the context of data-driven personalized bundle pricing and product recommendation in the online setting.

As demonstrated by leading market forecasts, the online channel stands to inherit a significant proportion of the market across all industries, and is also a rapidly growing avenue of opportunity for any business interested in offering customized consumer experiences. Thus, gaining the competitive edge this sector is of utmost importance to any firm's online success. We delve into this problem by developing a new modeling approach that combines many existing branches of literature and is thus innovative from an analytical standpoint and potentially promising from a practical one. Having constructed the analytical framework, it is also important to emphasize the need for tractability due to the particular context of this problem. Therefore, the development of approximation algorithms is crucial in order to apply this complex model to the online setting as originally intended. We find that the analytical guarantees of such approximation approaches are often conservative relative to their actual empirical performance on real data.

In addition to providing analytical insights into the structure of this complicated problem, it is important to also demonstrate the practical performance of approximation methods in order to justify the use of these approaches in business implementations. The case study analyses show that these algorithms are not only viable by providing real-time outputs, but also perform significantly better than their analytical guarantees and can obtain expected revenues of up to 98% of a full-knowledge oracle strategy that is an unattainable benchmark in practice. Developing methods whose outputs that are not only efficient in real-time but also of good quality is imperative to practical applications. Furthermore, in-depth analysis of each case study in detail allowed us to extract a wide range of insights that could potentially be of great use to online sellers employing personalized targeting and pricing strategies.

Lastly, we extend the original framework of our problem to the setting where demand may be subject to various forms of uncertainty, which is increasingly becoming the case in many business practices with an online channel. We develop two approaches in this context, based on (i) dynamic learning and (ii) robust optimization. For the dynamic learning setting we construct a generalized approach in the form of a model for personalized bundle offers that incorporates learning, along with a corresponding approximation algorithm that applies to very general demand functions. This method generalizes many current optimization approaches for dynamic learning in the literature and provides analytical guarantees on the asymptotic behavior of the regret, which is important to developing practical applications. Our empirical studies confirm that this approach is effective and outperforms existing baseline methods, and is thus a feasible and promising option for business implementation. However, we also establish that demand learning is inherently subject to error, particularly in a personalized dynamic setting. Therefore, we extend our analysis of demand uncertainty to the robust optimization setting, in which we develop the robust counterparts to the product recommendation problem under both polyhedral and ellipsoidal uncertainty sets. We demonstrate that under increasingly constrained inventory settings, it is crucial to capture demand uncertainty to mitigate the effects of error in estimation. Furthermore, we conclude that the cost of robustness is marginal relative to the likelihood of infeasibility in these scenarios.

Thus, this thesis demonstrates the benefits of adapting realistic business problems to revenue management in order to both expand analytical developments in this field, as well as develop methods to improve and significantly impact current business operations across a wide range of industries. A potential extension to this work could include the consideration of other learning algorithms, as well as the incorporation of external factors such as competition between firms and their corresponding effect on personalized pricing strategies.

Appendix A

Appendix of Chapter 2

A.1 Modeling Notation

We provide the following summary of model parameters for the dynamic programming formulation of the personalized bundle pricing and recommendation model presented in Section 2.2. We provide the formulation below again for reference:

$$
\begin{aligned}
\underset{S_{k,t} \subset \hat{S},\, p_{S_{k,t}},\, \bar{p}_i^t}{\text{maximize}} \quad & V_{k,t}(\mathbf{I}^{k,t}) \\
\text{subject to} \quad & V_{k,t}(\mathbf{I}^{k,t}) = \sum_{S \subset \hat{S}} \xi_S^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \Bigg(\Big(\sum_{i \in S} \bar{p}_i^t \Big) \\
& \qquad + \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}}\right) \cdot \mathbb{1}_{\{S_{k,t} \subset S\}} + V_{k+1,t}(\mathbf{I}^{k,t} - \mathbf{e}_S) \Bigg) \\
& V_{K^t+1,t}(\mathbf{I}) = V_{1,t+1}(\mathbf{I}) \qquad \forall K^t,\ t = 1, ..., T \\
& (1-\epsilon)\bar{p}_{S_{k,t}} \leq p_{S_{k,t}} \leq \bar{p}_{S_{k,t}} \qquad \forall (k,t), S_{k,t} \subset \hat{\mathcal{S}}
\end{aligned}
$$

Data Parameters for Dynamic Programming Formulation	
$\hat{S}$	A demand group; related items whose respective demand depends on the prices of the other items. This set is indexed by $i = 1, ..., n$ and can include substitutable, complementary and independent items.
K^t	The total number of consumers that arrive during period t, indexed by $k = 1, ..., K^t$.
$\bar{p}_i^t$	The nominal price of a given product $i \in \hat{S}$ during period t.
$\bar{p}_{S_{k,t}}$	The nominal price of bundle $S_{k,t}$; a scalar value given by $\bar{p}_{S_{k,t}} = \sum_{i \in S_{k,t}} \bar{p}_i^t$.
$\bar{\mathbf{p}}^t$	The vector of all nominal prices $[\bar{p}_1^t \quad \bar{p}_2^t \quad ... \quad \bar{p}_n^t]$ for the demand group $\hat{S}$ during period t.
$\bar{\mathbf{p}}_{S_{k,t}}$	The vector appended with the nominal price of the bundle: $\bar{\mathbf{p}}_{S_{k,t}} = [\bar{\mathbf{p}}^t \quad \bar{p}_{S_{k,t}}] = [\bar{p}_1^t \quad \bar{p}_2^t \quad ... \quad \bar{p}_n^t \quad \bar{p}_{S_{k,t}}]$.
$\mathbf{p}_{S_{k,t}}$	The vector appended with the discounted price of the bundle: $\mathbf{p}_{S_{k,t}} = [\bar{\mathbf{p}}^t \quad p_{S_{k,t}}] = [\bar{p}_1^t \quad \bar{p}_2^t \quad ... \quad \bar{p}_n^t \quad p_{S_{k,t}}]$.
$\xi_S^{k,t}\left(\mathbf{p}_{S_{k,t}}\right)$	The probability that consumer k will purchase product set S if the personalized bundle $S_{k,t}$ is offered at the price $p_{S_{k,t}}$.
$\mathbf{e}_S$	A bundle (or single product) unit vector that is 0 for all $i \notin S$ and 1 for all $i \in S$.
$\mathbf{I}^{k,t}$	Inventory of all SKUs at time of arrival of consumer k during period t. This a vector representing the state of the system, explicity defined by $\mathbf{I}^{k,t} = [I_1^{k,t} \quad I_2^{k,t} \quad ... \quad I_n^{k,t}]$.
Decision Variables (Output)	
$S_{k,t}$	The optimal bundle to recommend for consumer k at their arrival during period t.
$p_{S_{k,t}}$	The optimal price of the recommended bundle $S_{k,t}$ for consumer k during period t.

Table A.1: This table summarizes the data parameters and the decision variables for the full model formulation written in (A.1) above.

A.2 Proofs of Multiplicative Algorithm Results

A.2.1 Proof of Theorem 1

Proof of Lower Bound:

Proof. Proof of Theorem 1 We are interested in providing an analytical guarantee on the competitive ratio between the multiplicative algorithm and the optimal clairvoyant strategy. Specifically, we want to attain a lower bound on the performance of the following model, whose objective we will now refer to as $\{\text{MultAlg}\}_{\forall(k,t)}$:

$$\begin{aligned}
\{\text{MultAlg}\}_{\forall(k,t)} = \underset{S_{k,t} \subset \hat{S},\ p_{S_{k,t}}}{\text{maximize}} \quad & \Bigg\{ \left[\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot \psi\left(\frac{I_i^{k,t}}{I_i^0}\right) \right] \\
& + \quad \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot (p_{S_{k,t}} - \bar{p}_{S_{k,t}}) \cdot \min_{i \in S_{k,t}} \psi\left(\frac{I_i^{k,t}}{I_i^0}\right) \Bigg\} \\
\text{subject to} \quad & (1-\epsilon)\bar{p}_{S_{k,t}} \leq p_{S_{k,t}} \leq \bar{p}_{S_{k,t}} \qquad \forall(k,t), S_{k,t} \subset \hat{S}
\end{aligned}$$

For any given sequence of customers $\{k,t\}_{\forall(k,t),t=1}^{T}$, we have the following primal $\{\text{Clairvoyant}\}$ problem that has full knowledge of all arrival types in advance, as presented in Section 2.2:

$$
\begin{aligned}
\underset{y^{k,t}_{S_{k,t}}}{\text{maximize}} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left(\left[\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \right] \right. \\
& \qquad \left. + \phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) \cdot \left(p_{S_{k,t}} - \bar{p}_{S_{k,t}} \right) \right) \cdot y^{k,t}_{S_{k,t}} \\
\text{subject to} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K^t} \sum_{S_{k,t} \subset \hat{S}} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \right] \cdot y^{k,t}_{S_{k,t}} \leq I_i^0 \quad \forall i \\
& \sum_{S_{k,t} \subset \hat{S}} y^{k,t}_{S_{k,t}} = 1 \qquad \forall (k,t) \\
& y^{k,t}_{S_{k,t}} \geq 0 \qquad \forall (k,t),\ S_{k,t} \subset \hat{S}
\end{aligned} \tag{A.1}
$$

By weak duality we aim to find the following lower bound on the competitive ratio between our algorithm and the clairvoyant primal problem:

$$
\frac{\{\text{MultAlg}\}_{\forall (k,t)}}{\{\text{Clairvoyant}\}} \geq \frac{\{\text{MultAlg}\}_{\forall (k,t)}}{\{\text{Dual}\}}
$$

We let the price of the bundle $S_{k,t}$ offered to consumer (k,t) be $p_{S_{k,t}}$, defined explicitly by the bundle discount price ratio $d_{S_{k,t}}$ as follows,

$$
d_{S_{k,t}} = \frac{p_{S_{k,t}}}{\bar{p}_{S_{k,t}}}.
$$

Thus in order to derive the desired bound on the ratio of the primal problem using weak duality, we consider its dual given by $\{\text{Dual}\}_{\forall (k,t)}$ below,

$$
\begin{aligned}
\min_{\theta_i^t, \lambda^{k,t}} \quad & \sum_{i=1}^{n} I_i^0 \cdot \theta_i + \sum_{t=1}^{T} \sum_{k \in K^t} \lambda^{k,t} \\
\text{subject to} \quad & \lambda^{k,t} \geq \sum_{i=1}^{n} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})(\bar{p}_i^t - \theta_i) \right] \\
& \quad + \sum_{i \in S_{k,t}} \left[\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot (d_{S_{k,t}} - 1) \right] \quad \forall (k,t), S_{k,t} \subset \hat{S} \\
& \theta_i \geq 0 \quad \forall i
\end{aligned} \tag{A.2}
$$

For the dual problem in equation (A.2), based on the choices of consumers in the sequence $\{k,t\}^{T}_{\forall (k,t), t=1}$, we utilize the result from Proposition 1 to consider the following dual feasible solution, where I_i^0 is the initial inventory of product i and $\bar{p}_i^0$ is the corresponding initial nominal price setting:

$$\hat{\theta}_i = \bar{p}_i^0 \left(1 - \psi\left(\frac{I_i^T}{I_i^0}\right)\right), \qquad \forall i$$

$$\hat{\lambda}^{k,t} = \sum_{i=1}^{n} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot \psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] + \sum_{i \in S_{k,t}} \left[\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot (d_{S_{k,t}} - 1)\right], \ \forall (k,t)$$

We now want to find the expected value of this dual feasible solution as it will give us an upper bound on the expected objective of the primal problem by weak duality. Since we have a fixed sequence $\{k,t\}_{\forall(k,t),t=1}^{T}$ the expectation is taken relative to each consumer's purchase decision, given the current state of inventory $\{I_1^{k,t}, I_2^{k,t}, ..., I_n^{k,t}\}$. By Lemma 4, we obtain the following expression for the expectation over the dual feasible variables $\hat{\lambda}^{k,t}$:

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t} \hat{\lambda}^{k,t}\right] = \sum_{i=1}^{n}\left(\sum_{t=1}^{T} \sum_{l=I_i^t+1}^{I_i^{t-1}} \bar{p}_i^t \cdot \psi\left(\frac{l}{I_i^0}\right)\right) - \sum_{t=1}^{T} M^t,$$

We define the time-dependent constant M^t with the following expression:

$$M^t = \max_{S_{k,t} \subset \hat{S}, d_{S_{k,t}}} \sum_{k=1}^{K^t} \sum_{i \in S_{k,t}} \phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot (1 - d_{S_{k,t}})$$

We thus get the following form for our expected dual objective denoted {Dual}:

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t} \hat{\lambda}^{k,t} + \sum_{i=1}^{n} I_i^0 \cdot \hat{\theta}_i\right] = \sum_{i=1}^{n}\left[\sum_{t=1}^{T} \bar{p}_i^t \sum_{l=I_i^t+1}^{I_i^{t-1}} \psi\left(\frac{l}{I_i^0}\right) + I_i^0 \cdot \bar{p}_i^0 \left(1 - \psi\left(\frac{I_i^T}{I_i^0}\right)\right)\right] - \sum_{t=1}^{T} M^t$$

We want to now compare the expected objective values of the dual problem calculated above to the expected value of the proposed heuristic approach, which we defined as $\{\text{MultAlg}\}_{\forall(k,t)}$. The expected revenue can be written as follows from Proposition 5, denoted $\{\text{MultAlg}\}_{\forall(k,t)}$:

$$\sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{i=1}^{n} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \bar{p}_i^t - \sum_{t=1}^{T} M^t$$

We can now revisit the original goal to use weak duality and finally derive the following desired ratio:

$$\frac{\{\text{MultAlg}\}_{\forall(k,t)}}{\{\text{Clairvoyant}\}} \geq \frac{\sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{t=1}^{T}M^t}{\sum_{i=1}^{n}\left[\sum_{t=1}^{T}\bar{p}_i^t\sum_{l=I_i^t+1}^{I_i^{t-1}}\psi\left(\frac{l}{I_i^0}\right) + I_i^0\cdot\bar{p}_i^0\left(1-\psi\left(\frac{I_i^T}{I_i^0}\right)\right)\right] - \sum_{t=1}^{T}M^t}$$

However, this bound is path-dependent and relies on knowledge of the final inventory levels in order to calculate a value. We want to now develop a bound that depends solely on the initial conditions to compare our algorithm to the clairvoyant approach. We therefore work to bound it further to develop a worst-case analytical guarantee that is dependent only on initial inventory levels and expected demand (by using arrival rate estimates for consumer types to calculate $\sum_{t=1}^{T}\sum_{k=1}^{K^t}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})$). We recall the time-dependent price trajectory definitions:

$$\begin{aligned}
\alpha_i^t &= \frac{\hat{p}_i^t}{\bar{p}_i^0} \quad \forall i,t, \text{ as determined by formulation (2.8) of the upper-level problem} \\
&\qquad\qquad \text{in Section 2.3.1,} \\
\beta_i^t &= \frac{\bar{p}_i^t}{\bar{p}_i^0} \quad \forall i,t, \text{ as determined by formulation (2.4) of the Clairvoyant problem.}
\end{aligned} \tag{A.3}$$

Based on the above expression, α_i^t and β_i^t are the extent of the discount on the full price of item i in period t from its initial setting at $\bar{p}_i^0$, which is common to both the Clairvoyant and our upper-level method from formulation 2.8. Note that both algorithms are provided with these nominal price discounts in advance. Thus, we get the result below:

$$\begin{aligned}
&\frac{\{\text{MultAlg}\}_{\forall(k,t)}}{\{\text{Clairvoyant}\}} \geq \\
&\geq \frac{\sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\hat{p}_i^t - \sum_{t=1}^{T}M^t}{\sum_{i=1}^{n}\left[\sum_{t=1}^{T}\bar{p}_i^t\cdot\sum_{l=I_i^t+1}^{I_i^{t-1}}\psi\left(\frac{l}{I_i^0}\right) + I_i^0\cdot\bar{p}_i^0\left(1-\psi\left(\frac{I_i^T}{I_i^0}\right)\right)\right] - \sum_{t=1}^{T}M^t} \\
&= \frac{\sum_{i=1}^{n}\bar{p}_i^0\cdot\sum_{t=1}^{T}\sum_{k=1}^{K^t}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\alpha_i^t - \frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T}M^t}{\sum_{i=1}^{n}\bar{p}_i^0\cdot\left[\sum_{t=1}^{T}\beta_i^t\cdot\sum_{l=I_i^t+1}^{I_i^{t-1}}\psi\left(\frac{l}{I_i^0}\right) + I_i^0\left(1-\psi\left(\frac{I_i^T}{I_i^0}\right)\right)\right] - \frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T}M^t}, \quad \text{by Eq. (A.3),} \\
&\geq \min_{(I_i^0,I_i^T):I_i^T\leq I_i^0-1} \frac{\sum_{t=1}^{T}\sum_{k=1}^{K^t}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\alpha_i^t - \frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T}M^t}{\sum_{t=1}^{T}\beta_i^t\cdot\sum_{l=I_i^t+1}^{I_i^{t-1}}\psi\left(\frac{l}{I_i^0}\right) + I_i^0\left(1-\psi\left(\frac{I_i^T}{I_i^0}\right)\right) - \frac{1}{\sum_{i=1}^{n}\bar{p}_i^0}\sum_{t=1}^{T}M^t}, \quad \text{Lemma 6.}
\end{aligned}$$

By considering the discount factors α_i^t and β_i^t for each product i in period t, we are able to isolate the constant $\bar{p}_i^0$ and reduce the outer summation using a minimization, as a result of Lemma 6. We now introduce a change of variable by considering $x = \frac{I_i^T}{I_i^0}$ and get the following equivalent expressions.

$$\min_{(I_i^0, x): x \leq 1 - \frac{1}{I_i^0}} \frac{\sum_{t=1}^{T} \sum_{k=1}^{K^t} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \alpha_i^t - \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}{\sum_{t=1}^{T} \beta_i^t \cdot \sum_{l=I_i^t+1}^{I_i^{t-1}} \psi\left(\frac{l}{I_i^0}\right) + I_i^0 \left(1 - \psi(x)\right) - \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}$$

$$= \min_{(I_i^0, x): x \leq 1 - \frac{1}{I_i^0}} \frac{\frac{l}{I_i^0} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \alpha_i^t - \frac{l}{I_i^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}{\frac{l}{I_i^0} \sum_{t=1}^{T} \beta_i^t \cdot \sum_{l=I_i^t+1}^{I_i^{t-1}} \psi\left(\frac{l}{I_i^0}\right) + 1 - \psi(x) - \frac{l}{I_i^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}$$

In the second expression we scale all of the terms by $\frac{1}{I_i^0}$, so that we can apply the property below, which is the result of Lemma 7.

$$\frac{1}{I_i^0} \sum_{l=I_i^T+1}^{I_i^0} \psi\left(\frac{l}{I_i^0}\right) \leq \frac{1}{I_i^0} + \int_{\frac{I_i^T+1}{I_i^0}}^{1} \psi(y) dy$$

By applying this to the previous expression we get the following result,

$$\min_{(I_i^0, x): x \leq 1 - \frac{1}{I_i^0}} \frac{\frac{l}{I_i^0} \sum_{t=1}^{T} \sum_{k=1}^{K^t} \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \alpha_i^t - \frac{l}{I_i^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}{\frac{l}{I_i^0} + \sum_{t=2}^{T} \beta_i^t \cdot \int_{l=I_i^t+1}^{I_i^{t-1}} \psi(y) dy + 1 - \psi(x) - \frac{l}{I_i^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}$$

Finally we introduce $I_{\mathsf{min}}^0 = \min_i I_i^0$ (and symmetrically also I_{max}^0, and β_{max}). We also define $R_{\mathsf{min}}^{k,t} = \min_i \ \phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \cdot \alpha_i^t$ and by definition conclude the following result,

$$\min_{(I_i^0, x): x \leq 1 - \frac{1}{I_i^0}} \frac{\frac{l}{I_{\mathsf{max}}^0} \sum_{t=1}^{T} \sum_{k=1}^{K^t} R_{\mathsf{min}}^{k,t} - \frac{l}{I_{\mathsf{min}}^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}{\beta_{\mathsf{max}} \cdot \frac{l}{I_{\mathsf{min}}^0} \cdot \int_{x=1+\frac{1}{I_{\mathsf{min}}^0}}^{1} \psi(y) dy + 1 - \psi(x) - \frac{l}{I_{\mathsf{max}}^0} \cdot \frac{1}{\sum_{i=1}^{n} \bar{p}_i^0} \sum_{t=1}^{T} M^t}$$

This completes the proof of Theorem 1. $\square$

A.2.2 Proof of Proposition 1

Proposition 1. *For the dual problem presented in formulation (A.2), the following is a dual feasible solution, where I_i^0 is the initial inventory of product i and $\bar{p}_i^0$ is the initial nominal price setting for product i:*

$$\hat{\theta}_i = \bar{p}_i^0 \left(1 - \psi\left(\frac{I_i^T}{I_i^0}\right)\right) \quad \forall i$$

$$\begin{aligned}\hat{\lambda}^{k,t} &= \sum_{i=1}^{n} \left[\phi_i^{k,t}(\boldsymbol{p}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot \psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\ &+ \sum_{i \in S_{k,t}} \bar{p}_i^t \cdot \left[\phi_{S_{k,t}}^{k,t}(\boldsymbol{p}_{S_{k,t}})\left(d_{S_{k,t}} + \psi\left(\frac{I_i^{k,t}}{I_i^0}\right) - 1\right) - \phi_{S_{k,t}}^{k,t}(\bar{\boldsymbol{p}}_{S_{k,t}}) \cdot \psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \quad \forall (k,t)\end{aligned}$$

Proof. Proof of Proposition 1 Given the formulation of $\{\text{Dual}\}_{\forall(k,t)}$ presented in equation (A.2), we want to show the following two conditions:

$$\begin{aligned}(1) \qquad \lambda^{k,t} &\geq \sum_{i=1}^{n} \left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})(\bar{p}_i^t - \theta_i)\right] \\ &+ \sum_{i \in S_{k,t}} \left[\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\left(\bar{p}_i^t \cdot d_{S_{k,t}} - \theta_i\right) - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\left(\bar{p}_i^t - \theta_i\right)\right] \quad \forall k, t, S_{k,t} \in \hat{S} \\ (2) \qquad \theta_i &\geq 0 \quad \forall i\end{aligned}$$

Let us first focus on the more challenging condition (1). We define a new term as follows:

$$\theta_i^t = \bar{p}_i^t \left(1 - \psi\left(\frac{I_i^T}{I_i^0}\right)\right) = \bar{p}_i^t - \bar{p}_i^t \cdot \psi\left(\frac{I_i^T}{I_i^0}\right) \quad \forall i, t \tag{A.4}$$

Note that this new term θ_i^t is based on the nominal price $\bar{p}_i^t$ for product i in period t. Thus, $\bar{p}_i^t \leq \bar{p}_i^0$, because all nominal prices follow a markdown trajectory over time. Therefore,

$$\theta_i^t \leq \theta_i^0 = \hat{\theta}_i \tag{A.5}$$

We can now show feasibility using this new terminology as follows:

$$
\begin{aligned}
\hat{\lambda}^{k,t} &= \sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&\quad + \sum_{i\in S_{k,t}}\bar{p}_i^t\cdot\left[\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\left(d_{S_{k,t}}+\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)-1\right)-\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&= \sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&\quad + \sum_{i\in S_{k,t}}\left[\left(\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\left(\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})-\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\right)\right)+\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\left(d_{S_{k,t}}-1\right)\right] \\
&\geq \sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{T}}{I_i^0}\right)\right] \\
&\quad + \sum_{i\in S_{k,t}}\left[\left(\bar{p}_i^t\cdot\psi\left(\frac{I_i^{T}}{I_i^0}\right)\left(\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})-\phi_{S^t}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\right)\right)+\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\left(d_{S_{k,t}}-1\right)\right] \\
&= \sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot(\bar{p}_i^t-\theta_i^t)\right]+\sum_{i\in S_{k,t}}\left[\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\cdot(\bar{p}_i^t\cdot d_{S_{k,t}}-\theta_i^t)-\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})(\bar{p}_i^t-\theta_i^t)\right] \\
&\geq \sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot(\bar{p}_i^t-\hat{\theta}_i)\right]+\sum_{i\in S_{k,t}}\left[\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\cdot(\bar{p}_i^t\cdot d_{S_{k,t}}-\hat{\theta}_i)-\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})(\bar{p}_i^t-\hat{\theta}_i)\right]
\end{aligned}
$$

We get the second inequality from the fact that $\psi(\cdot)$ is concave and increasing and $I_i^T \leq I_i^t \quad \forall t = 1,...,T$. Note that it is key here that the expression $\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}}) \geq 0$, ensuring that all the quantities by which $\psi(\cdot)$ is multiplied are positive. The third equality comes directly from the definition of θ_i^t in equation (A.4); finally this leads to the last inequality by applying equation (A.5). For condition (2) concerning $\hat{\theta}_i$, showing feasibility is trivial. As stated, $\psi(\cdot)$ is a concave monotone increasing function defined on $[0,1]$, so $\bar{p}_i^0 \cdot \psi(\cdot) \leq \bar{p}_i^0$. Thus $\hat{\theta}_i = \theta_i^0 \geq 0 \quad \forall i$ by definition. □

A.2.3 Proof of Lemma 4

Lemma 4. *For a fixed arrival sequence $\{k,t\}_{\forall(k,t),t=1}^T$, the expected value of the expectation of the duals variables $\hat{\lambda}^{k,t}$ is defined by the expression:*

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\hat{\lambda}^{k,t}\right] = \sum_{i=1}^{n}\left(\sum_{t=1}^{T}\sum_{l=I_i^t+1}^{I_i^{t-1}}\bar{p}_i^t\cdot\psi\left(\frac{l}{I_i^0}\right)\right)+\sum_{t=1}^{T}M^t\cdot\left(L^t-1\right) \tag{A.6}$$

Proof. For a fixed arrival sequence $\{k,t\}_{\forall(k,t),t=1}^{T}$, we want to find the expected value of the objective function of the dual of the Clairvoyant problem. We define a binary variable $Q_i^{k,t}=1$ if item i is purchased at time t, and is 0 otherwise. We first use this to consider the expectation $\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\hat{\lambda}^{k,t}\right]$ over consumer choices below:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\hat{\lambda}^{k,t}\right]= \\
=&\mathbb{E}\Bigg[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\Bigg(\sum_{i=1}^{n}\left[\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&\quad+\sum_{i\in S_{k,t}}\left[\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\left(\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})-\phi_{S^t}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\right)+\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\left(d_{S_{k,t}}-1\right)\right]\Bigg)\Bigg] \\
\leq&\mathbb{E}\Bigg[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\Bigg(\sum_{i=1}^{n}\left[Q_i^{k,t}\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&\quad+\sum_{i\in S_{k,t}}\left[\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\cdot\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\left(\frac{\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})}{\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})}-1\right)+\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\cdot\bar{p}_i^t\left(d_{S_{k,t}}-1\right)\right]\Bigg)\Bigg] \\
\leq&\mathbb{E}\Bigg[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\Bigg(\sum_{i=1}^{n}\left[(I_i^{k,t}-I_i^{k+1,t})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right] \\
&\quad+\sum_{i\in S_{k,t}}\left[\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\cdot\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\cdot L^t\left(1-d_{S_{k,t}}\right)+\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\cdot\bar{p}_i^t\left(d_{S_{k,t}}-1\right)\right]\Bigg)\Bigg] \\
=&\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\sum_{l=I_i^t+1}^{I_i^{t-1}}\bar{p}_i^t\cdot\psi\left(\frac{l}{I_i^0}\right)\right) \\
&\quad+\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\sum_{i\in S_{k,t}}\left[\bar{p}_i^t\cdot\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\left(1-d_{S_{k,t}}\right)\left(L^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)-1\right)\right]\right] \\
\leq&\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\sum_{l=I_i^t+1}^{I_i^{t-1}}\bar{p}_i^t\cdot\psi\left(\frac{l}{I_i^0}\right)\right)+\sum_{t=1}^{T}\hat{M}^t\cdot\left(L^t-1\right), \\
=&\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\sum_{l=I_i^t+1}^{I_i^{t-1}}\bar{p}_i^t\cdot\psi\left(\frac{l}{I_i^0}\right)\right)+\sum_{t=1}^{T}M^t\cdot\left(L^t-1\right)
\end{aligned}
\tag{A.7}
$$

Note that the first inequality comes from the fact that the discount $d_{S_{k,t}}=\frac{p_{S_{k,t}}}{\bar{p}_{S_{k,t}}}<1$ and $\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})\geq\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$. The second inequality comes from applying Lipschitz continuity to the expression $\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})-\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$, from which we derive the following value of L^t:

$$L^t = C^t \cdot \frac{1}{\epsilon^t} \sum_{i=1}^{n} \bar{p}_i^t \cdot (1-\gamma) \ \ \forall t = 1, ..., T,$$

where $\epsilon^t = \min_{S_{k,t} \in \hat{S}} \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$ for all $t = 1, ..., T$, and the maximum discount of any given bundle is lower bounded by a constant, resulting in $d_{S_{k,t}} \geq \gamma$. We derive this in detail in Proposition 2. In the third inequality we use the fact that $\psi(\cdot) \in [0,1]$ and finally that the discount ratio $d_{S_{k,t}}$ will be at most some $\delta < 1$ within any given period t, providing us with the following expression:

$$M^t = \max_{S_{k,t} \in \hat{S}} \quad \sum_{k=1}^{K^t} \sum_{i \in S_{k,t}} \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}}) \cdot \bar{p}_i^t \cdot (1-\delta)$$

□

A.2.4 Proof of Proposition 2

Proposition 2. *The Lipschitz continuity factor L^t representing the quantity $\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$ is given by,*

$$L^t = L^t = C^t \cdot \frac{1}{\epsilon^t} \sum_{i=1}^{n} \bar{p}_i^t \cdot (1-\gamma), \ \ \textit{where } \epsilon^t = \min_{S_{k,t} \in \hat{S}} \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}}) \ \ \forall t$$

where the minimum in ϵ is taken over all arrivals $k = 1, ..., K^t$ in a given period t, γ is the maximum possible bundle discount available in any period, and C^t is the Lipschitz continuity constant that depends on the nominal price settings $\bar{p}_i^t$.

Proof. Proof of Proposition 2 We are interested in approximating the quantity,

$$\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}}) - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$$

used in the calculation of equation (A.7) of the expectation of $\hat{\lambda}^{k,t}$. We apply the definition of Lipschitz continuity to the demand function $\phi(\cdot)$ to derive L^t as follows:

$$
\begin{aligned}
|\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}}) - \phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})| &\leq C^t \cdot |\bar{p}_{S_{k,t}} - p_{S_{k,t}}|, \text{ by Lipschitz continuity of } \phi(\cdot), \\
\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})\left|\frac{\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}})}{\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})} - 1\right| &\leq C^t \cdot |\bar{p}_{S_{k,t}} - p_{S_{k,t}}| \\
\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})\left|\frac{\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}})}{\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})} - 1\right| &\leq C^t \cdot \bar{p}_{S_{k,t}} \cdot \left|1 - \frac{p_{S_{k,t}}}{\bar{p}_{S_{k,t}}}\right| \\
\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})\left|\frac{\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}})}{\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})} - 1\right| &\leq C^t \cdot \sum_{i=1}^{n} \bar{p}^t_i \cdot |1 - d_{S_{k,t}}|, \text{ by definition of } \bar{p}_{S_{k,t}} \text{ and } d_{S_{k,t}}, \\
\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})\left|\frac{\phi^{k,t}_{S_{k,t}}(\mathbf{p}_{S_{k,t}})}{\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})} - 1\right| &\leq C^t \cdot \phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}}) \cdot \frac{1}{\epsilon^t} \cdot \sum_{i=1}^{n} \bar{p}^t_i \cdot (1 - d_{S_{k,t}}), \epsilon^t = \min_{S_{k,t} \in \hat{S}} \phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}}) \ \forall t
\end{aligned}
$$

Note that ϵ^t is defined differently in each period t depending on both the settings of the nominal prices $\bar{p}^t_i$ and the bundles $S_{k,t}$ (and corresponding discounts $d_{S_{k,t}}$) offered in that period. In the calculation of the expectation of $\hat{\lambda}^{k,t}$ in equation (A.7) we use an additional $\phi^{k,t}_{S_{k,t}}(\bar{\mathbf{p}}_{S_{k,t}})$ in order to construct our period dependent profit loss term M^t and thus introduce ϵ^t to maintain the final inequality with this extra term. Furthermore, it is a realistic business assumption that the maximum discount given on any product i in any period t is bounded, giving us $d_{S_{k,t}} \geq \gamma$. Thus we define our desired Lipschitz continuity factor by,

$$L^t = C^t \cdot \frac{1}{\epsilon^t} \sum_{i=1}^{n} \bar{p}^t_i \cdot (1 - \gamma) \ \ \forall t = 1, ..., T.$$

Note that the above definition can be further simplified to a single constant with a maximization over all time periods t. □

A.2.5 Proof of Lemma 5

Lemma 5. *Given a fixed consumer arrival sequence* $\{k,t\}^T_{\forall(k,t),t=1}$*, the expected value of the objective function of* $\{MultAlg\}_{\forall(k,t)}$ *is given by the expression:*

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\sum_{k=1}^{K^t}\phi^{k,t}_i(\boldsymbol{p}_{S_{k,t}}) \cdot \bar{p}^t_i - \sum_{l=I^t_i+1}^{I^{t-1}_i} \bar{p}^t_i \cdot \psi\left(\frac{l}{I^0_i}\right)\right) = \sum_{i=1}^{n}\sum_{t=1}^{T} \bar{p}^t_i \cdot \left(\sum_{k=1}^{K^t}\phi^{k,t}_i(\boldsymbol{p}_{S_{k,t}}) - \sum_{l=I^t_i+1}^{I^{t-1}_i} \psi\left(\frac{l}{I^0_i}\right)\right)$$

Proof. Given a fixed consumer arrival sequence $\{k,t\}_{\forall(k,t),t=1}^{T}$, we can derive the objective value of the multiplicative approximation algorithm as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\{\text{MultAlg}\}_{\forall(k,t)}\right] = \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\right) + \left(\phi_{S_{k,t}}^{k,t}(\mathbf{p}_{S_{k,t}})p_{S_{k,t}} - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\bar{p}_{S_{k,t}}\right)\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\right) - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\bar{p}_{S_{k,t}}\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\left(\sum_{i\in S_{k,t}}\bar{p}_i^t\right)\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{i\in S_{k,t}}\phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{i\in S_{k,t}}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\min_{i\in S_{k,t}}\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{K^t}\left(\sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{i=1}^{n}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t\cdot\psi\left(\frac{I_i^{k,t}}{I_i^0}\right)\right)\right] \\
&= \sum_{i=1}^{n}\sum_{t=1}^{T}\left(\sum_{k=1}^{K^t}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}})\cdot\bar{p}_i^t - \sum_{l=I_i^t+1}^{I_i^{t-1}}\bar{p}_i^t\cdot\psi\left(\frac{l}{I_i^0}\right)\right) \\
&= \sum_{i=1}^{n}\sum_{t=1}^{T}\bar{p}_i^t\cdot\left(\sum_{k=1}^{K^t}\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) - \sum_{l=I_i^t+1}^{I_i^{t-1}}\psi\left(\frac{l}{I_i^0}\right)\right)
\end{aligned}
$$

The second equality comes directly from the definition of a nominal bundle price $\bar{p}_{S_{k,t}}$. The third inequality comes from the fact that by definition, $\phi_i^{k,t}(\mathbf{p}_{S_{k,t}}) \geq \phi_{S_{k,t}}^{k,t}(\bar{\mathbf{p}}_{S_{k,t}})$ as it encompasses the bundle purchase probability at nominal price. In the fourth inequality we remove the minimum, giving us a lower bounding quantity since this term is negative. □

A.2.6 Proof of Lemma 6

Lemma 6. *Given a fixed set of constants a_i $\forall i = 1, ..., n$ and corresponding variables x_i and y_i, the following property holds:*

$$\frac{\sum_{i=1}^{n} a_i \cdot x_i}{\sum_{i=1}^{n} a_i \cdot y_i} \geq \underset{i}{min} \frac{x_i}{y_i}$$

Proof. Proof of Lemma 6 We want to show that we can lower bound the ratio of two sums with the same weights a_i and different variable values x_i and y_i using a minimum over the ratios of all the variable pairs x_i, y_i. Let us first define the following term,

$$\hat{\alpha} = \min_i \frac{x_i}{y_i}$$

By the definition of α we know that $x_i \geq \hat{\alpha} \cdot y_i$ $\forall i$. Therefore, we get the following result as desired,

$$\frac{\sum_{i=1}^{n} a_i \cdot x_i}{\sum_{i=1}^{n} a_i \cdot y_i} \geq \frac{\sum_{i=1}^{n} a_i \cdot (\hat{\alpha} \cdot y_i)}{\sum_{i=1}^{n} a_i \cdot y_i} = \hat{\alpha} = \min_i \frac{x_i}{y_i}.$$

□

A.2.7 Proof of Lemma 7

Lemma 7. *Given a monotone increasing function $\psi(\cdot)$ and an increasing set of constants $x = x_0, ..., x_N$, the following condition holds,*

$$\sum_{x=x_0}^{x_N - 1} \psi(x) \leq \int_{x=x_0}^{x_N} \psi(y) dy$$

Proof. Proof of Lemma 7 By definition of the values of x, we know that $x_0 \leq x_1 \leq ... \leq x_N$. Since $\psi(\cdot)$ is a monotone increasing function we have that $\psi(x) \geq \psi(x_i)$ $\forall x \in [x_i, x_i + 1]$. If we integrate this expression over $[x_i, x_i + 1]$ for a fixed value of x_i we get,

$$\int_{x_i}^{x_i+1} \psi\left(x_i\right) = \psi(x_i) \leq \int_{x_i}^{x_i+1} \psi(x)dx$$

Rewriting the left hand expression through a summation we precisely get the desired result,

$$\sum_{x=x_0}^{x_N-1} \psi\left(x\right) \leq \int_{x=x_0}^{x_N} \psi(y)dy.$$

$\square$

Appendix B

Appendix of Chapter 3

B.1 Supplemental Figures for Airline Case Study

<u>Distribution of Consumers Across Tier Levels</u>

In Figure B-1 below is a summary of the distribution of consumers in the airline case study across the four possible tier levels, as provided by the airline. Note that very few consumers attain the higher levels, which require a great deal of very frequent travel.

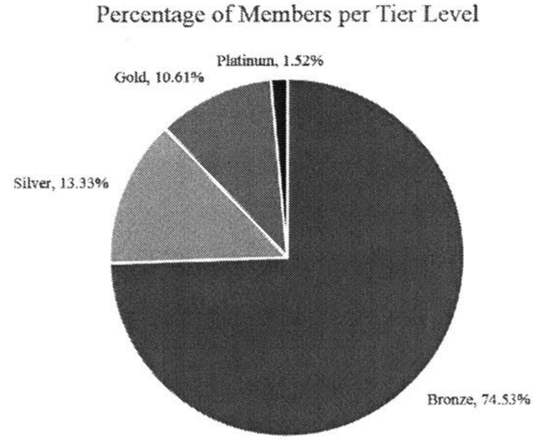

Figure B-1: This chart summarizes the distribution of consumers across the various airline tier levels.

Description of Persona Clusters

We used k-means clustering to develop the persona groups in the airline case study. We ultimately constructed 7 distinct clusters based on both personalized and itinerary context related features. The below descriptions in Figure B-2 explain how some of these features influenced the various cluster compositions.

Consumer Cluster	Description
Premium Business Traveler	**High tier level, short time to departure, high fare, traveling on week days, single passenger**
Economy Business Traveler	High tier level, short time until departure, lower fare, traveling on week days, single passenger
Families	**At least 3 passengers in booking, longer time before departure, most previous trips in econ.**
Leisure Couples	Tier level Skywards or Silver, 2 passengers in booking, previous trips in econ. class
High-End Couples	**Tier level Gold or Platinum, 2 passengers in booking, at least 1 trip in prem. class**
Leisure Individuals	Tier level Skywards or Silver, single passenger, longer time before departure
High-End Individuals	**Gold or Platinum Tier, single passenger, min. 1 trip in prem. class, longer time before departure**

Figure B-2: This table details some of the descriptive features that define each of the airline persona clusters.

Price Elasticities by Persona Type

We extracted the following price elasticity relationships between each (persona type, ancillary service) pair as a result of fitting our pairwise models of propensity-to-buy. Note that Figure B-3 is fairly intuitive: lower-elasticity types are more frequent and higher tier consumers such as business travelers in premium, as well as higher-end leisure single travelers and couples. Similarly, family groups have typically higher elasticities across all products unanimously.

Price Elasticities	Wifi access	Premium on board entertainment	Priority security	Priority boarding	Priority baggage handling	Comfort seats /upgrade seats	Checked /Excess baggage	VIP lounge access	Gourmet meals	Bonus miles 2000	Bonus miles 4000
BusTravPREM	-1.11	-1.60	-	-	-	-	-	-	-	-1.07	-2.03
BusTravECON	-1.58	-1.86	-1.22	-1.51	-2.63	-1.82	-2.43	-1.87	-1.53	-1.10	-2.43
Family	-2.26	-2.32	-3.15	-2.66	-3.83	-3.04	-2.60	-3.55	-2.78	-1.89	-3.39
LastMinGroup	-2.56	-3.73	-3.21	-3.21	-3.14	-3.48	-2.06	-4.00	-3.99	-2.06	-3.41
CoupleNormal	-1.92	-3.00	-2.96	-2.68	-2.84	-2.93	-2.07	-2.47	-2.83	-1.25	-2.76
CoupleHighEnd	-1.20	-1.86	-1.34	-1.07	-2.17	-1.85	-1.41	-1.85	-1.95	-1.93	-2.98
LeisTravNormal	-2.33	-2.17	-2.35	-2.15	-2.96	-1.59	-2.56	-2.22	-2.65	-1.19	-2.79
LeisTravHighEnd	-0.68	-1.18	-0.79	-1.29	-1.83	-1.71	-1.47	-1.10	-1.20	-1.03	-1.69

Figure B-3: This table details the pairwise price elasticities between each consumer persona cluster and ancillary service.

Expected Gains in Revenue and Sales for Various Levels of Lost Sales

By introducing parameter α, we are able to capture the percentage of consumers who are entirely unaware of the existence of ancillary services, and simply exit the ticket transaction without considering them at all. By considering varying levels of α we can objectively weigh the effect of product recommendation of ancillary goods in this setting. The below tables are the expected lifts in revenue, Figure B-4, and expected lifts in sales volume, Figure B-5, over the baseline method which offers all of the ancillary products at their full prices. Notice that as we reach higher levels of lost sales, the method proposed by our personalized pricing and recommendation model shows significantly more improvement over this baseline.

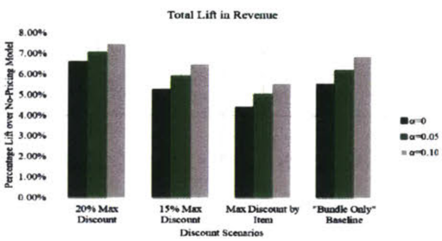

Figure B-4: This graph shows the expected improvement in revenue by implementing our model over the no-pricing benchmark across various levels of lost sales.

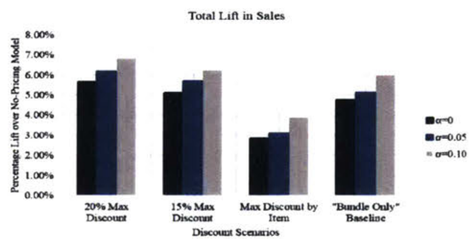

Figure B-5: This graph shows the expected improvement in sales volume by implementing our model over the no-pricing benchmark across various levels of lost sales.

Empirical Model Performance with Simulated Inventory

While inventory constraints were not inherently present in this data set, we narrowed our scope to consider ancillary services with reasonably limited stock, such as in-flight wi-fi, excess checked baggage, VIP lounge access, and seating upgrades. Across all of these products, we measured the average expected revenue achieved by each of our various approximation heuristics and benchmark methods relative to the full-knowledge Clairvoyant model and summarized the results in Figure B-6 below.

Note that using a dynamic pricing approach over a rolling LP model with static price settings improves expected revenue by 5.4%. Incorporating inventory considerations as opposed to implementing a myopic profit-maximizing approach improves the expected revenue by 3.7%. Furthermore, our best-performing approximation method,

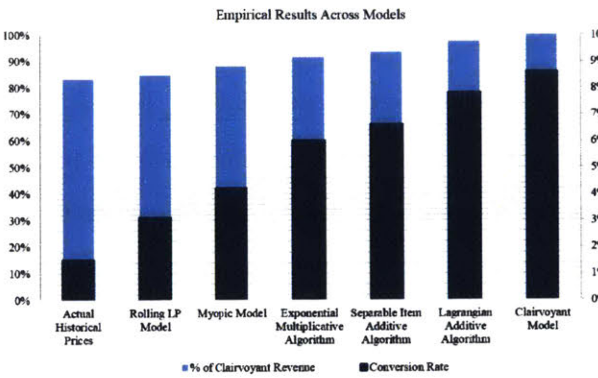

Figure B-6: This table summarizes the average empirical performance of the various approximation methods and benchmarks with simulated inventory, relative to the full-knowledge strategy.

the ALA, achieves on average up to 98.4% of the Clairvoyant revenue across all products, personas and inventory scenarios. When compared to the currently implemented baseline approach of offering all ancillary products at full prices, we obtain up to a 12.6% expected gain in revenue overall.

B.2 Supplemental Figures for Retail Case Study

Distribution of Expenditure and Market Basket Size per Transaction

In order to develop our personalization metrics we analyzed population-level behavior of the cumulative expenditure and number of consumer visits in the entire two year selling period. To initially understand the consumer spending behavior in this data set, we analyzed the distribution of market basket sizes, as shown in Figure B-7, and corresponding transaction-level expenditure, as shown in Figure B-8.

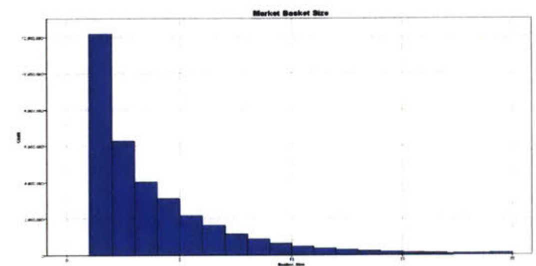

Figure B-7: This plot shows distribution of market basket sizes per transaction across the entire data set.

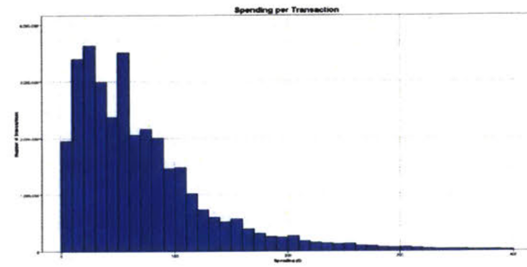

Figure B-8: This plot shows distribution of expenditure per transaction across the entire data set.

Distribution of Cumulative Expenditure Across Entire Population

Having noted the general low transaction-level spending and market basket composition size, we decided to focus on cumulative metrics in order to construct personalization in this data set. The plot below in Figure B-9 summarizes the cumulative expenditure for all the consumers in the retail data set.

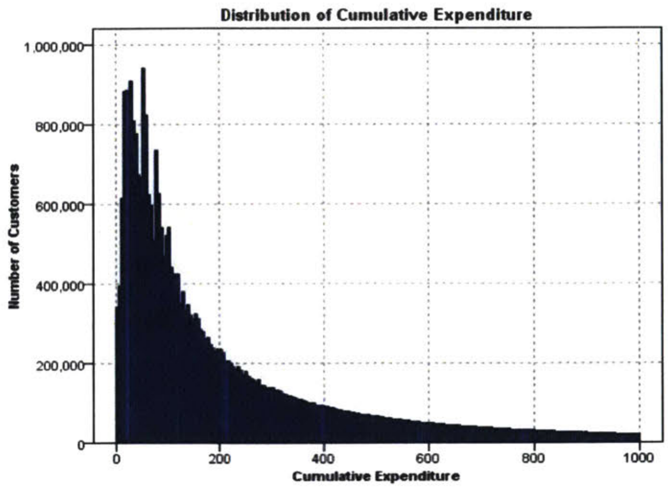

Figure B-9: This histogram summarizes cumulative consumer expenditure across all of the consumer population in the entire two year selling period.

Cumulative Number of Consumer Visits by Loyalty Group

As the retail data was innately missing individualized features outside of customer IDs and transaction IDs, we developed our own metrics of personalization in order to fit the desired models of propensity-to-buy. We focused on cumulative expenditure and cumulative number of shopping visits to distinguish consumers into various loyalty groups.

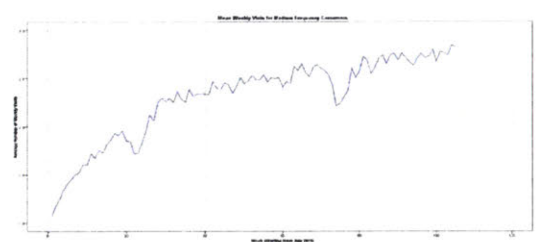

Figure B-10: This plot shows the cumulative expenditure over time of the medium frequency loyalty group.

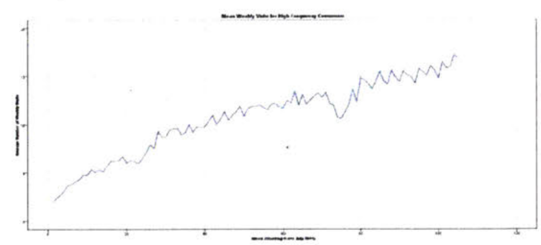

Figure B-11: This plot shows the cumulative expenditure over time of the high frequency loyalty group.

In Figures B-10 and B-11 we show the time-dependent behavior of cumulative consumer expenditure across the two more frequency loyalty groups. Note that high frequency consumers spend above the mean amount as determined by the previous histogram in Figure B-9.

Sold Quantities and Charged Amounts for Products in Demand Group 4

We chose to analyze products in the seasonal home decor department due to their inherent full pricing strategies. These products, as demonstrated in Figures B-12 and B-13 generate the most revenue when sold as bundles, but also the least sales. From a business perspective, the goal of our bundle recommendation system is precisely to convert the lower sales in Figure B-12 in order to improve revenues of the higher charged quantities in Figure B-13.

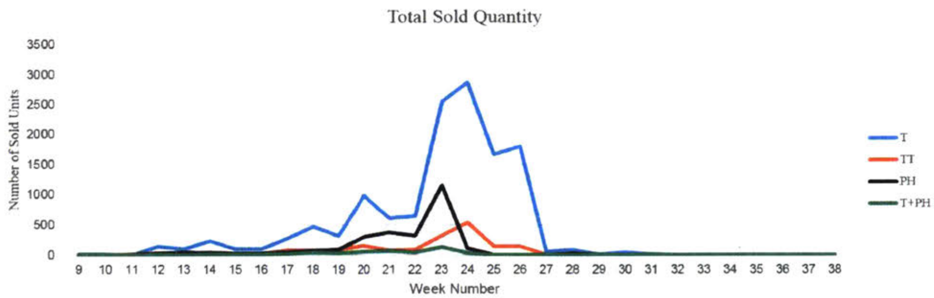

Figure B-12: This plot shows the sold quantity of product combinations from Demand Group 4.

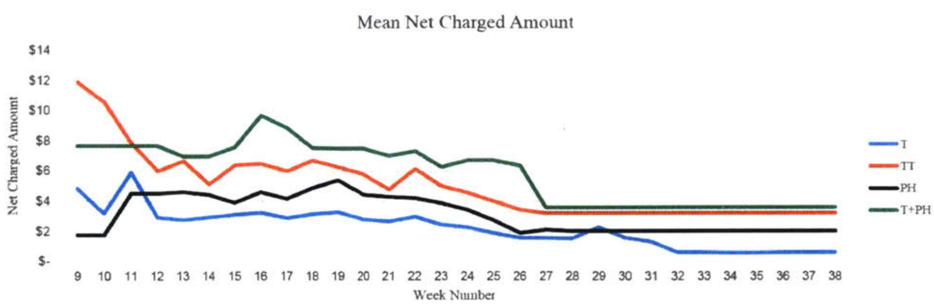

Figure B-13: This plot shows the charged amount for product combinations from Demand Group 4.

Appendix C

Appendix of Chapter 4

C.1 Positioning Relative to Existing Literature

Relevant Literature on Contextual Multi-Armed Bandit Theory

Paper	Reward Structure	Number of Pulls	Reward Dependence	Problem Constraints	Similarity	Difference	Regret Characterizations	Description
Auer'02	Linear	1	None	None	Contextual (changing contexts)	Non-linear	Dependence on K	First bandit paper
Yadkori'11	Linear	1	None	None	Contextual (changing contexts)	Conditional R-sub Gaussian	No dependence on K	Improved bounds using self normalized bounds
Badanidiyuru'13	None	1	None	Knapsack constraint	Knapsack constraint	No contexts	Dependence on K	First paper on bandits with knapsack
Badanidiyuru'14	None	1	None	Knapsack constraints	Arbitrary contexts allowed; reward can be non-linear.	Regret on policy space, Distribution of reward, context		First paper for contextual bandits with knapsacks
Agarwal'14	Non-linear	1	None	Convex constraint	Convex constraint	Contexts are known before starting the algorithm	Independence from K	First paper with concave rewards and knapsack constraints
Agarwal'16	Non-linear	1	None	Knapsack constraints	Extension to concave objective, constraints	Reward and consumption IID drawn from a joint distribution. Policy space to bound regret		Main focus on computational efficiency
Agarwal'16	Linear	1	None	Knapsack constraints	Linear reward which is generated from IID sample	Reward is linear	Independence from K	Extension of BâĂŽ14 but only for linear case. No policy space needed
Qin'14	Non-linear	Subset	None	None	Multiple pulls	Each arm score is created which can be observed.	Dependence on K	
This Work	Non-linear	Can be either	Can be either	Knapsack constraints			Regret can	

Table C.1: This table summarizes the relevant literature on contextual multi-armed bandit theory and positions our work relative to the existing models in this field.