Three Essays in Technology and Revenue Management

by

Sam Kirshner

A thesis submitted to the

School of Business

in conformity with the requirements for

the degree of Doctorate of Philosophy

Queen's University

Kingston, Ontario, Canada

December 2014

Abstract

In this dissertation, I apply optimization methods and game theory to address three problems in technology and revenue management. In the first essay, I analyze how brand commitment and product failure impact a firm's upgrade strategy in the presence of a stochastically evolving technological frontier. The essay explores the optimal timing of upgrades across a variety of market parameters and establishes the market conditions in which firms should invest in brand commitment to lengthen the product upgrade cycle. The model also demonstrates that firms with high brand commitment must balance the benefits of pent-up demand with potential loss due to product failure. The second and third essay focus on the allocation of resources and products in the presence of demand uncertainty and consumer behavior, respectively. In the second essay, I develop a methodology to approximate the value of capacity in the network airline revenue management problem. The value of capacity is used to control the sale of products to consumers requesting products over a finite time horizon. The advantage of this methodology is the scalability, which we demonstrate by solving for capacity values on an industrial sized network. In the third essay, I study a consumer-to-consumer exchange market. I prove that there exists market conditions where the equilibrium prices allow a unique opitmal allocation of products amongst participants.

Acknowledgments

I cannot begin to express my appreciation and gratitude towards my advisors Yuri Levin and Mikhail Nediak. Both of you are exceptional mentors. I am very lucky to have had you guide me through my doctorate. I would like to specifically thank Yuri for his wisdom and for always looking out for my future as an academic. Your advice and knowledge has been invaluable. I would like to specifically thank Mikhail for his limitless patience and for always making himself available to help me when I needed it. I would also like to thank my committee members. Tanya, thank you for your encouragement and for allowing my defence to be an enjoyable moment. Arcan, it has been a great pleasure to teach with you. I appreciate the guidance that you have imparted to me over coffee and meals. I also would like to thank the entire PhD office for their assistance and for putting up with my shenanigans.

My dissertation would not have been completed without the help and support of my family and several of my closest friends. Mother, your love and constant encouragement kept me motivated throughout my degree. Father, I will never forget the number of nights that you spent helping me to clarify and articulate ideas. Alicia and Archer, together, you provided endless comic relief. Finally, Shira, Julia, Aaron, and Sarah, I thank you both for acting as readers and editors for my thesis and for being my closest friends over the last five years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The research in this dissertation addresses three important problems in technology and revenue management. In the first study, I analyze the impact of branding on the timing of a firm's optimal product upgrade strategy in industries where market demand for a durable product is influenced by the stochastic improvements of technology. Formulating the timing of product upgrade releases as a Markov Decision Process, I prove that the optimal upgrade strategy can be characterized by threshold policy either based on pent-up demand for a future product release or the technological lag of the product. The research demonstrates that brand commitment offers substantial profit improvements over the lifecycle of the product in industries characterized by fast pace of technology, high upgrade costs, or small incremental market growth. The research also explores the market conditions in which investing in raising brand commitment is profitable. Unexpectedly, investments in raising brand commitment are most valuable among firms with an existing high degree of brand commitment.

The model is generalized to account for the possibility that a new product may fail to meet consumer expectations due to unforeseen manufacturing, design, or supply chain issues. When considering potential for product failure, a threshold policy on pent-up demand remains optimal, while a technology lag-based threshold policy is generally sub-optimal. Thus, for markets where branding is an important determinant of consumer purchasing behavior, an upgrade policy constructed on threshold values of pent-up demand is a more robust approach to managing product upgrades. Analysis of the model demonstrates that product failure has a marginal impact on the upgrade strategy of firms with low degrees of brand commitment. Conversely, product failure has a significant impact on upgrade frequency for firms with high degrees of brand commitment, since such firms have to balance the advantages of pent-up demand with potential loss of sales. Research implications are discussed in the context of the premium smartphone market.

In the second study, I analyze a network capacity control problem, in which consumers purchase products based on various combinations of resources. This creates a stochastic demand for a fixed set of perishable resources over a large network. My research addresses the intractable nature of the problem by generating time-dependent bid prices to approximate the value of each resource. The methodology is derived directly from the network revenue management optimal control problem by introducing two approximations. The first approximation uses an affine functional form to represent the value function and time-dependent bid prices. The second approximation reduces the number of variables and constraints in the problem through an implicit treatment of the reservation control policy. The most significant contribution of this methodology is the use of splines and second-order cone programming to reduce the infinite collection of constraints to a finite set, in order to produce efficient time-dependent bid prices. The spline representation of bid prices allows the number of variables to depend solely on the number of resources, as opposed to the size of the planning horizon. The resulting model is an approximate second-order cone program (ASOCP).

The ASOCP is tested against the state of the art discrete time-dependent bid price policy over randomly generated networks. The run-time of discrete time dynamic bid prices is shown to increase dramatically as the capacity levels across the network increase in variability. On the other hand, the ASOCP algorithm timing is robust to variation in capacity, which implies that the methodology is applicable to situations with frequent updating. The ASOCP is subsequently tested against re-optimized deterministic bid price policies on randomly generated and industrial airline networks. The ASOCP generates higher revenue compared to the re-optimized static bid price, thus demonstrating the practicality and scalability of the methodology.

In the third study, I develop an approach to modeling a general exchange market with an arbitrary number of participants and products. The model features quantity competition among participants, who are differentiated by product preference and price sensitivities to available products. In order to maximize surplus, participants in the exchange market decide which products to buy and sell. Using linear complementarity theory, the research shows that the solution for an arbitrary number of participants trading an arbitrary number of products is always feasible. I prove that there exists special cases where the quantity of goods traded is unique. The generality of the setup may allow future research to incorporate the model into a larger supply chain framework with secondary markets applications.

Chapter 2

Timing Product Upgrades with Stochastic Technological Advancements, Brand Commitment, and Pent-Up Demand

2.1. Introduction

When firms time releases of next generation products, they need to account for the effects of uncertainty in advancements of component technology on the demand. Shorter intervals between upgrades enable firms to capture incremental demand at a higher cost. Longer intervals risk missed sales if customers opt to purchase products with improved technology from competitors. A variety of models have been proposed to optimize the timing of upgrades in this setting ([9]; [36]; [43]; [15]; [45]). We contribute to this literature by incorporating branding and product failure into a decision model for timing upgrades. We demonstrate that consumers' willingness to wait given a technology lag and willingness to purchase given a product failure drives the optimal upgrade strategy. Since these behaviors have opposite effects on upgrade frequencies, understanding consumer dynamics is essential for successfully executing a product upgrade strategy.

In mass consumer markets for technological products, customers have a predisposition for purchasing a particular brand. This preference may result from a firm's reputation for quality, advertising, social influence, research or a prior experience with other products offered by the firm. The strength of this preference may differ widely within a firm's customer base and across different firms within a common market. Customers may have a small positive bias towards one brand that may dissipate if a competing product has the slightest edge in terms of performance. Thus, if the consumer does not purchase a product due to technology performance lag, she will prefer an outside option over purchasing from the firm. On the other hand, consumers may have a deeply held commitment to purchasing a firm's product. In industries where the evolution of technology occurs at a rapid pace, consumers anticipate that preferred firms will at some point release an upgraded product. If a firm's product performance lags leading edge products, committed customers will wait for an upcoming product release from their preferred brand. This is particularly true in situations where consumers strongly identify with the brand's image ([18]; [23]).

In this paper, we characterize the strength of a firm's brand commitment as the willingness of consumers to wait for future product releases rather than leave the firm's market. High brand commitment can significantly mitigate the rate at which firm's sales erode due to a gap in technology between their incumbent product and more advanced competitive offerings, since consumers wait for the release of the upgraded product. The number of consumers who have delayed purchasing in anticipation of a future product release is the firm's pent-up demand for the next generation product. Over time consumers who had postponed their purchases may lose patience, in which case they will opt to purchase from a competitor. Thus, pent-up demand for the future product release is a dynamic process that initially increases with the gap in technology. However, pent-up demand contracts if the time between upgrades is sufficiently long such that the number of committed customers exiting the market exceeds those entering.

Smartphones provide one of the most compelling examples of how the dynamics of upgrades for a mass technology product are influenced by branding. This industry has been singled out as having the most rapid pace of product upgrades due to the high rate of technological innovation. Consequently, to remain competitive, firms must frequently upgrade their products to keep pace with advancements of components suppliers of memory, CPU, modems, displays, camera modules, GPS, batteries and sensors (see Online Appendix A.2). Nevertheless, smartphone vendors differ widely in their pace of model releases and the extent to which an incumbent product can remain competitive despite the release of numerous models from competitors equipped with leading-edge technology.

International smartphone sales are currently dominated by Apple and Samsung, which have divergent strategies in terms of their upgrade frequency. Samsung releases multiple upgrades of their premium smartphones throughout the year in order to capture incremental demand associated with improvements in technology ([71]). Conversely, Apple adheres to a schedule of releasing an upgrade of their iPhone approximately every 12 months. Although Apple generally experiences a decline in market share towards the latter stages of their product release cycle, longer duration between model releases allow the company to accumulate pent-up demand from committed consumers ([40]). For firms with less brand commitment, falling behind in technology leads to lost sales that are not recaptured with the release of a new model ([31]; [29]). While the strong commitment of Apple's iPhone customer base can be partly attributed to network effects and switching costs, empirical evidence suggests that branding plays the most important role in attracting Apple's committed customer base. For example, sales data shows that the aggregation of pent-up demand for new product launches occurs even in countries where the iPhone becomes available for the first time and where other Apple products have few inroads ([39]). Although pent-up demand significantly impacts the profitability of upgrades, research on upgrade strategies has yet to analytically establish the connection between brand commitment and technological advancement. There is a similar gap in knowledge among business analysts. For example, despite the proven success of Apple's product strategy in the past, many analysts were skeptical about whether the company could afford to wait a year between the release of the iPhone 5 in September 2012 and the iPhone 5s in the Fall of 2013. During this period, Apple experienced declining sales due to the vast number of premium smartphones with state of the art technology launched by competitors such as Samsung, LG and HTC (see Online Appendix A.2). Nevertheless, when Apple released the iPhone 5s in September of 2013, it sold close to 9 million phones globally in the opening weekend. Apple was able to regain their cumulative market share of approximately 45% of US smartphones in 2013 despite the drop in sales over several months prior to the 5s product launch ([13]; [28]).

High brand commitment enables the firm to execute an upgrade strategy that reduces launch costs without an excessive loss of sales due to the aggregation of pent-up demand. However, if there is an unanticipated failed product launch due to a product defect, firms with high levels of pent-up demand are at greater risk of suffering a disproportionate loss in their profit. This problem is prevalent when stochastic improvements in the technology of several components complicate the design and production process of state of the art products. The probability of a failed product launch also tends to increase in relation to the size of the upgrade that the firm is performing in terms of technological advancements ([10]; [54]).

The inherent risk associated with the aggregation of pent-up demand in the smartphone industry is vividly illustrated by the downfall of Blackberry and Nokia. Both firms went from market leaders to laggards in a short timeframe. Blackberry, in particular, had one of the highest measures of brand commitment in the industry which enabled it to aggregate significant demand for their future release of a smartphone with a functional web browser. The product was delayed for over two years and by the time it came out most of their customers had defected. Solidifying Blackberry's downfall was their decision to release a touch screen model (the Blackberry Z10) despite the fact that Blackberry's remaining pent-up demand was comprised of customers who wanted a phone with the company's proprietary QWERTY keyboard ([57]). Similarly, Nokia's overestimation of brand value to consumers and succession of product failures are credited as major reasons for the firm's rapid collapse ([61]).

The objective of this paper is to analyze the influence of brand commitment on the frequency and profitability of product upgrades. We establish the existence of a state-dependent threshold based on the level of pent-up demand, which determines whether upgrading is optimal in the current period. We prove that this threshold is monotonically nonincreasing in the lag in technology between the incumbent product and the leading-edge. Thus, we can construct a threshold policy in terms of the technology lag similar to the policies described in the literature ([9]; [36]; [43]; [15]).

We extend the baseline model to examine how the upgrade decision is affected by the risk of launching a failed product. Product failures result in a portion of customers, comprised of pent-up demand and new arrivals, who do not buy the product due to the product's suboptimal performance. We demonstrate that the optimal upgrade strategy maintains a threshold policy on pent-up demand, but a threshold policy based on the lag in technology is not guaranteed to be optimal. For markets where branding is an important determinant of consumer purchasing behavior, this suggests that an upgrade policy constructed on critical values of pent-up demand is a more robust approach to managing product upgrades compared to a policy based solely on the gap in technology.

Extensive numerical experiments demonstrate how this model performs under a variety of industry and firm-specific conditions related to the pace of technology, market growth, and cost of upgrades. The results highlight that brand commitment offers substantial profit improvements in industries characterized by either a fast pace of technology, high upgrade costs, or low incremental market growth. We also use the model to assess how an investment to increase brand commitment by a fixed amount will impact a firm's profit. Unexpectedly, when a firm can make a set investment to raise their level of brand commitment by a fixed amount, these investments are most valuable if the degree of brand commitment is already high. We show that the risk of failed product launches has a greater impact on the expected profit of firms with higher levels of brand commitment. For such firms, the risk of failures increases the frequency of upgrades, since the firm has to balance the advantages of pent-up demand with the potential sales loss due to failed products.

The remainder of this paper is organized as follows. §2 provides an overview of how our research contributes to the relevant literature. §3 describes the modeling assumptions and develops the firm's decision model. §4 discusses the optimality of the threshold policies based on pent-up demand and technology. §5 provides a detailed analysis of how the model performs under various industry and firm-specific conditions, highlighting managerial insights into the product upgrade problem. §6 extends the model to account for the possibility of a failed product launch. Finally, §7 provides a discussion on the main findings of the paper and directions for future research. All mathematical proofs can be found in Online Appendix A.1.

2.2. Related Literature

There is a vast body of literature on defining and measuring the relationship between branding and customer purchasing intentions. A common perception in technology industries is that the purchase decisions of customers are more influenced by the relative value proposition of competing products rather than factors related to branding. However, research shows that brand preference has a very powerful and enduring influence on consumers both prior to and after they purchase a technology product. Prior to the purchase, consumers place a higher premium on positive information about the product of their preferred brand while minimizing the importance of negative information ([53]). Likewise, after the purchase, consumers are much more likely to be satisfied with their choice and less likely to suffer buyer's remorse if they purchased a product from their preferred brand ([17]).

Brand preference is reinforced by factors related to brand identity, network effects, or switching costs that deepen consumer commitment to purchasing products from their preferred brand ([67]). With respect to consumer electronics, empirical studies and consumer surveys show that brand identification and switching costs become increasingly important determinants of purchase intent as a product category matures ([32]). Cross-cultural studies have shown that this is particularly true in the case of smartphones even when competing brands release new models with novel or significantly enhanced technology ([39]).

The literature on brand commitment's role in moderating purchase intent given a product failure is fairly divided. [4] shows that consumers with brand commitment have a bias which cause them to downplay negative product reviews. Whereas the experiments conducted by [56] show that a product failure will have greater impact on the probability of purchase for firms with high brand commitment. In particular, high brand commitment leads consumer to develop stronger attitudinal ties to the firm and to have greater expectations. We add to the literature by considering the role of commitment leading to pent-up demand within the context of product upgrades.

To date there are no models on the timing of product upgrades that consider factors related to branding or the aggregation of pent-up demand in the decision process. Most models focus on a firm's trade-off between performance and time to market. This trade-off was first explored by [19] for a fixed sales window based on market size potential, competitive offerings, and the pace of product improvements. [50] extend this research for a firm that introduces multiple generations of a product over a finite selling horizon in order to maintain or improve market share. [59] consider the problem in terms of an industry's clockspeed based on a firm's input costs related to product development, production, and inventory. In the aforementioned models, the advancement of technology is deterministic and market demand is independent of the level of technology. Thus, these models are not necessarily applicable to the upgrade problem for firms in industries where market potential increases due to stochastic improvements in technology driven by component suppliers.

[9], [36], [43], [15], and [45] all study the timing of upgrades with exogenous stochastic technology improvements. [9] study the timing of technology adoption over an infinite time horizon. Given a fixed cost of technology adoption, they demonstrate that a firm should adopt the state of the art technology once the lag in technology exceeds a threshold value. [15] extend [9] to investigate the adoption of two complementary technologies over time. Their model analyzes the interaction between co-dependent technologies, and focuses on situations where an increase in the lag of one technology leads to an upgrade in the dependent technology. In both models, profit is a linear function of the technology levels, implying that profit is strictly improving in a firm's incumbent technology level. These models are most applicable to manufacturing firms that adopt new technology in order to reduce the cost of production, rather than industries where the performance of products increases with technology.

Our formulation and assumptions are closely related to [36] and [43]. Similar to [36], the firm's optimization problem is structured as a Markov decision problem where the firm makes a binary decision between launching and delaying the next generation product release. Although the research by [36] has the added benefit of incorporating demand diffusion, their model is only applicable to a monopolistic environment where technology always increases the firm's potential market. Thus, delays in product upgrade result in discounted rather than lost sales. [43] consider a decision model from the perspective of a follower firm in a duopoly. The follower firm makes their upgrade decision in response to a market leader, given the assumption that the follower firm's demand rate depends on the performance gap between their own and the market leader's technology levels. The presence of the market leader is comparable to the outside option embedded in our model with respect to the impact of a lag in technology on demand. However, similar to [19], [50], and[59] market demand in Liu and Ozer (2009) is independent of technology improvements and technology improvements only cause erosion in the firm's market share. Thus, beyond the inclusion of branding and product failure, our formulation of the problem differs from [36] and [43] because technological advancement can result in both incremental market growth as well as attrition in the firm's market potential.

[45] analyze the decision of whether to pre-announce launch dates for the next generation product, or simply release upgrades on the go, in the presence of strategic customers. The results show that when facing strategic customers, firms should alternate between minor and major product upgrades. Furthermore, the firm should announce that there will be a long duration between the minor and major upgrade. The intuition behind the strategy is that the entire market will purchase after a major release. However, the firm must guarantee that there will be a sufficient duration between upgrades in order to incentivize the market to purchase the product with minor improvements. The stochastic advancement of technology encourages the entire market to re-purchase the product, but does not bring new consumers to the market. Thus, the model is most applicable to industries where the entire consumer base would consider purchasing the first generation product. Indeed, [45] show that if technology improvement is slow and prices are high, the firm will release a single generation and exit the market after the entire consumer base purchases the product.

As mentioned, an important contribution of this research is to explore how the risk of a failed product launch impacts product upgrades. Although firms purchase most of their new technology from component suppliers, they are often faced with significant design, production and supply chain challenges related to the integration, manufacturing and distribution of a new product ([69]; [51]). This problem is particularly prevalent in consumer technology industries where the launch of each new product is extensively scrutinized by a plethora of online technical and consumer oriented review sites in terms of performance metrics and product flaws. To date, other than documenting the scope of the problem, no research has analyzed how the risk of a failed product launch impacts the frequency and profitability of product upgrades.

2.3. Model Description

Consider a firm selling a single product comprised of component-based technology in discrete periods over a finite planning horizon. The component technology is supplied by third party vendors and advances stochastically over time, increasing the market potential for the firm. The firm sells the product at a fixed profit margin, in an environment which is sufficiently competitive such that there is always at least one competing product released with leading-edge component technology. At the start of each period, the firm decides whether to continue selling the incumbent product or to release the next generation product equipped with the leading-edge component technology. To focus the model on the timing of upgrades and avoid issues related to product rollover strategies, we make the common assumption that the firm upgrades the product to include leading-edge componentry and withdraws the old product from the market ([36]; [15]; [45]). To optimally time the upgrade decision, the firm must consider how the aggregation of pent-up demand due to brand commitment mediates the relationship between sales, the costs of the upgrade, and the potential attrition in the firm's customer base.

At a given technology level, a segment of the overall market has a preference for the firm based on factors related to branding such as marketing, reputation for quality, and complementary product network effects. This segment is commonly known as the firm's serviceable market. In each period, consumers from the firm's serviceable market look to purchase the firm's product with a baseline level of component technology. If the firm's product lags the leading-edge in a given period, then not all of the consumer arrivals will result in sales. Consumers who do not buy a product may delay their purchase to the next period rather than leave the market. If the firm has a high degree of brand commitment, these consumers are more likely to delay their purchase decision to the next period. Consequently, firms with brand commitment can build pent-up demand for a future product, while continuing to sell a product based on older technology. However, the firm can also experience erosion in the aggregation of pent-up demand, as a portion of the customers who have been waiting may exhaust their patience and leave the firm's market. Similar to the no purchase option in [41], consumers leaving the firm's serviceable market is interpreted as consumers purchasing a product with more advanced technology from a competitor.

2.3.1 Modeling Assumptions

At the start of period t, the level of pent-up demand is d_t, the firm's remaining serviceable market size is n_t, and the firm's product has a technological lag of z_t relative to the capabilities of the leading-edge. These parameters form the state components observed by the firm and are restricted to the set of nonnegative numbers.

In each period the portion of the remaining serviceable market n_t that arrives looking to purchase a product is $a(n_t)$. The number of arrivals who purchase the product in period t is dependent upon the firm's incumbent level of technology relative to the leading-edge. Given a technological lag z_t, the proportion of arrivals that purchase a product in period t is $\rho(z_t)$. We make the following assumptions on the consumer arrivals and sales process.

Assumption 1 (Arrivals and Sales). For market size n_t and technology lag z_t, the total number of sales in period t is $a(n_t)\rho(z_t)$. The proportion of total arrivals $a(n_t)$ is nondecreasing in the market size n_t with $\frac{da(n_t)}{dn_t} \leq 1$ and $a(0) = 0$. The proportion of arrivals that purchase, $\rho(z_t)$, is nonnegative, nonincreasing in technology lag z_t, and if there is no lag in technology, $\rho(0) = 1$.

Assumption 1 states that if the market size increases, then an additional consumer in the market at period t will not result in more than one additional consumer entering the market in period t. In a period where the firm's product contains the leading-edge technology (i.e., $z_t = 0$), all consumers that enter the market purchase the firm's product. If $z_t > 0$, then $a(n_t)\bar{\rho}(z_t)$ consumers, where $\bar{\rho}(z_t) = 1-\rho(z_t)$, look to purchase from the firm in period t, but do not buy. A portion of the consumers who do not purchase the product may delay leaving the market in favor of a competing product to see if the firm releases a product upgrade in the next period. The proportion of consumers who delay their purchasing decision depends on the degree of the firm's brand commitment, which is represented by the parameter $\theta \in [0, 1]$. If the firm has perfect brand commitment, then $\theta = 1$ and consumers who did not purchase the incumbent product wait for the upgrade rather than leave the market. If the firm has no brand commitment, then $\theta = 0$, and consumers never postpone their purchase decisions in hopes for an upgraded product in the future. If $\theta \in (0,1)$, then the consumers have a certain willingness to wait for the firm to upgrade to leading-edge technology. If consumers delay their purchasing decisions, they will only buy the product from the firm once the firm provides an upgrade.

Assumption 2 (Pent-Up Demand). If in period t the firm does not release a product upgrade with technology lag z_t, servicable market n_t, and d_t consumers who delayed their purchase in period $t-1$, then the number of consumers who delay their purchasing decisions in period t is $\theta(d_t + a(n_t)\bar{\rho}(z_t))$. If the firm upgrades in period t, then d_t consumers purchase the new product in period t and there is no pent-up demand at the end of the period, since $a(n_t)\bar{\rho}(z_t) = 0$.

Finally, once the firm realizes their sales and future pent-up demand, component vendors announce the technological progress made over the course of the period. The advancement of component technology is represented by the random variable ξ, where the advancement of technology increases the serviceable market of the firm.

Assumption 3 (Technological Progress). The advancements of technology over time are statistically independent and identically distributed as a nonnegative integer random variable ξ with a finite expectation, i.e. $E[\xi] < \infty$. An advancement of technology ξ results in an incremental market growth of $g(\xi)$, where $g(\xi)$ is nonnegative and nondecreasing in ξ.

2.3.2 Firm's Optimal Control Problem

The decision of whether or not to release the next generation product is represented by the binary decision variable x_t. If $x_t = 1$, then the firm continues to sell the incumbent product. If $x_t = 0$, then the firm launches the next generation product, pays fixed launch cost K, which incorporates various activities related to switching to a new product, such as withdrawing older generation products from the market, distributing and restocking the new product, marketing and rebranding, as well as costs related to sales, support and administration.

From the description of the modeling framework and the assumptions, the state dynamics proceed as follows. If the firm invests in an upgrade, then there will be no pent-up demand in the following period, since the pent-up demand d_t and new arrivals $a(n_t)$ all purchase the product. By Assumption 2, if the firm does not upgrade, pent-up demand will be $\theta(d_t + a(n_t)\bar{\rho}(z_t))$. If the firm has a technology lag of z_t at the start of period t and ξ is the stochastic increase in technological capability at the end of period t, then the firm's technology level at the start of period $t+1$ is ξ if the firm releases the next generation product in period t and $z_t + \xi$ otherwise. The firm's serviceable market at the start of period $t+1$ is $n_t - a(n_t) + g(\xi)$. Given that the firm earns profit π for each sale and future sales are discounted by $\delta \in [0, 1)$, the firm's optimization problem in period t is

$$V_t(d_t, n_t, z_t) = \max_{x_t \in \{0,1\}} \Big\{ \pi a(n_t)\rho(z_t) + \big(1 - x_t\big)\big(\pi\big(d_t + a(n_t)\bar{\rho}(z_t)\big) - K\big) + \delta E\big[V_{t+1}(x_t\theta(d_t + a(n_t)\bar{\rho}(z_t)), n_t - a(n_t) + g(\xi), x_t z_t + \xi)\big] \Big\}, \tag{2.3.1}$$

where without loss of generality $V_{T+1}(d_{T+1}, n_{T+1}, z_{T+1}) = 0$, for all $d_{T+1}, n_{T+1} \in \mathbb{R}^+$ and $z_{T+1} \in \mathbb{Z}^+$.

2.4. Pent-Up Demand and the Optimal Upgrade Policy

Although the optimal upgrade policy is often characterized in terms of the firm's technology level relative to the frontier, we develop an alternative threshold policy based on pent-up demand. The policy states that it is optimal to introduce the next generation product once the firm has built up a sufficient level of delayed demand among their committed consumers. Before characterizing the optimal policy, we remark on the monotonicity properties of the state dynamics and the profit-to-go with respect to pent-up demand.

Lemma 1. For two system states that evolve under the same sequence of technological advancements, if $n^1_{t_0} \leq n^2_{t_0}$ then $n^1_t \leq n^2_t$ for all $t \in \{t_0, \ldots, T\}$. If the two system states also have the same upgrade decisions, then

- for system states $(d^1_{t_0}, n^1_{t_0}, z_{t_0})$ and $(d^2_{t_0}, n^2_{t_0}, z_{t_0})$, if $d^1_{t_0} \leq d^2_{t_0}$ and $n^1_{t_0} \leq n^2_{t_0}$, then $d^1_t \leq d^2_t$ for all $t \in \{t_0, \ldots, T\}$, and
- for system states $(d_{t_0}, n_{t_0}, z^1_{t_0})$ and $(d_{t_0}, n_{t_0}, z^2_{t_0})$, if $z^1_{t_0} \leq z^2_{t_0}$, then $z^1_t \leq z^2_t$ for all $t \in \{t_0, \ldots, T\}$.

The monotonicity property of the state dynamics implies that under the same realization of technology, if two firms have the same product introduction timing, then the firm with greater pent-up demand or better technology will maintain their advantage until both firms upgrade their products, whereas the firm with greater market potential will maintain the advantage until time T. In addition, the lemma establishes that the advantage resulting from greater pent-up demand and market potential is nonincreasing with time.

Proposition 1. The profit-to-go $V_t(d_t, n_t, z_t)$ is increasing in the state component d_t and n_t and is nonincreasing in state component z_t.

To develop the intuition of the monotonicity of the profit-to-go function with respect to pent-up demand, we define market loss as the number of consumers who leave the market without purchasing from the firm. All else equal, if two firms with different levels of pent-up demand experience the same consumer arrival pattern, the proportions of new arrivals that result in sales and market loss will be the same for both firms. This implies that the firms will proceed to build up or lose the same amount of pent-up demand from consumer arrivals over the remaining planning horizon. The monotonicity with respect to pent-up demands holds because the size of the firm's pent-up demand does not impact the future behavior of arrivals. Therefore, greater pent-up demand simply implies that the firm has more opportunities for sales, which increases the profit-to-go.

Proposition 2. If for fixed n_t, and z_t there are two systems with pent-up demands d_t^1 and d_t^2 respectively, where $d_t^1 \leq d_t^2$, then $V_t(d_t^2, n_t, z_t) \leq V_t(d_t^1, n_t, z_t) + \pi(d_t^2 - d_t^1)$.

Proposition 2 utilizes the results from Lemma 1 to bound the difference in the profit-to-go. The tightness of the bound depends on the upgrade strategy of the firm. Given that for both firms the loss of consumers in the serviceable market is the same, the firm with the demand advantage (either from a greater market size or more pent-up demand) will not be able to accumulate more demand going forward. Thus, the longer the firm with the demand advantage waits to upgrade the product, the greater their potential to lose consumers to competitors. If the firms never upgrade, then the pent-up demand advantage is never utilized, and both firms achieve equal profits.

We now demonstrate that pent-up demand can be utilized to determine whether it is optimal to upgrade the product at time t. The policy states that for each pair (n_t, z_t), there exists a threshold level of pent-up demand $D_t^*(n_t, z_t)$, which dictates the firm's upgrade decision. It is optimal for the firm to introduce (delay) the next generation product if and only if the pent-up demand is above (below) the value $D_t^*(n_t, z_t)$.

Proposition 3. There exists a threshold $D_t^*(n_t, z_t)$ such that it is optimal to upgrade the product if and only if the pent-up demand $d \geq D_t^*(n_t, z_t)$.

The threshold policy provided by Proposition 3 defines a set of switching curves based on pent-up demand over the state space $\{n_t, z_t\}$. Since the optimal policy is to upgrade when $d \geq D_t^*(n_t, z_t)$, if the level of pent-up demand is exactly at or above the plotted value, then it is optimal for the firm to invest in an upgrade. The next proposition demonstrates monotonicity of the threshold policy in z_t.

Proposition 4. For a fixed value of n_t, the optimal threshold value $D_t^*(n_t, z_t)$ is non-increasing in the technology lag z_t.

Proposition 4 demonstrates that the threshold value $D_t^*(n_t, z_t)$ is nonincreasing in technology lag z_t. This implies that as the technology lag grows, there is greater pressure on the firm to upgrade their product, and that at smaller technology lags, firms require a large amount of aggregate delayed demand to warrant an upgrade. The monotonic behavior of the threshold policy implies that the optimal policy can alternatively be expressed as thresholds in the lag of technology $Z_t^*(d_t, n_t)$.

Corollary 1. An equivalent threshold policy can be constructed using $Z_t^*(d_t, n_t)$ where it is optimal to upgrade the product if and only if the technology lag $z \geq Z_t^*(d_t, n_t)$.

2.5. Experiments for Baseline Model

In this section, we present the results from our numerical experiments. For consistency with the literature, the experiments utilize the same functional form for the market growth and technological advancement as [36]. The advancement of component technology in each period is binary, improving by 1 unit with probability λ and remaining the same with probability $1-\lambda$. Therefore, $\lambda \in [0,1]$ represents the pace of technological advancement. The firm's potential market size grows linearly with the advancement of technology. Given an existing market size of n_t, the market size in the following period is $n_{t+1} = n_t + \eta\xi$. The parameter η represents the incremental growth in the firm's serviceable market due to an increase in technology. The market arrival process for a serviceable market size of n_t is $a(n_t) = \alpha n_t$ and the erosion in sales due to a technology lag is $\rho(z_t) = e^{-\gamma z_t}$, where $\alpha \in [0,1]$ and $\gamma \geq 0$. The parameter α captures the speed at which consumers from the serviceable market look to purchase a product, while the parameter γ is consumer sensitivity to a lag in technology.

Since the level of pent-up demand and the market potential is continuous, interpolation is used to discretize the values for the future state dynamics. Define the $D^f(d_t, n_t, z_t)$ and $N^f(n_t)$ as the floor function of pent-up demand $D(d_t, n_t, z_t)) =$

Table 2.1: Baseline Parameters

α	γ	λ	δ	η	π	K	N_0	T
0.4	0.4	0.8	0.9	5	1	2.5	20	50

$\theta\big(d_t + a(n_t)\bar{s}(z_t)\big)$ and market potential $N(n_t) = n_t - a(n_t) + g(\xi)$, and define

$$D^c(d_t, n_t, z_t) = \begin{cases} \lceil D(d_t, n_t, z_t) \rceil & D(d_t, n_t, z_t) > D^f(d_t, n_t, z_t) \\ \lceil D(d_t, n_t, z_t) \rceil + 1 & D(d_t, n_t, z_t) = D^f(d_t, n_t, z_t), \end{cases} \text{ and}$$

$$N^c(n_t) = \begin{cases} \lceil N(n_t) \rceil & N(n_t) > N^f(n_t) \\ \lceil N(n_t) \rceil + 1 & N(n_t) = N^f(n_t). \end{cases}$$

For non-integer values of D and N, if the probabilities of being in the states D^i and N^j for $i, j \in \{f, c\}$ are given respectively by $q_D^f(d_t, n_t, z_t) = D^c(d_t, n_t, z_t) - D(d_t, n_t, z_t)$, $q_D^c(d_t, n_t, z_t) = D(d_t, n_t, z_t) - D^f(d_t, n_t, z_t)$, $q_N^f(n_t) = N^c(n_t) - N(n_t)$, $q_N^c(n_t) = N(n_t) - N^f(n_t)$, then the approximate dynamic program of (2.3.1) is

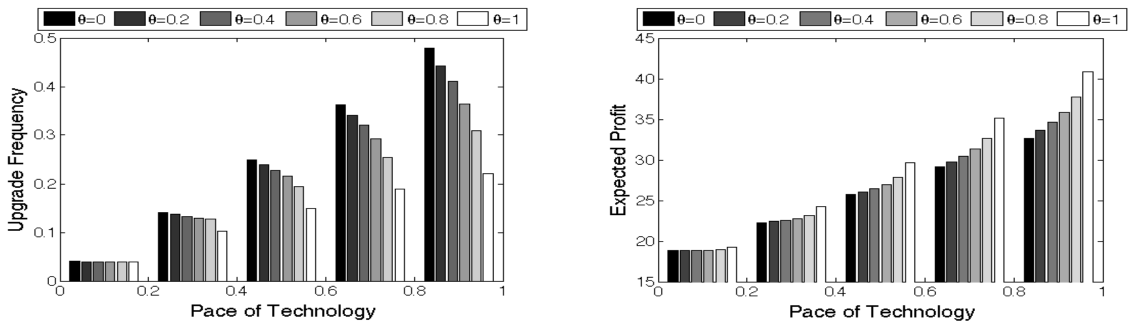
$$V_t(d_t, n_t, z_t) = \max\Big(I_t(d_t, n_t, z_t), W_t(d_t, n_t, z_t)\Big),$$

$$I_t(d_t, n_t, z_t) = \pi\big(d + a(n_t)\big) - K + \delta \sum_{\xi \in \{0,1\}} p(\xi) \sum_{j \in \{c,f\}} q_N^j(m, z_r, z_r) V_{t+1}(0, N^j(n_t), \xi), \text{ and}$$

$$W_t(d_t, n_t, z_t) = \pi a(n_t) s(z_t) + \delta \sum_{\xi \in \{0,1\}} p(\xi) \sum_{i,j \in \{c,f\}} q_D^i(d_t, n_t, z_t) q_N^j(n_t) V_{t+1}(D^i(d_t, n_t, z_t), N^j(n_t), z_t + \xi).$$

To explore the impact of the various parameters in the models, we solve the approximate dynamic program over a planning horizon of 50 periods. The experiments considered the parameter set $\alpha \in \{0.2, 0.3, \ldots, 0.6\}$, $\gamma \in \{0.2, 0.3, \ldots, 0.6\}$, $\eta \in \{3, 4, 5, 6, 7\}$, $K \in \{2, 3, 4, 5\}$, and $\theta \in \{0, 0.1, \ldots, 1\}$, with the remaining baseline parameter values given by Table 2.1. For each problem instance, 1000 simulations were generated in order to track the average number of upgrades over the planning

Figure 2.1: Comparative statistics for various degrees of brand commitment and pace of technological improvement.

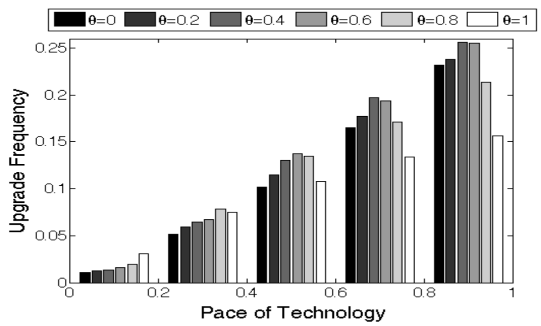


(a) Upgrade Frequency (b) Expected Profit

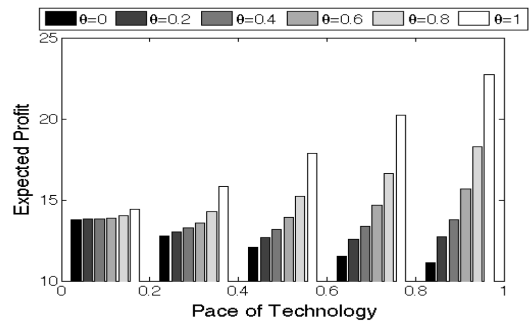
horizon. Each simulation started in the state $(0, N_0, 0)$ and was executed using the SHARCNET serial throughput cluster Kraken. Intuitively the experiments demonstrated that greater brand commitment leads to less upgrades, greater sales from pent-up demand, and thus greater profit. To understand the market conditions where brand commitment has the greatest influence, we provide a sensitivity analysis on the model's parameters.

Impact of the Pace of Technology. Figure 2.1 (a) displays a breakdown of the average number of upgrades over the planning horizon with respect to $\lambda \in \{0.1, 0.3, \ldots, 0.9\}$ and brand commitment averaged across the market scenario values. The figure shows that the average number of upgrades is correlated with the advancement of technology at each level of brand commitment. As the speed of technology increases, firms upgrade more frequently to prevent erosion in sales and capture incremental market growth associated with the advancement of technology. Since the difference in the vertical bars between various degrees of brand commitment increases with λ, this implies that at higher brand commitment, the frequency of upgrades over the planning horizon is less influenced by the pace of technology. As

Figure 2.2: Comparative statistics for various degrees of brand commitment and pace of technological improvement at high launch cost and low incremental market growth.



(a) Upgrade Frequency

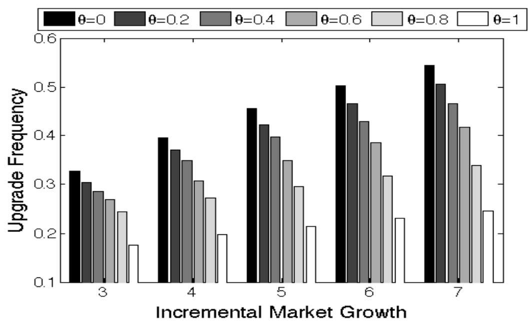


(b) Expected Profit

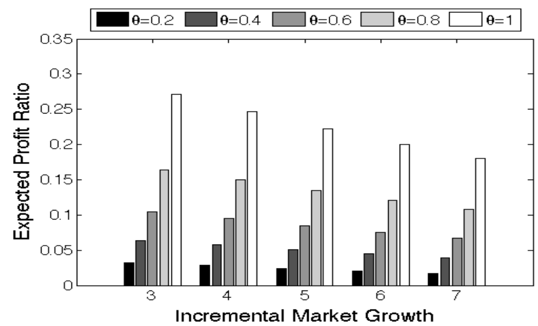
expected, Figure 2.1 (b) shows that less frequent upgrades due to the pace of technology is correlated with increasing levels of profit. The figure also demonstrates that firms are typically more profitable when technology improvements are faster, even in the absence of brand commitment.

Although figure 2.1 shows that firms with greater brand commitment upgrade less over the planning horizon than firms with lower brand commitment, and that profits improve with faster paces of technology, these results do not necessarily hold in all test instances. Figure 2.2 shows that firms with low brand commitment may upgrade less frequently than firms with high brand commitment and may experience a decline in profit due to a faster pace of technology when market growth is low ($\eta = 3$) and upgrade costs are high ($k = 4$). These results demonstrate that market growth and launch cost have a significant influence over both upgrade frequency and profitability. Since brand commitment has a greater influence at faster paces of technology, the pace of technology for the analysis of market growth and launch cost is restricted to the set $\lambda \in \{0.6, 0.7, \ldots, 1\}$.

Figure 2.3: Comparative statistics for various degrees of brand commitment and incremental market growth.



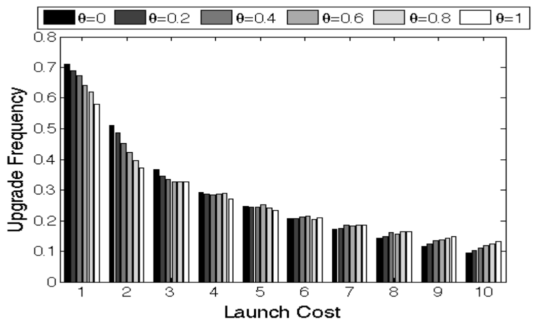
(a) Upgrade Frequency



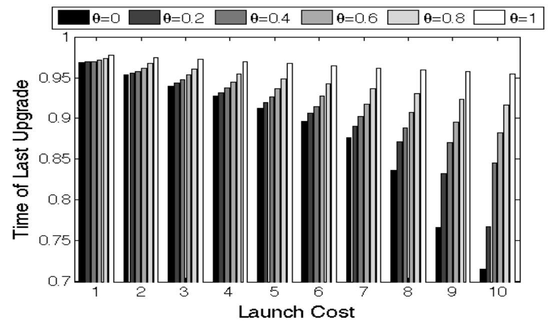
(b) Expected Profit

Impact of Incremental Market Growth. A greater level of incremental market growth η over the planning horizon increases the market potential due to the advancement of technology but also creates the potential for lost sales if a firm continues to sell an incumbent product with older technology. As can be seen in Figure 2.3 (a), as η rises, the frequency of upgrades for firms with lower brand commitment increases at a faster rate than for those with higher brand commitment. If a firm has sufficiently high levels of brand commitment, in the absence of an upgrade, it can maintain their market potential by aggregating pent-up demand. On the other hand, firms with low to moderate brand commitment must rely on more frequent upgrades to achieve these objectives. A higher market growth factor will increase profits for all firms irrespective of their level of brand commitment. Thus, we study the expected profit ratio of a firm with brand commitment level θ relative to a firm with no brand commitment. Figure 2.3 (b) shows that a slower market growth factor increases the comparative profit ratio of firms with higher brand commitment relative to a firm with no brand commitment. Indeed, although firms with no brand commitment have to upgrade more frequently due to sales erosion, they only capture modest gains in

Figure 2.4: Comparative statistics for various degrees of brand commitment and launch costs.



(a) Upgrade Frequency



(b) Last Upgrade Relative to Planning Horizon

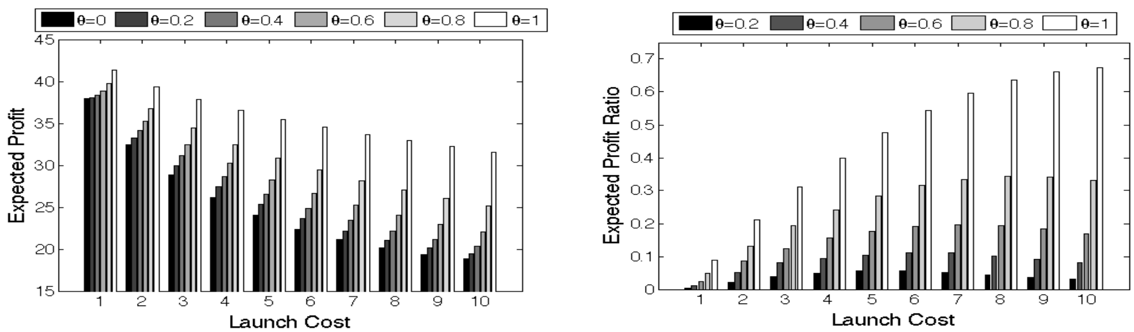
incremental sales associated with the demand for the new technology. Therefore, increases in η favor a frequent upgrade strategy, which is why the expected profit ratio for firms with high brand commitment decreases.

As observed by [3], there are diminishing returns in the functionality of products from incremental advancements in technology as the market matures. Thus, the incremental market growth related to technology will tend to decrease as the market reaches maturity. Based on the results in Figure 2.3 (b), we infer that brand commitment becomes increasingly important to a firm's profitability in a mature market, since such markets are characterized by diminishing values of η.

Impact of Launch Cost. We explore how the cost of the product launch impacts the frequency of upgrades as well as when a firm's last upgrade occurs over the planning horizon specified in the model. Figure 2.4 (a) shows the average number of upgrades for different launch costs at various levels of brand commitment. As expected, the number of upgrades is inversely related to the costs of the product launch for all firms. However, the variability in the frequency of upgrades between firms with different levels of brand commitment shrinks as the launch costs increase. This finding may be misleading when considered in isolation because firms with low level of brand commitment stop upgrading long before the end of the planning horizon. When Figures 2.4 (a) and (b) are examined together, they show that as the variability in upgrade frequency between firms with different levels of commitment contracts with the increase in launch costs, the corresponding variability in when the last upgrade occurs expands. These results indicate that as the launch cost increases, firms with less brand commitment are unable to continue launching products, and are no longer viable in the market. On the other hand, when faced with high launch costs, firms with high brand commitment accumulate significant levels of pent-up demand, which reduces their frequency of upgrade but at the same time enables them to remain viable in the market throughout the planning horizon. This explains why the variability in profits between firms with different levels of brand commitment, displayed in Figure 2.5, increases dramatically in relation to the launch cost. Taken together, Figures 2.4 and 2.5 show that when the launch cost of upgrades is disproportionally high relative to profits, only firms with high levels of brand commitment are able to upgrade profitably throughout the lifecycle of the product in fast paced industries.

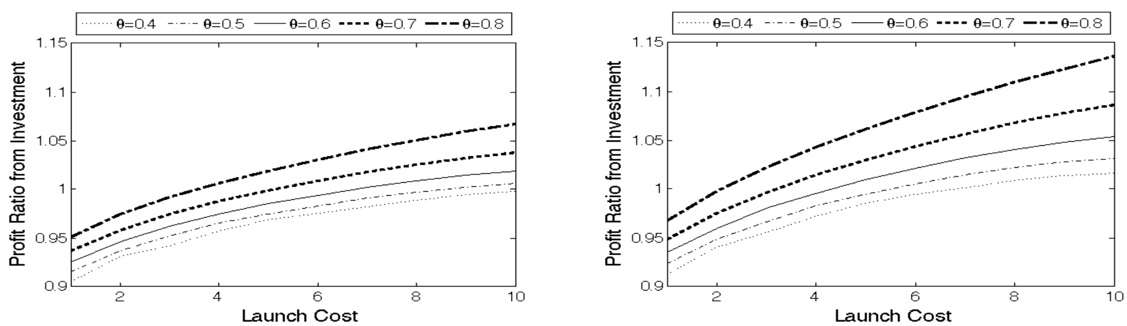
Figure 2.5 also suggest that a competitive advantage in the launch cost of one firm over another can be offset by an increase in brand commitment. For example, a firm with 80% brand commitment and a launch cost of 10 earns the same amount of profit as a firm with 20% brand commitment and a launch cost of 5. These findings may explain why the firms that are long-term survivors in markets for technology-intensive products are not always the lowest-cost producers, but often include firms with well established brands ([16]). They also raise the issue of whether firms can profit from

Figure 2.5: Profit plots for various degrees of brand commitment and launch costs.



(a) Expected Profit (b) Expected Profit Relative to No Brand Commitment

Figure 2.6: Expected profit improvement from an investment in raising brand commitment.



(a) 10% Increase in Brand Commitment (b) 15% Increase in Brand Commitment

investing in brand commitment at the expense of increasing their launch cost.

Investments in Brand Commitment. Firms are often faced with the decision of whether to increase the level of brand commitment by improving the post-sales customer experience through investments associated with developing or launching the product. Such investments include improving the quality of product finishings, making the product more user friendly, or enhancing the after sales customer support infrastructure ([27]; [72]). Figure 2.6, plots the profit improvement from increasing the upgrade costs by a fixed amount to raise brand commitment by a fixed percentage. For the purpose of this example, we consider the profit from a \$1 investment to raise brand commitment by 10% and 15% relative to the firm's profit level without investment for a spectrum of launch costs. Contrary to what a manager might expect, Figure 2.6 shows that the outcome from these investments is a function of a firm's baseline level of brand commitment: only firms which already have high baseline levels of brand commitment are likely to profit from such investments. For example, at a 15% increase in brand commitment, the investment is almost never profitable for a firm below 40% commitment. By comparison when commitment is 80%, firms will accrue profits from the investment for most of the launch costs analyzed.

In the absence of the model, managers might expect that firms with low to moderate brand commitment would benefit more from these investments than firms that already have high levels of brand commitment. An explanation for this unexpected finding is that in an industry where markets for incumbent products erode quickly due to stochastic increases in technology, it takes high levels of pent-up demand to decrease the frequency of upgrades in a meaningful way. Therefore, for firms with low to moderate levels of brand commitment, minimizing the costs of the upgrade would be a preferable strategy to one which incrementally increases brand commitment but at higher costs. Conversely, for firms with high levels of brand commitment which already benefit from a reduction in the frequency of upgrades, minimizing lost sales due to market erosion over the longer sales window for an incumbent product would be more beneficial, even if it adds to the cost of product launches.

2.6. Product Failure and Investments in Brand Commitment

In industries for mass consumer technology products, failed product launches are a frequent occurrence. This is particularly true in the current sales environment, where product reviews from professional and social online media sites are disseminated to consumers, often before products are launched ([14]). For technology products that integrate multiple components from different vendors, a firm's risk of a failed product launch varies throughout the product's lifecycle. The risk of a failure can escalate with the size of the upgrade in terms of complexity and variation of the componentry ([63]; [49]). For example, a product upgrade that involves integrating many new chips with smaller process nodes scales dramatically in terms of complexity.

The magnitude of a product failure influences the degree of lost sales associated with the product launch. For example, a malfunctioning sensor in a multi-attribute product may have a modest impact on sales, whereas a defective battery could result in a much steeper loss of sales. Therefore, the firm's upgrade strategy, which dictates the potential market size for a product release, and the potential of a failure, which can vary in magnitude, influence the firm's expected profit from an upgrade. The issue of how a firm's upgrade strategy influences the profit loss from a failed product launch has not been explored in the literature. In this section of the paper, we explore this issue and determine how a possibility of a failed product impacts the threshold policy of the upgrade decision.

To incorporate the degree of product failure, the state space of the model is expanded to include the indicator f_t which takes on an integer value in the set $\{0, 1, \ldots, F\}$. The probability that an arriving consumer considers purchasing from the brand is denoted by $p(f_t)$. The probability $p(f_t)$ is decreasing in f_t, where $p(0) = 1$

and $p(F) \geq 0$. If the firm launches a product upgrade which obtains a failure level of $f_{t_0} = j$, then the sales in period t_0 will be $p(j)\big(d_{t_0} + a(n_{t_0})\big)$. If the next upgrade is at time τ, then the sales and pent-up demand in each period $t \in \{t_0+1, \ldots, \tau\}$ will be $p(j)a(n_t)\rho(z_t)$ and $\theta\big(d + p(j)a(n_t)\bar{\rho}(z_t)\big)$ respectively. The likelihood and degree of a product failure is stochastic and can be dependent on the lag in technology. An "innovative" versus "incremental" product improvement has inherent risk in successfully launching the product. Although the upgrade is based on component technology, firms have to integrate the component technology into the existing product architecture. In addition, there is a potential for manufacturing issues or disruptions in the supply chain which can lead to product failure. The probability of a product failure of degree j is $q_j(z_t)$, where $\sum_j q_j(z_t) = 1$ for all $z_t \in \mathbb{Z}^+$. Thus, the firm's optimization problem is

$$V_t(d_t, f_t, n_t, z_t) = \max\{I_t(d_t, f_t, n_t, z_t), W_t(d_t, f_t, n_t, z_t)\} \tag{2.6.1}$$

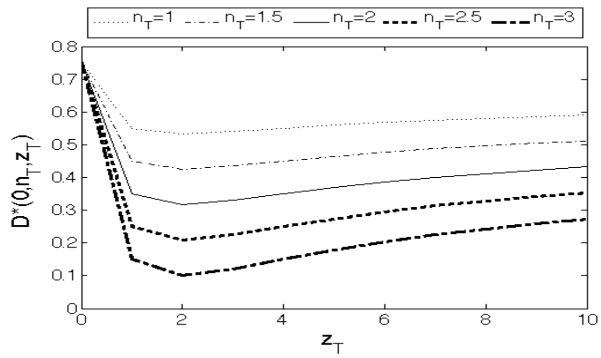
$$\begin{aligned} I_t(d_t, f_t, n_t, z_t) = \sum_{j=0}^{F} q_j(z_t) \bigg(& \pi p(j)\big(d_t + a(n_t)\big) - K \\ & + \delta E\big[V_{t+1}(0, j, n_t - a(n_t) + g(\xi), \xi)\big] \bigg) \end{aligned} \tag{2.6.2}$$

$$\begin{aligned} W_t(d_t, f_t, n_t, z_t) = {} & \pi p(f_t)a(n_t)\rho(z_t) \\ & + \delta E\big[V_{t+1}(\theta(d_t + p(f_t)a(n_t)\bar{\rho}(z_t)), f_t, n_t - a(n_t) + g(\xi), z_t + \xi)\big], \end{aligned} \tag{2.6.3}$$

where without loss of generality $V_{T+1}(d_{T+1}, f_{T+1}, n_{T+1}, z_{T+1}) = 0$, for all $d_{T+1}, n_{T+1} \in \mathbb{R}^+$ and $f_{T+1}, z_{T+1} \in \mathbb{Z}^+$.

The potential for releasing a failed product implies that the threshold results (Proposition 3, Proposition 4, and Corollary 1) established for the baseline model

Figure 2.7: State-dependent threshold policy with potential for a product failure



may not hold. The following result demonstrates that the upgrade policy can still be characterized as a state-dependent threshold policy based on pent-up demand, where it is optimal to upgrade if and only if $d \geq D(f_t, n_t, z_t)$.

Proposition 5. The optimal policy (2.6.1)-(2.6.3) is to upgrade if and only if $d \geq D(f_t, n_t, z_t)$.

An interesting property of the optimal policy given the potential for a product failure is that the optimal threshold value $D(f_t, n_t, z_t)$ is not guaranteed to be monotonically nonincreasing in z_t. We demonstrate the failure of a technology based threshold policy by constructing a simple example.

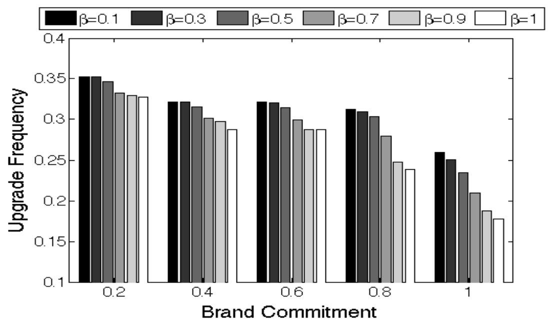
Example 1. Consider the parameters and functional forms $F = 1$, $K = 0.75$, $\pi = 1$, $p(1) = 0.1$, $q_0(z_t) = 2/(2 + z_t)$, $\rho(z_t) = 1/(1 + z_t)$, and $a(n_t) = n_t$. The parameters imply that a product is either a hit (a failure of degree 0) or a miss (a failure of degree 1). The probability of a failure of degree 1 is $1 - q_0(z_t)$ and the probability of purchase given a miss is $p(1)$. Consider a firm in the last stage of the planning horizon, where the incumbent product was a hit, i.e. $t = T$ and $f_T = 0$. Figure **??**, which plots the state dependent threshold policy $D^*_T(0, n_T, z_T)$ for various levels of n_T and z_T, clearly demonstrates that the threshold values are not monotonically nonincreasing in z_T.

Since the pent-up demand threshold value is not monotonically nonincreasing in z_t, a single technology lag threshold upgrade policy may not exist. The presence of multiple technology threshold values reveals that the structure of the profit-to-go function is more complex in the presence of product failures. For high regions of technology lag, a firm may want to continue accumulating pent-up demand instead of releasing an upgrade. Although the lag in technology continues to grow, the firm may be able to increase the level of pent-up demand, making an upgrade profitable once again. This can provide incentive for firms to further delay product upgrades, raising the optimal pent-up demand threshold values. On the other hand, if there is a substantial risk associated with long durations between upgrades, then product failures may force firms to release their upgrades *more* frequently. Therefore, we extend the experiments from section 2.5 to investigate the influence of product failures in combination with brand commitment on the frequency and profitability of upgrades. We set $F = 1$ so that the firm either releases a product which is a hit or the firm releases a product which is a failure from the perspective of the consumer. The probability of purchase for a failed product is $p(1) = \bar{\beta}$, where $\bar{\beta} \in [0, 1]$. The functional form for the probability that a product is a failure is taken as $q_1(z_t) = 1 - e^{-\phi z_t}$. Given that brand commitment has a greater impact at faster technology speeds, at higher costs, and faster sales erosion we shifted some baseline parameters to the values in Table 2.2.

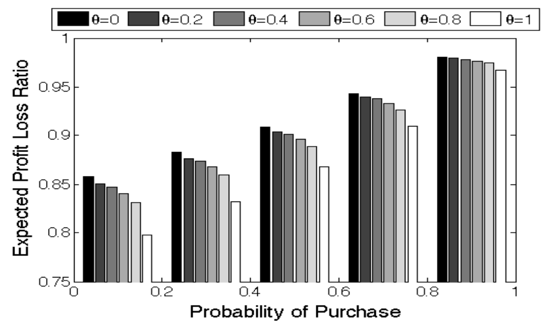
Table 2.2: Experiments with Failure Parameters

α	γ	λ	δ	η	π	K	N_0	T
0.4	0.6	0.9	0.9	6	1	6	20	50

Figure 2.8: Upgrade frequency and profit loss across brand commitment and probability of purchase given a failure.



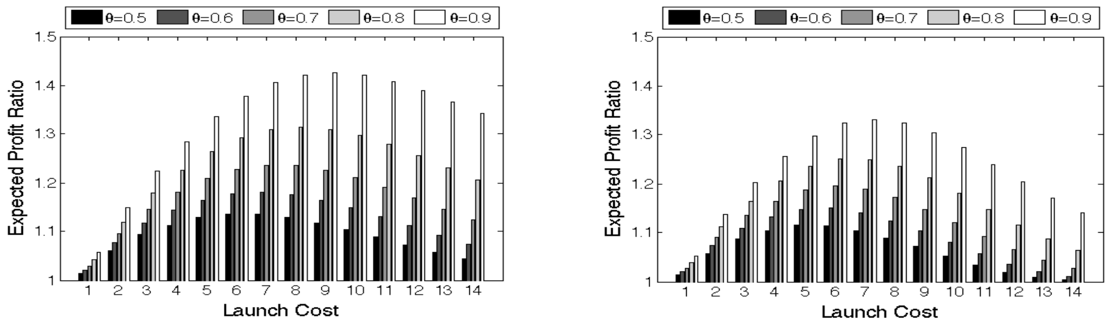
(a) Upgrade Frequency



(b) Expected Profit Loss Ratio

Figure 2.8 (a) displays the upgrade frequency due to the risk of releasing a failed product for different levels of brand commitment. Note that $\beta = 1$ is equivalent to the baseline model where there is no risk of failure. Several important observations are discernible from Figure 2.8 (a). Firms are forced to upgrade more frequently as the probability of purchase given a failure decreases (lower β). Furthermore, this effect becomes more pronounced as the level of brand commitment increases. As the magnitude of product failures escalates, the changes in frequency is marginal at low degrees of brand commitment, whereas, at high levels of brand commitment, frequency increases substantially. The reasons for this are two fold. First, firms increase the risk of having a failed product by delaying the upgrade. Second, firms that delay an upgrade to aggregate pent-up demand are at risk of suffering disproportionately from lost sales. Thus, the optimal upgrade policy for firms with high brand commitment balances the benefits of pent-up demand with the losses due to product failures. Figure 2.8(b) shows the profit loss ratio, which is the expected profit at β as a percentage of the expected profit with no risk of failure, across different values of brand commitment and probability of purchase. Since firms with high brand commitment

Figure 2.9: Expected profit relative to no brand commitment at different launch costs and probability of purchase.



(a) $\beta = 0.9$ (b) $\beta = 0.1$

have to upgrade more frequently compared to the baseline model, as β increases, the losses due to product failures increase substantially for firms with higher levels of brand commitment.

Figure 2.9 examines the relative benefit of brand commitment defined as the ratio of the expected profit for $\theta \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ to the expected profit of a firm with no brand commitment at high ($\beta = 0.9$) and low ($\beta = 0.1$) probabilities of purchase. These graphs show that the relative benefits of brand commitment with respect to increasing the launch cost become less pronounced as the magnitude of the impact of a product failure increases, i.e. as β decreases. At the same time, the peak values of the relative profit ratios shifts in the direction towards lower costs. Although at high costs, firms with high brand commitment still enjoy a profit advantage over firms with no commitment, this does not hold for firms with low-moderate commitment levels. This is evident from the example of $K = 14$ and $\beta = 0.1$, where firms with commitment levels between 0.5 and 0.7 earn marginal profit improvements over firms without any brand commitment. In industries where a product failure leads to a large proportion of consumers to exit the firm's market, investments in brand commitment are less valuable. This is an important insight given that brand commitment could be seen as a tool to buffer the potential loss of consumers due to a failed product.

The profit loss associated with an upgrade strategy based on aggregating pent-up demand is highly sensitive to the probability of purchase. Since brands with higher levels of commitment have higher expectations from their consumers, a lower probability of purchase due to a product failure is likely to be greater for firms with high commitment ([56]). In addition, these findings represent a best case scenario for losses due to a product failure, since the model assumes that the firm is able to relaunch a successful upgrade in the following period. Furthermore, the model assumes that a product failure does not result in the firm losing commitment from future consumers. In reality, a failed product launch has a negative impact on a company's reputation and will hurt the firm's level of brand commitment going forward. Thus, the results underestimates the expected profit loss for firms with high brand commitment.

In terms of the smartphone industry, Blackberry is a compelling example of the risk that a firm with high brand commitment faces by having a failed product launch in the presence of significant pent-up demand. By comparison, LG, which releases multiple smartphone models throughout the year and does not count on a committed customer base, experienced a number of disappointing product launches in 2012. This only impacted their sales for that calendar year and had little effect on their sales for product launches in successive years. To date Apple has managed to avoid disappointing product launches with the iPhone, although the viral spread of the Bendgate rumor associated with the launch of the iPhone 6+ is a potent reminder of how quickly news on a product flaw can spread through social media sites. Fortunately for Apple, they were able to effectively counter the rumor through social media and third party reviews which demonstrated that the iPhone 6+ did not bend from regular use.

2.7. Conclusion and Future Research

The primary contribution of this research is to analyze the impact of brand commitment and product failure on the timing and profitability of product upgrades. We demonstrate that a firm's optimal upgrade policy is based on a threshold value of pent-up demand, which is monotonically nonincreasing in relation to the lag in technology. We conduct extensive experiments to show how the model performs under various firm and industry related conditions. The results lead to important managerial insights regarding the impact of pace of technology, market growth, launch costs and investments in brand commitment on the timing of upgrades and a firm's profits. We extend the baseline model to examine how the possibility of a failed product launch impacts the results based on a firm's level of brand commitment. The potential of a failure changes the structure of the optimal policy, and shows that a threshold policy based on pent-up demand is more robust than a policy based on a critical value in the lag in technology. We show that incorporating the risk of a failed product launch drastically alters the upgrade strategy and profits for firms with high brand commitment.

There are several avenues for future research related to our formulation of the upgrade problem in terms of brand commitment and of pent-up demand. The aggregation of pent-up demand has important implications for inventory procurement in terms of volume discounts and predictability of customer arrivals. Factoring improved economies of scale into the decision process may provide further insights into the benefits of accumulating pent-up demand prior to a product launch. These effects may help to explain why those firms which build-up high levels of pent-up demand for a product launch often enjoy much higher margins than the firms that upgrade more frequently.

Another avenue of research is to explore how the emergence of secondary markets impacts the upgrade decision for firms with high brand commitment. Recently, there has been a great deal of interest in the role that secondary markets play in incentivizing previous customers to purchase upgrades by enabling them to monetize the residual value of their previous purchases. Firms with high brand commitment are likely to be the main beneficiaries of this trend since their customers are more likely to repurchase their products. Although brand commitment lengthens the duration between upgrades, firms with higher rates of brand commitment may have more incentive to accelerate their product cycles to maximize repurchases from their customer base. Therefore, to the extent that this trend becomes more prevalent, it may impact the dynamics of the timing of product upgrades.

Finally, this work could benefit from more precise data on how quickly firms in fast paced technology industries can recover from a failed product launch and how it impacts brand commitment on a go forward basis. While there are some general assumptions and theoretical models about how a failed product launch impacts brand commitment and vice versa, there is little in the way of empirical research that investigates the issue systematically. The importance of this issue will only grow as the pace of technology and product complexity continues to accelerate in consumer electronics and other related industries.

Chapter 3

Scalable Dynamic Bid Prices for Network Revenue Management in Continuous Time

3.1. Introduction

Stochastic network capacity control problems are characterized by customers arriving at random instances in time requesting products, which require different combinations of perishable resources. The most common application of network capacity control problems in revenue management (RM) is the control of reservations for airline tickets. In the airline RM problem, the single-leg flights correspond to the perishable resources and the available origin-destination (OD) itineraries comprise the set of products. Airlines can maximize their revenue by determining whether or not to accept an itinerary request. Given sufficient resources, the optimal policy for an airline is to accept a booking request for an itinerary if its fare is greater than the opportunity cost of required resources. The opportunity cost of each flight depends on the number of available seats in the network and the remaining time in each flight's booking horizon. Although the optimal controls can theoretically be found using the Bellman equation, even for a modest sized airline network the problem suffers from the curse of dimensionality.

The intractability of the airline RM problem has lead to an extensive literature proposing various heuristics and approximate solutions. One fundamental approach is to derive bid prices, which approximate the value of each resource, using a deterministic linear program (LP). The resulting control policy is to sell the itinerary if its fare is greater than the sum of the bid prices of seats on flights in the requested itinerary. The shortcoming of the bid price policy based on the LP is that the bid prices are static over the booking horizon. Thus, the bid prices do not adjust as the system evolves until the LP is re-optimized using the updated capacity levels.

[1] proposed an alternative bid price approach, the dynamic linear program (DLP), by modeling the revenue function with an affine approximation in order to obtain time-dependent bid prices. [1] derived the DLP directly from the Bellman's equation, developed appropriate bounds on the performance of the method, demonstrated structural properties, and proposed a solution approach based on column generation. However, approximating the revenue function and generating time-dependent bid prices remains difficult for large networks due to the number of variables and constraints in the DLP problem. Network models that generate dynamic (time or capacity dependent) bid prices have become an active area of RM research, since [1]'s seminal paper. Research on dynamic bid prices extends existing formulations to incorporate other aspects of RM or provides novel computational approaches to improve performance.

[65] decomposed itinerary request decision by flight leg using Lagrangian relaxation to create capacity-dependent bid prices. [38] subsequently extended this model

by also incorporating time-dependence into the inventory-sensitive bid prices. Extensions of existing models include [73], who integrated [44]'s customer choice based linear program into [1]'s initial model and [37], who incorporated customer choice behavior into their Lagrangian relaxation time-sensitive bid prices model. [48] formulated an inventory-sensitive bid price model with customer choice and demonstrated that their model encapsulates the choice based linear programs by [73], [37], and [44]. Since we focus strictly on time-dependent bid prices, the phrases 'dynamic bid prices' and 'time-dependent bid prices' are used interchangeably throughout this paper.

Both [64] and [68] present new methods for producing time-dependent bid prices, recognizing that the number of variables and constraints as well as the column generation procedure prevent the DLP from being implemented in practice. [64] showed that the column generation problem for the dual to [1]'s DLP can be solved in polynomial time by using a minimum cost network flow problem. Using the network flow structure, they were able to reduce the number of constraints from exponential to linear in the number of flight legs. [68] use a Dantzig-Wolfe reformulation to provide the same reduction found in [64]. [68] also develop a novel dynamic disaggregation algorithm to solve the reduced programs, utilizing the fact that bid prices only change towards the end of the booking horizon.

Although network RM problems are traditionally modeled based on discrete time, in reality, customer arrivals occur in continuous time. As [11] point out in their overview of RM and pricing, continuous time models of RM problems are more appropriate than their discrete counterparts given the growth of e-commerce and the Internet. Bitran and Caldentey's observation could not be more relevant for network

RM in the current business environment as exemplified by travel websites such as Expedia, Priceline, Travelocity and Orbitz which attract over 7.5 millions people each day.

While continuous time formulations are used extensively in dynamic pricing research, only [5] have used a continuous time framework to obtain and study bid prices in the network RM setting. Akan and Ata highlight that the non-optimality of classical bid-price controls stems from discreteness and implement a generalized bid price control, which combines bid-prices with a capacity usage limit process to determine how much of the demand rate to accept. Using a nonnegative stochastic process to characterize bid prices, they demonstrated that the optimal generalized bid price process forms a martingale. While Akan and Ata's model is not readily applicable for solving large-scale problems, their model is general and provides several important theoretical insights into the structure of bid prices. Rather than representing the bid prices as stochastic processes, our model (like [1]) utilizes an affine functional form to approximate the value function and develop bid prices. Working in continuous time diminishes the impact of time on the number of variables and the number of constraints, which helps address the shortcomings of previous approaches regarding the problem size. The challenge with the continuous time formulation is replacing the infinite collection of constraints with a finite set and selecting a finite-dimensional family of functions to approximate the bid prices. Applying properties from the optimality conditions of the problem, we construct a scalable computational procedure by representing the derivatives of the bid price functions with cubic splines and reformulating the problem as a second-order cone program. The cubic splines allow the number of the bid price variables to be independent of the size of the booking horizon making it scalable to larger networks.

Our numerical experiments demonstrate that the revenue benefits from discrete time dynamic bid prices dissipate in a continuous time setting as the problem scale increases. In comparison, by developing an approach from continuous time, our method is able to maintain revenue improvements over an LP bid price policy that is updated several thousands of times as the network size increases. By implementing our method on the Porter Airlines network, which consists of 182 flight legs, 2196 products, and 70 seats on each flight, we establish that our continuous time formulation can generate effective time-dependent bid prices on an industrial scale.

The remainder of the paper is organized as follows. In §3.2, we introduce notation, formally define the network RM problem and present the Hamilton-Jacobi-Bellman conditions, which serves as a starting point for our model. In §3.3, we derive our bid price approach directly from the optimal control problem of the network RM problem by introducing two approximations into the problem. The first approximation introduces approximation of the value function using time-dependent bid prices and the second reduces the number of variables and constraints in the problem by an implicit treatment of the accept/reject controls. We establish the optimality conditions and structural properties for the approximate optimal control problem in §3.4. In §3.5m we use the established properties to select a finite-dimensional family of functions of time that can approximate the derivatives of bid prices and reduce the infinite collection of constraints to a finite set. The resulting model is an approximate second-order cone program. Three sets of numerical experiments illustrating our model's effectiveness are presented in §3.6. The first two sets of experiments compare discrete and continuous dynamic bid prices against LP bid prices over a randomly generated network in continuous time. The third set of experiments computes bid prices from our solution method over the Porter Airlines network to demonstrate the potential of our method's performance in an industrial setting. Finally, in §3.7, we provide a brief summary and directions for future research. Please see Online Appendix B.1 for proofs.

3.2. Network Revenue Management Problem in Continuous Time

The capacity control approach to network RM can be described as follows. Consider a firm which has m perishable resources with capacities c_i, $i \in \mathcal{I} \equiv \{1, \ldots, m\}$ which expire at time T. The firm uses these resources to offer n products during the time interval $[0, T]$. The price of each product $j \in \mathcal{J} \equiv \{1, \ldots, n\}$ is fixed at p_j, and demands for the products are described by a collection of independent (possibly non-homogeneous) Poisson processes with continuous intensities $\lambda_{j,t}$. One unit of demand for product $j \in \mathcal{J}$ requires $a_{i,j}$ units of resource $i \in \mathcal{I}$. We assume that the entries of $A = \|a_{i,j}\|$ are 0 or 1 and that of $\mathbf{c} = (c_1, \ldots, c_m)'$ are positive integers. In this case, column A_j of A is an incidence vector of resources required by product j. The described capacity and demand structure is a common assumption in RM models for such applications as airline and hotel RM.

In particular, it is still a common practice among airlines to fix prices $\mathbf{p} = (p_1, \ldots, p_n)'$ and only control whether a particular product is offered at a given time within the selling horizon. Thus, the capacity control process $u_{j,t}$, $j \in \mathcal{J}$ at time t, specifies whether to accept ($u_{j,t} = 1$) or reject ($u_{j,t} = 0$) an incoming demand request for product j. Let $N_t = (N_{1,t}, \ldots, N_{n,t})'$ be a vector of Poisson processes with intensities $(u_{1,t}\lambda_{1,t}, \ldots, u_{n,t}\lambda_{n,t})$ modulated by the controls. Let $\mathcal{U}$ be the class of all non-anticipating control processes $\mathbf{u}_t = (u_{1,t}, \ldots, u_{n,t})'$ which satisfy

$$\int_0^T A dN_t \leq \mathbf{c} \quad \text{(a.s.)} \tag{3.2.1}$$

and $u_{j,t} \in \{0,1\}$, $j \in \mathcal{J}$, $t \in [0,T]$. The expected total revenue of policy $\pi \in \mathcal{U}$ is

$$V^\pi(\mathbf{c}, 0) = E_\pi\left[\int_0^T \mathbf{p}' dN_s\right]$$

The capacity control problem is to find a policy π^* which attains $V^*(\mathbf{c}, 0) = \sup_{\pi\in\mathcal{U}} V^\pi(\mathbf{c}, 0)$. The optimal policy for this intensity control problem is of Markovian form and Theorem VII.1 of [12] provides sufficient Hamilton-Jacobi-Bellman (HJB) conditions. Let $\mathbf{r} = (r_1, \ldots, r_m)$ be the vector of remaining capacities, and $\mathbf{r}_t = \mathbf{c} - AN_t$ be the remaining capacity process. Constraints (3.2.1) are equivalent to the condition that $A\mathbf{u}_t \leq \mathbf{r}_t$ holds almost surely. Let $U(\mathbf{r}) = \{\mathbf{u} \in \{0,1\}^n : A\mathbf{u} \leq \mathbf{r}\}$, and $V^*(\mathbf{r}, t)$ be the value function, i.e. the optimal expected total revenue given capacities $\mathbf{r}$ at time $t \leq T$. Then the HJB conditions are given by the following differential equations

$$\frac{\partial}{\partial t} V^*(\mathbf{r}, t) + \max_{\mathbf{u}\in U(\mathbf{r})} \left\{ \sum_{j\in\mathcal{J}} \lambda_{j,t} u_j \left[p_j - \left(V^*(\mathbf{r}, t) - V^*(\mathbf{r} - A_j, t)\right)\right] \right\} = 0, \tag{3.2.2}$$

for all $t \in [0,T]$ and $\mathbf{r} \in \mathcal{R} = \{\mathbf{r} \in \mathbb{Z}_+^m : \mathbf{r} \leq \mathbf{c}\}$, with the boundary conditions

$$V^*(\mathbf{r}, T) = 0, \quad \forall\, \mathbf{r} \in \mathcal{R}, \text{ and} \tag{3.2.3}$$

$$V^*(\mathbf{0}, t) = 0, \quad \forall\, t \in [0,T]. \tag{3.2.4}$$

We observe that the maximum in (3.2.2) gives the optimal controls of the threshold form

$$u_j^*(\mathbf{r}, t) = \begin{cases} 1, & \text{if } p_j \geq V^*(\mathbf{r}, t) - V^*(\mathbf{r} - A_j, t), \\ 0, & \text{otherwise.} \end{cases} \tag{3.2.5}$$

The policy implies that given sufficient resources it is optimal to fulfill the demand request as long as the revenue received from this request is greater than the opportunity cost of the required resources in terms of the expected revenues from future sales. On substituting optimal controls (3.2.5) into (3.2.2), we get differential equations of the form

$$\frac{\partial}{\partial t} V^*(\mathbf{r}, t) + \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \lambda_{j,t} \max \left\{ p_j - \big(V^*(\mathbf{r}, t) - V^*(\mathbf{r} - A_j, t) \big), 0 \right\} = 0. \tag{3.2.6}$$

This a system of ordinary differential equations for $V^*(\cdot, t)$ (treated as an $|\mathcal{R}|$-dimensional vector) with continuous dependence on t and Lipschitz-continuous on $V^*(\cdot, t)$. A known result from the theory of differential equations (see Corollary 2.4.5 of [66]) implies that there exists a unique solution, in the Sobolev space $W^{1,1}[0, T]$. The functions in this space are almost everywhere differentiable on $[0, T]$, Lebesgue integrable along with their derivatives (belong to $L^1[0, T]$), and equal to the integrals of their derivatives (absolutely continuous). The norm in $W^{1,1}[0, T]$ is the sum of $L^1[0, T]$ norms (the integral of the absolute value) of the function and its derivative. Unfortunately, the exact solution of (3.2.6) cannot be computed in most practical problems because of the extremely high number of dimensions of $V^*(\cdot, t)$. In the next section, we discuss approximations to $V^*(\mathbf{r}, t)$ which can be obtained with reasonable computational effort.

3.3. Approximate Optimal Control Problem

We start by contrasting our approach to approximating the network capacity control problem with that of [1], who starts with a discrete time version of the problem where it is assumed that the time unit is sufficiently small for $\sum_{j \in \mathcal{J}} \lambda_{j,t} \leq 1$ to hold and for the intensities $\lambda_{j,t}$ to be interpreted as probabilities of booking request arrivals in period t. A discrete time analogue of (3.2.2) is

$$V^{DT}(\mathbf{r}, t) = V^{DT}(\mathbf{r}, t{+}1) + \max_{\mathbf{u} \in U(\mathbf{r})} \left\{ \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \lambda_{j,t} u_j \left[p_j - \left(V^{DT}(\mathbf{r}, t{+}1) - V^{DT}(\mathbf{r} - A^j, t{+}1) \right) \right] \right\},$$
$$\forall\ t = 0, \ldots, T-1,\ \mathbf{r} \in \mathcal{R},$$

with the boundary conditions of the same form as (3.2.3)-(3.2.4). The discrete time HJB conditions can be equivalently restated as a linear program of the form

$$\begin{aligned}
\min \quad & \tilde{V}^{DT}(\mathbf{c}, 0) \\
\text{s.t.} \quad & \tilde{V}^{DT}(\mathbf{r}, t) \geq \tilde{V}^{DT}(\mathbf{r}, t+1) + \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \lambda_{j,t} u_j \left[p_j - \left(\tilde{V}^{DT}(\mathbf{r}, t+1) - \tilde{V}^{DT}(\mathbf{r} - A^j, t+1) \right) \right], \\
& \quad \forall\ t = 0, \ldots, T-1,\ \mathbf{r} \in \mathcal{R}, \mathbf{u} \in U(\mathbf{r}) \\
& \tilde{V}^{DT}(\mathbf{r}, T) = 0, \quad \forall\ \mathbf{r} \in \mathcal{R}, \\
& \tilde{V}^{DT}(\mathbf{0}, t) = 0, \quad \forall\ t = 1, \ldots, T-1,
\end{aligned}$$

where $\tilde{V}^{DT}(\mathbf{r}, t)$'s are the variables. This linear program is still intractable for instances of realistic size because of the extremely large number of variables. Therefore, [1] restricts the values of these variables to a subspace represented by a linear combination of a certain collection of basis functions. A particularly tractable case is provided by the linear function of the capacity vector $\tilde{V}^{DT}(\mathbf{r}, t) = v_{0,t} + \sum_{i\in\mathcal{I}} v_{i,t} r_i$. The values $v_{i,t}$ have an appealing interpretation of the *dynamic bid prices*. A dynamic control policy resulting from this representation reduces to accepting a booking request if and only if its fare p_j is greater or equal to the sum of bid prices of the required resources $\sum_{i\in A_j} v_{i,t}$ at the time t of the request arrival. With a slight abuse of notation, $i \in A_j$ represents all resources used by product j.

Our approach differs from that of [1] by the use of continuous time methods in approximating the value function as well as the use of a different starting point – equation (3.2.6) rather than (3.2.2). Consider a variational problem of the form

$$\min \quad \tilde{V}(\mathbf{c}, 0) \tag{3.3.1}$$

$$\text{s.t.} \quad \frac{\partial}{\partial t}\tilde{V}(\mathbf{r}, t) + \sum_{j\in\mathcal{J}:A_j\leq\mathbf{r}} \lambda_{j,t} \max\left\{p_j - (\tilde{V}(\mathbf{r}, t) - \tilde{V}(\mathbf{r} - A_j, t)), 0\right\} \leq 0, \tag{3.3.2}$$

$$\text{a.e. } t \in [0, T], \ \forall\ \mathbf{r} \in \mathcal{R},$$

$$\tilde{V}(\mathbf{r}, T) = 0, \quad \forall\ \mathbf{r} \in \mathcal{R}, \tag{3.3.3}$$

and trajectories $\tilde{V}(\cdot, t)$ in the space $W^{1,1}[0, T]$, where “a.e.” stands for “almost everywhere” in the Lebesgue measure. Constraint (3.3.2) is a differential inequality that limits the derivatives of $\tilde{V}(\cdot, t)$ from above. Constraints (3.3.2)-(3.3.3) used with $\mathbf{r} = 0$ imply that $\tilde{V}(\mathbf{0}, t) \geq 0$ for all $t \in [0, T]$. This removes the need to explicitly impose the equivalent of boundary condition (3.2.4) in the variational problem. Intuitively, the value of $\tilde{V}(\mathbf{c}, 0)$ is the smallest when all inequalities are satisfied as equalities. As the result, the optimal value of $\tilde{V}(\mathbf{c}, 0)$ is equal to $V^*(\mathbf{c}, 0)$. The proof of the following formal statement is straightforward:

Lemma 2. $V^*(\cdot, t)$ is the optimal solution to (3.3.1)-(3.3.3).

The idea of the proof is to observe that $V^*(\cdot, 0)$ is a feasible trajectory for (3.3.1)-(3.3.3), and any feasible solution to (3.3.1)-(3.3.3) has a property $\tilde{V}(\mathbf{r}, t) \geq V^*(\mathbf{r}, t)$ for all $\mathbf{r} \in \mathcal{R}$ and $t \in [0, T]$.

We now consider a *restriction* of the problem (3.3.1)-(3.3.3), where:

1. The feasible trajectories are restricted to a subspace represented as

$$\tilde{V}(\mathbf{r}, t) = v_{0,t} + \sum_{i \in \mathcal{I}} v_{i,t} r_i.$$

2. The max operator is approximated from above by the following lemma (the proof is immediate):

Lemma 3. For any g, $\frac{1}{2}\big((g^2 + \epsilon^2)^{\frac{1}{2}} + g\big) \to \max\{g, 0\}$ from above uniformly as $\epsilon \to 0$. Moreover,

$$0 \leq \frac{1}{2}\big((g^2 + \epsilon^2)^{\frac{1}{2}} + g\big) - \max\{g, 0\} \leq \frac{\epsilon}{2}$$

where the upper bound is attained for $g = 0$.

Constraint (3.3.2) uses terms of the form $\max\{g_{j,t}, 0\}$ where

$$g_{j,t} = p_j - \big(\tilde{V}(\mathbf{r}, t) - \tilde{V}(\mathbf{r} - A_j, t)\big) = p_j - \sum_{i \in A_j} v_{i,t}$$

represents a profit estimate from product j based on the bid prices at time t. We replace $\max\{g_{j,t}, 0\}$ with the expression

$$M_j^{\epsilon}(\mathbf{v}_t) = \frac{1}{2}\Big\{\Big(\Big[p_j - \textstyle\sum_{i \in A_j} v_{i,t}\Big]^2 + \epsilon^2\Big)^{\frac{1}{2}} + p_j - \textstyle\sum_{i \in A_j} v_{i,t}\Big\},$$

which approximates it from above within $\frac{\epsilon}{2}$. The resulting approximate optimal control problem (AOCP) is

$$\min \quad v_{0,0} + \sum_{i\in\mathcal{I}} v_{i,0}c_i \tag{3.3.4}$$

$$\text{s.t.} \quad \dot{v}_{0,t} + \sum_{i\in\mathcal{I}} r_i \dot{v}_{i,t} + \sum_{j\in\mathcal{J}:A_j\le\mathbf{r}} \lambda_{j,t} M_j^{\epsilon}(\mathbf{v}_t) \le 0, \text{ a.e. } t\in[0,T],\ \forall \mathbf{r}\in\mathcal{R}, \tag{3.3.5}$$

$$v_{i,T} = 0,\ \forall\ i\in\mathcal{I}\cup\{0\}. \tag{3.3.6}$$

Since AOCP results from restricting the feasible set of the problem (3.3.1)-(3.3.3), we conclude the following:

Lemma 4. The optimal value of AOCP is an upper bound for $V^*(\mathbf{c},0)$.

Although AOCP belongs to the general class of control problems for unbounded differential inclusions, we are able to derive rather simple optimality conditions for it using results of [46]. The inclusion is unbounded since $\dot{v}_{i,t}$ on a feasible state trajectory can be arbitrarily low.

3.4. Optimality Conditions for AOCP

For convenience, we let the function of t, $\mathbf{v}_t$ and $\dot{\mathbf{v}}_t$ representing the left-hand-side of (3.3.5) be denoted as $f_{\mathbf{r}}(t,\mathbf{v}_t,\dot{\mathbf{v}}_t)$ and its last term, which does not depend on $\dot{\mathbf{v}}_t$, as $f^0_{\mathbf{r}}(t,\mathbf{v}_t)$. Because of its additive structure, $f^0_{\mathbf{r}}(t,\mathbf{v}_t)$ can be expressed as $f^0_{\mathbf{r}}(t,\mathbf{v}_t) = \sum_{j\in\mathcal{J}:A_j\le\mathbf{r}} \lambda_{j,t} M_j^{\epsilon}(\mathbf{v}_t)$. We use this representation to make the statement of the optimality conditions more compact:

Theorem 1. If $\mathbf{v}^*_t = (v^*_{i,t},\ i\in\mathcal{I}\cup\{0\})$ is an optimal bid price trajectory for the problem (3.3.4)-(3.3.6), then there exist an adjoint trajectory $\mathbf{z}_t = (z_{i,t},\ i\in\mathcal{I}\cup\{0\})$ and multipliers $\mu_{\mathbf{r},t}\ge 0$, $\mathbf{r}\in\mathcal{R}$ for almost all t such that the adjoint differential inclusion (the Euler-Lagrange inclusion) holds for almost all t:

$$\dot{z}_{0,t} = 0, \tag{3.4.1}$$

$$\dot{z}_{i,t} = \sum_{\mathbf{r}\in\mathcal{R}} \mu_{r,t} \sum_{j:A_j\leq\mathbf{r},i\in A_j} \lambda_{j,t} \frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t^*), \quad i \in \mathcal{I}, \tag{3.4.2}$$

$$z_{0,t} = \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t}, \tag{3.4.3}$$

$$z_{i,t} = \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} r_i, \quad i \in \mathcal{I}, \tag{3.4.4}$$

$$\mu_{\mathbf{r},t} = 0, \quad \mathbf{r} \notin B_t, \tag{3.4.5}$$

where $B_t = \{\mathbf{r} \in \mathcal{R} : f_{\mathbf{r}}(t, \mathbf{v}_t^*, \dot{\mathbf{v}}_t^*) = 0\}$ is the set of active constraints in (3.3.5); the Weierstrass-Pontryagin maximum condition holds for almost all t:

$$\sum_{i\in\mathcal{I}\cup\{0\}} z_{i,t}\dot{v}_{i,t}^* = \max \sum_{i\in\mathcal{I}\cup\{0\}} z_{i,t} w_i \tag{3.4.6}$$

$$\text{s.t.} \quad w_0 + \sum_{i\in\mathcal{I}} r_i w_i + f_{\mathbf{r}}^0(t, \mathbf{v}_t^*) \leq 0, \quad \mathbf{r} \in \mathcal{R}; \tag{3.4.7}$$

and the transversality condition holds

$$z_{0,0} = 1, \tag{3.4.8}$$

$$z_{i,0} = c_i, \quad i \in \mathcal{I}. \tag{3.4.9}$$

Moreover, collection $\{\mu_{\mathbf{r},t}, \ \mathbf{r} \in \mathcal{R}\}$ is an optimal solution to the dual of the LP problem in (3.4.6)-(3.4.7).

The optimality conditions of Theorem 1 are essential for further analysis of the optimal solution to AOCP. In particular, we obtain the following monotonic property:

Corollary 2. The adjoint trajectory has the following properties: $z_{0,t}$ is constant and equals 1, i.e.

$$z_{0,t} = \sum_{\mathbf{r} \in \mathcal{R}} \mu_{\mathbf{r},t} = 1,$$

and $\dot{z}_{i,t} \leq 0$, $i \in \mathcal{I}$ for almost all t with a strict inequality unless $z_{i,t} = 0$.

The second property asserted in Corollary 2 means that $z_{i,t}$, $i \in \mathcal{I}$ is strictly decreasing unless it is zero. This corollary also suggests an interpretation to the adjoint variables $z_{i,t}$ and multipliers $\mu_{r,t}$. Since $\mu_{r,t}$ are nonnegative and add to 1 they can be interpreted as probability distribution over all possible capacity vectors. Strictly positive $\mu_{r,t}$'s identify active constraints, and their values represent relative contributions of the corresponding constraints to the optimal objective value. Value of $z_{i,t}$ is the expected value of the ith component of the capacity vector over this distribution. Additional insights are revealed from the form of adjoint equation (3.4.2). From the proof of Corollary 2, we see that the value of $-\frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t)$ is zero for $i \notin A_j$, identical for $i \in A_j$ and belongs to the interval $[0, 1]$. Moreover, it approximates an indicator function of the event that the revenue from product j exceed the sum of bid prices of resources in A_j. Since the latter is the acceptance rule for demand requests, the negative of the inner summation in (3.4.2) approximates the intensity of utilization of resource i when capacities are given by $\mathbf{r}$. Therefore, $\dot{z}_{i,t}$ decreases approximately at the rate of expected intensity of resource i utilization where the expectation is with respect to probability mass $\mu_{r,t}$. According to a general control-theoretic interpretation, adjoint variables play the role of shadow prices by measuring the impact on the value function of a unit change in the state variables on the optimal trajectory of the system. Applying this general interpretation to our problem, we conclude that $z_{i,t}$ measures, on the optimal bid-price trajectory, the impact of a unit change in bid price $v_{i,t}$ on the value of the upper bound. Corollary 2 and equation (3.4.2) imply that this impact is strictly decreasing at the expected rate of resource utilization until it reaches zero and stays constant after that. The second result is a monotonicity property of the bid prices:

Corollary 3. The optimal bid price trajectory of AOCP is decreasing and nonnegative in every component $v^*_{i,t}$, $i \in \mathcal{I} \cup \{0\}$. Moreover, bid price $v_{i,t}$, $i \in \mathcal{I}$ remains constant from time 0 until t'_i such that $z_{i,t'_i} = 1$.

In addition to the managerial implication that dynamic bid prices produced by the continuous time model decrease over time, this result reveals that each bid price remains constant for some period of time starting from the beginning of the planning horizon. The length of these intervals is determined from the adjoint variables. Corollary 3 also suggests a way to eliminate a bulk of constraints from the problem. Consider any decreasing bid price trajectory $\mathbf{v}_t$. Observe that $f^0_{\mathbf{r}+\mathbf{e}_i}(t, \mathbf{v}_t) = f^0_{\mathbf{r}'}(t, \mathbf{v}_t)$ whenever the set of depleted resources in $\mathbf{r}$ and $\mathbf{r}'$ is the same. If $\mathbf{r}' \geq \mathbf{r}$ we have $f_{\mathbf{r}+\mathbf{e}_i}(t, \mathbf{v}_t, \dot{\mathbf{v}}_t) \geq f_{\mathbf{r}'}(t, \mathbf{v}_t, \dot{\mathbf{v}}_t)$ because $\dot{\mathbf{v}}_t \leq \mathbf{0}$ and $\dot{\mathbf{v}}_t$ enters into these functions with nonnegative coefficients. This immediately implies the following

Corollary 4. If we modify the problem (3.3.4)-(3.3.6) by adding a constraint that all state components $v_{i,t}$, $i \in \mathcal{I} \cup \{0\}$ are decreasing and removing all constraints in (3.3.5) corresponding to $\mathbf{r}$ which have some component greater than 1, then the resulting problem has the same optimal solution as (3.3.4)-(3.3.6).

The resulting reduced set of constraints is indexed by the set $\bar{\mathcal{R}} = \{\mathbf{r} \in \mathcal{R} : r_i \leq 1,\ i \in \mathcal{I}\}$. We refer to the modified AOCP problem where a feasible bid price trajectory is constrained to be monotone and the constraint index set $\mathcal{R}$ is replaced by $\bar{\mathcal{R}}$ as AOCPM.

Corollaries 2-4 parallel findings in [1]. The interpretation of $\mu_{r,t}$ from Corollary 2 as a probability distribution is analogous to the interpretation of the dual variables in the DLP as state action probabilities. The monotonicity property of the time-dependent bid prices from Corollary 3 was established for DLP in Theorem 2 of [1]. One advantage of the continuous time formulation that transpires in Corollary 3 is a relatively easy-to-prove property of constant bid prices at the beginning of the planning horizon. The motivation for adding the monotonicity constraints in Corollary 4 stems from [1]'s numerical experiments, which demonstrated that these additional constraints vastly improve the solution speed of the problem. In addition, [1] proves that the objective value of the DLP is bounded from above by the static LP. We show that AOCP has a similar property up to an additive term which is proportional to the accuracy ϵ of approximation for the max operator. The static LP is a problem which maximizes the total expected revenues from all products subject to the constraints that the expected sales are less than the expected demand, and that the network has sufficient capacity. If Y_j is the expected number of seats sold for product j then the static LP is

$$V^{LP}(\mathbf{c}, 0) = \max_{\mathbf{Y}} \sum_{j} p_j Y_j \tag{3.4.10}$$

$$\text{s.t.} \sum_{j: i \in A_j} Y_j \leq c_i, \quad \forall\, i \in \mathcal{I} \tag{3.4.11}$$

$$0 \leq Y_j \leq \int_0^T \lambda_{j,t} dt, \quad \forall\, j \in \mathcal{J}. \tag{3.4.12}$$

The following proposition is established by constructing a feasible solution to the static LP problem (3.4.10)-(3.4.12) from an optimal solution to AOCP.

Proposition 6. An optimal solution to AOCP yields a feasible solution to the standard LP and its value is bounded from above by $V^{LP}(\mathbf{c}, 0) + \frac{\epsilon}{2}\int_0^T \sum_{j\in\mathcal{J}} \lambda_{j,t} dt$.

The proposition shows that the bound on $V^*(\mathbf{c}, 0)$ becomes no worse than $V^{LP}(\mathbf{c}, 0)$ as ϵ goes to zero. From the proof, it is also evident that the actual bound provided by AOCP may be even better. However, the main advantage of AOCP is not in the bound it provides but in the continuous time computational approaches to finding dynamic bid prices.

Remark 1. The theorems and lemmas presented in this section assume the existence of an optimal solution to AOCP. Augmenting AOCP with a lower bound constraint on $\dot{v}_{i,t}$, $i \in \mathcal{I} \cup 0$ places it into a general class of the optimal control problems P discussed in §2.6 of [66]. Proposition 2.6.2 in [66] states that P has a minimizer. With the additional bound on $\dot{v}_{i,t}$, $i \in \mathcal{I} \cup 0$, all assumptions of Proposition 2.6.2 [66] are satisfied and consequently, there exists an optimal solution. For this bounded AOCP, the Euler-Lagrange inclusion is similar to that of Theorem 1 except (3.4.3) and (3.4.4) become respectively:

$$z_{0,t} = \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} - \xi_{0,t}$$

$$z_{i,t} = \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} r_i - \xi_{i,t}, \quad i \in \mathcal{I},$$

where $\xi_{i,t} \geq 0$ play the role of Lagrange multipliers for the boundary constraints. The modified optimality conditions can be used to establish structural results for the bounded AOCP. The choice of the lower bound on $\dot{v}_{i,t}$ should not a priori exclude any vertices of the constraint set (3.4.7). One such option is $\dot{v}_{i,t} \geq -f^0_{\mathbf{e}}(t, \mathbf{0})$ (where $\mathbf{e}$ is the m-dimensional vector of ones). No vertex is excluded because $f^0_{\mathbf{e}}(t, \mathbf{0}) \geq f^0_{\mathbf{r}}(t, \mathbf{v})$

for any $\mathbf{v} \geq 0$, $\mathbf{r} \in \mathcal{R}$.

3.5. Finite-Dimensional Computational Strategies for AOCP

In this section we discuss finite-dimensional computational strategies for solving AOCP. There are several key points that need to be taken into account. First, we need to select an appropriate family of functions of time that can approximate feasible state trajectories $\mathbf{v}_t$, $t \in [0, T]$ of the AOCP. This family must be described by a finite number of parameters. Second, we need to devise a constraint generation strategy that permits us to replace an infinite collection of constraints in differential inclusion (3.3.5) with a finite set. The two points are related, because the continuity and differentiability properties of functions in the family may affect the number of constraints that have to be explicitly considered. The nature of relation follows from constraints (3.3.5) that involve both the bid price trajectory and its derivative. Indeed, if a bid price trajectory satisfies constraints (3.3.5) only for a given finite number of time points, then the continuity of the trajectory and its derivative permits us to reduce constraint violations for other time points. The proposed computational strategy addresses both of these points and leads to a finite-dimensional linear program with second-order cone constraints.

3.5.1 Reduction to a Finite-Dimensional Problem

Bid-price trajectories $\mathbf{v}_t$, $t \in [0, T]$ of the AOCP belong to the normed space $W^{1,1}[0, T]$ of a.e. differentiable functions. Since constraints of AOCP involve both $\mathbf{v}_t$ and its derivative $\dot{\mathbf{v}}_t$, we need to develop a computationally tractable representation for either object. It is sufficient to start with the derivative. Indeed, if $\dot{\mathbf{v}}_t$ is approximated in terms of L_1-norm within a sufficiently small tolerance by an integrable function $\dot{\mathbf{v}}_t^A$, then $\mathbf{v}_t$ is approximated by $\mathbf{v}_t^A = -\int_t^T \dot{\mathbf{v}}_t^A dt$ within a small tolerance in the stronger pointwise sense (L_∞-norm). The objective function for the approximate problem is:

$$v_{0,0}^A + \sum_{i \in \mathcal{I}} v_{i,0}^A c_i. \tag{3.5.1}$$

We recall that any function in $L_1[0,T]$ can be approximated by a twice continuously differentiable function (see, for example, Theorem 2.16 from [42] claiming a stronger result about approximation by infinitely differentiable functions). Within the class of such functions, we focus on the cubic splines:

$$\dot{v}_{i,t}^A = -\sum_{l=0}^{3} a_{i,l}^k \left(\frac{t - t_k}{t_{k+1} - t_k} \right)^l, \forall\, t \in [t_k, t_{k+1}], \tag{3.5.2}$$

where t_k, $k \in \mathcal{K} \equiv \{1, \dots, K\}$ is a collection of appropriate *knot points* and $a_{i,l}^k$, $l \in \mathcal{L} \equiv \{0, \dots, 3\}$ are the coefficients of the spline on the interval $[t_k, t_{k+1})$ for resource $i \in \mathcal{I}$. Continuity of this approximation and its derivatives is ensured by the constraints:

$$a_{i,0}^{k+1} = \sum_{l=0}^{3} a_{i,l}^k, \tag{3.5.3}$$

$$\frac{1}{t_{k+2} - t_{k+1}} a_{i,1}^{k+1} = \sum_{l=1}^{3} \frac{l}{t_{k+1} - t_k} a_{i,l}^k, \tag{3.5.4}$$

$$\frac{2}{(t_{k+2} - t_{k+1})^2} a_{i,2}^{k+1} = \sum_{l=2}^{3} \frac{l(l-1)}{(t_{k+1} - t_k)^2} a_{i,l}^k, \tag{3.5.5}$$

that apply to every $k = 1, \ldots, K-1$ and $i \in \mathcal{I}$. The value of $\mathbf{v}_{i,t}^A$ is found as the integral of (3.5.2) on the interval $[t, T]$

$$v_{i,t}^A = -\int_t^T \dot{v}_{i,t}^A dt \tag{3.5.6}$$

which is a linear expression of the variables $a_{i,l}^k$, $k \in \mathcal{K}$, $l \in \mathcal{L}$.

Since we already know that AOCP is equivalent to AOCPM, it makes sense to restrict $\dot{\mathbf{v}}_{i,t}^A$ to be nonpositive or, equivalently, the cubic spline represented by $a_{i,l}^k$'s to be nonnegative. To enforce nonnegativity of the cubic splines, we use a method that was recently employed by [6] in the context of statistical estimation of the arrival rate using cubic splines. Applying the characterization of nonnegative functions in Tchebysheff systems (see [34]) to cubic splines over a finite set, Alizadeh et al. enforce nonnegativity using semidefinite matrices. For each spline component index $k \in \mathcal{K}$ and $i \in \mathcal{I}$, we use Theorem 1 of [6] claiming that the necessary and sufficient conditions for nonnegativity of cubic polynomial $\sum_{l=0}^{3} a_{i,l}^k \left(\frac{t-t_k}{t_{k+1}-t_k}\right)^l$ are provided by the following representation of the coefficients:

$$a_{i,0}^k = y_{i,0}^k, \tag{3.5.7}$$

$$a_{i,1}^k = 2y_{i,1}^k + x_{i,0}^k - y_{i,0}^k, \tag{3.5.8}$$

$$a_{i,2}^k = y_{i,2}^k + 2x_{i,1}^k - 2y_{i,1}^k, \tag{3.5.9}$$

$$a_{i,3}^k = x_{i,2}^k - y_{i,2}^k, \tag{3.5.10}$$

$$\frac{x_{i,0}^k + x_{i,2}^k}{2} \geq \sqrt{\left(\frac{x_{i,0}^k - x_{i,2}^k}{2}\right)^2 + (x_{i,1}^k)^2}, \tag{3.5.11}$$

$$\frac{y_{i,0}^k + y_{i,2}^k}{2} \geq \sqrt{\left(\frac{y_{i,0}^k - y_{i,2}^k}{2}\right)^2 + (y_{i,1}^k)^2}. \tag{3.5.12}$$

Constraints (3.5.11)-(3.5.12) are equivalent to restricting matrices $X_i^k = \begin{pmatrix} x_{i,0}^k & x_{i,1}^k \\ x_{i,1}^k & x_{i,2}^k \end{pmatrix}$ and $Y_i^k = \begin{pmatrix} y_{i,0}^k & y_{i,1}^k \\ y_{i,1}^k & y_{i,2}^k \end{pmatrix}$ to being positive semidefinite. This type of constraint is well studied in the area of optimization known as semidefinite programming (SDP). Together, constraints (3.5.3)-(3.5.12) enforce monotonicity of bid prices.

In a practical implementation, variables $a_{i,l}^k$ can be eliminated by substituting representation (3.5.7)-(3.5.10) into (3.5.2) and (3.5.6). In fact, for each $i \in \mathcal{I} \cup 0$ and time interval $[t_k, t_{k+1}]$ only six variables are needed to represent dynamic bid price $v_{i,t}^A$. The number of variables is effectively independent of time and depends on the number of resources and knot points. Moreover, after the substitution, the monotonicity of bid price $v_{i,t}^A$ on the entire interval $[t_k, t_{k+1}]$ is enforced by just two constraints (3.5.11)-(3.5.12). Given that the proposed approximation of the derivative $\dot{\mathbf{v}}_t^A$ requires a finite number of variables and the objective of AOCPM is linear, our problem reduces to a semi-infinite optimization problem (SIP). There are several standard approaches for solving SIP (see [47] for a recent review). Here, we construct a finite problem by restricting constraint collection (3.3.5) to a set of grid-points $\mathcal{T} \subset [0, T]$. We provide a discussion on the resulting error bound of limiting the SIP to a set $\mathcal{T}$ consisting of N_k evenly spaced grid-points over $[t_k, t_{k+1}]$, $k \in \mathcal{K}$ in the following section. The error bound can be used to check the approximation accuracy after solving for the spline coefficients through the ASOCP. In the subsequent development, we assume that $\mathcal{T}$ contains the set of spline knot points t_k, $k \in \mathcal{K}$ and $t_{K+1} = T$.

We complete the construction of the approximate problem by observing that constraints (3.3.5) restricted to $t \in \mathcal{T}$ can be converted to a second-order conic constraints by introducing variables $g_{j,t}$ and $h_{j,t}$, $j \in \mathcal{J}$, $t \in \mathcal{T}$ such that

$$h_{j,t} \geq \sqrt{g_{j,t}^2 + \epsilon^2}, \tag{3.5.13}$$

$$g_{j,t} = p_j - \sum_{i \in A_j} v_{i,t}^A. \tag{3.5.14}$$

Constraints (3.5.11), (3.5.12), and (3.5.13) equivalently require that the vectors

$$\left(\frac{x_{i,0}+x_{i,2}}{2}, \frac{x_{i,0}-x_{i,2}}{2}, x_{i,1}\right), \quad \left(\frac{y_{i,0}+y_{i,2}}{2}, \frac{y_{i,0}-y_{i,2}}{2}, y_{i,1}\right), \quad \text{and} \quad \left(h_{j,t}, g_{j,t}, \epsilon\right)$$

belong to the three-dimensional second-order (Lorentz) cone. Each constraint of the form (3.3.5) is replaced by

$$\dot{v}_{0,t}^A + \sum_{i \in \mathcal{I}} r_i \dot{v}_{i,t}^A + \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \lambda_{j,t} \frac{1}{2}(h_{j,t} + g_{j,t}) \leq 0 \tag{3.5.15}$$

for each $t \in \mathcal{T}$ and $\mathbf{r} \in \bar{\mathcal{R}}$. The maximum slack (equivalently, the minimum violation) in this constraint is obtained when $h_{j,t}$ is as small as possible, forcing the equality in (3.5.13).

The resulting approximate SOCP problem (ASOCP) is to minimize the linear objective (3.5.1) with respect to the variables X_i^k, Y_i^k, $a_{i,l}^k$, $\dot{v}_{i,t}^A$, $v_{i,t}^A$, $g_{j,t}$ and $h_{j,t}$ where the indices range over the sets $i \in \mathcal{I} \cup \{0\}$, $j \in \mathcal{J}$, $t \in \mathcal{T}$, $k \in \mathcal{K}$, $l \in \mathcal{L}$ subject to the spline representation constraints (3.5.2)-(3.5.10) for the trajectory and its derivative, the second-order cone constraints (3.5.11)-(3.5.13), as well as (3.5.14) and the modified differential inclusion constraints (3.5.15). All of the constraints in ASOCP, except those of the conic types, are linear. Although we use the SOCP representation to impose nonnegativity on the derivatives of the spline representation of the bid price trajectory, nonnegativity could have been enforced using B-Splines. However, as discussed by [6], B-Splines enforce constraints that are tighter than functional nonnegativity, potentially reducing the accuracy of approximation. ASOCP belongs to a very well studied class of SDP/SOCP problems (see, for example, [70, 7]), and there are efficient software packages that can solve fairly large-scale instances. A potential inefficiency in application of these solvers to ASOCP is the number of constraints in (3.5.15). Indeed, there is a constraint for every time instance in $\mathcal{T}$ and every capacity vector in $\bar{\mathcal{R}}$. We can keep the size of $\mathcal{T}$ fairly small, but the cardinality of $\bar{\mathcal{R}}$ is exponential in the number of resources: $|\bar{\mathcal{R}}| = 2^m$. Fortunately, we do not need all of these constraints to obtain highly effective dynamic bid prices. Next, we discuss a constraint generation procedure that can efficiently sample elements of $\bar{\mathcal{R}}$ corresponding to the most violated constraints.

3.5.2 Grid Size Selection

In this section we apply the convergence results on the error between the semi-infinite optimization problem (SIP) and discretized program presented in [60] to select an efficient grid size $\mathcal{T}$ for the ASOCP. The results demonstrate that $\mathcal{T}$ is determined by the order of magnitude of the derivatives of the bid prices as well as the parameter ϵ. Before proceeding to the discussion on the grid size selection, we summarize the main theorem and assumptions from [60].

Convergence for discretization of a semi-infinite program

The SIP is defined as

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in F = \{x \in \mathbb{R}^n | g(x, y) \leq 0, y \in Y\}, \end{aligned}$$

where f and g are real valued functions and Y is a compact infinite index set in $\mathbb{R}^m$. Given a finite grid Y_d, where $Y_d \subset Y$, the discretized problem of SIP is SIP(Y_d):

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in F(Y_d) = \{x \in \mathbb{R}^n | g(x, y) \leq 0, y \in Y_d\}, \end{aligned}$$

The fineness of the grid in Y_d is measured by the distance of its largest mesh, i.e. $d = \max_{y \in Y} \min_{\hat{y} \in Y_d} ||\hat{y} - y||$. Consider the following assumptions on the structure of g, Y, and Y_d.

1. There is a neighborhood $\bar{U}$ of $\bar{x}$ such that the function $\frac{\partial^2 g(x,y)}{\partial y^2}$ is continuous on $\bar{U} \times Y$.

2. The index set $Y \subset \mathbb{R}^m$ is compact, nonempty and given as a finite set of inequalities $Y = \{y \in \mathbb{R}^m | \zeta_i(y) \leq 0, i \in \mathcal{I}\}$, where ζ_i is two times continuously differentiable with respect to y.

3. The Mangasarian Fromovitz Constraint Qualification (MFCQ) holds for Y. This implies that for any $\bar{y} \in Y$ with the active index set $I(\bar{y}) := \{i \in I | \zeta_i(\bar{y}) = 0\}$ there exists a vector $\bar{\eta} = \eta(\bar{y})$ such that $\frac{\partial \zeta_i(\bar{y})}{\partial y} \bar{\eta} < 0, \; i \in I(\bar{y})$.

4. The grid Y_d is chosen such that for all d, $\max_{y \in S_l} \min_{\hat{y} \in Y_d \cap S_l} ||\hat{y} - y|| \leq d \; \forall l = 1, ..., k$, where S_l is a subset of Y defined as $\{y \in Y | \zeta_i = 0\}$ and $S_l \cap Y \neq \emptyset, l = 0, ..., k$.

Since there exists $\hat{y}_d \in S_l \cup Y_d$ such that $||\hat{y}_d - y|| \leq d$, if assumptions 1-5 hold, then Theorem 1 of [60] states that for some $L > 0$, a set of multipliers χ, and small d:

$$\begin{aligned} \max_{y \in Y} g(x_d, y) &\leq \frac{1}{2}(y_d - \hat{y}_d)^T \left(\frac{\partial^2 g(x_d, y_d)}{\partial y^2} - \sum_{i \in I_l} \chi_{d,i} \frac{\partial^2 \eta_i(y_d)}{\partial y^2} \right) (y_d - \hat{y}_d) + o(||y_d - \hat{y}_d||^2) \\ &\leq L||y_d - \hat{y}_d||^2 \\ &\leq Ld^2 \end{aligned}$$

Discretization Error of the ASOCP

Recall that by introducing the spline approximation of the derivative $\dot{\mathbf{v}}_t^A$, the AOCPM reduces to a SIP with the objective to minimize (3.5.1) subject to constraints (3.5.2)-(3.5.12) and (3.3.5). In this section we present a bound on discretizing time to a set of grid-points $\mathcal{T}$ and in the next section we provide a bound on the error for restricting the set of capacity vectors. The potential unboundedness of $\dot{\mathbf{v}}_t^A$ prevents us from selecting the grid-points $\mathcal{T}$ a priori and guaranteeing a pre-specified error tolerance. However, we can use the results of [60] to evaluate the discretization error of a given solution to the ASOCP.

Consider the bid-price trajectory $\mathbf{v}_t^M$, which is the solution to problem (3.5.1)-(3.5.12) and (3.3.5) restricted to $t \in \mathcal{T}$, and its associated spline coefficients $\mathbf{a}_i^k$, $k \in \mathcal{K}$. The set $\mathcal{T}$ has N_k evenly spaced grid-points over each interval $[t_k, t_{k+1}]$, so the fineness of the discretization over mesh k is measured by $d_k = \frac{1}{N_k}$. Using $\mathbf{a}_i^k$, we define

$$\dot{v}_{i,k}^B = \sum_{l=0}^{3} |a_{i,l}^k|, \qquad \ddot{v}_{i,k}^B = \sum_{l=1}^{3} |a_{i,l}^k| \frac{l}{t_{k+1} - t_k}, \qquad \dddot{v}_{i,k}^B = \sum_{l=2}^{3} |a_{i,l}^k| \frac{l(l-1)}{(t_{k+1} - t_k)^2},$$

which act as bounds on the absolute value of the first, second, and third derivative of the bid price trajectory $v_{i,t}^M$, $t \in [t_k, t_{k+1}]$. Defining the bound on the absolute values of $\lambda_{t,j}$ and its first and second derivatives over each spline interval by $\lambda_{j,k}^B$, $\dot{\lambda}_{j,k}^B$, and $\ddot{\lambda}_{j,k}^B$, we claim

Proposition 7. If $\mathbf{v}_t^M$ is the bid price trajectory for the discretized SIP with grid-points $\mathcal{T}$ and $\mathbf{v}^{A*}$ is the optimal bid price trajectory to the SIP, then

$$V^M(\mathbf{c}, 0) \leq V^{A*}(\mathbf{c}, 0) + \sum_{k=1}^{K} L_k d_k^2,$$

where

$$L_k = \frac{1}{2}\Bigg[\sum_{i \in \mathcal{I} \cup \{0\}} \dddot{v}_{i,k}^B + \sum_{j \in \mathcal{J}} \left(\lambda_{j,k}^B \sum_{i \in A_j} \left(\frac{\left(\dot{v}_{i,k}^B\right)^2}{2\epsilon} - \ddot{v}_{i,k}^B \right) + 2\dot{\lambda}_{j,k}^B \sum_{i \in A_j} \dot{v}_{i,k}^B + \ddot{\lambda}_{j,k}^B p_j \right) \Bigg]. \quad (3.5.16)$$

Given that the arrival rates are well behaved functions without sudden change or discontinuities, the terms containing the derivatives of λ will be small. This implies that the size of $\mathcal{T}$ that guarantees minimal violation of the constraint is mostly dependent upon the second derivative of M_j^ϵ. Furthermore, the value of 2ϵ is significantly smaller than the mesh size $t_{k+1} - t_k$ for any k, so $\frac{\dot{v}_{i,t}^2}{2\epsilon}$ is the dominant term bounding the size of $\frac{d^2}{dt^2}\left(f_{\mathbf{r}}(t, \mathbf{v}_t^M, \dot{\mathbf{v}}_t^M)\right)$. The presence of ϵ in the denominator of $\ddot{M}$

presents an interesting computational trade-off. Increasing ϵ improves the accuracy of the approximation, while also inducing a larger value of L, implying that higher levels of accuracy will require more computation time.

The proposition demonstrates that with a fine grid the error from the discretization is sufficiently small. The advantage of expressing (3.5.16) in terms of the spline intervals is that the different values of d can be set for each interval. Given that bid prices are fairly stationary over the majority of the booking horizon, the value of L is typically very low for most spline meshes. Consequently, a dense grid is only necessary towards the end of the booking horizon, where there is significant changes in the bid prices.

3.5.3 Constraint Generation Procedure

The proposed constraint-sampling algorithm works with a subset of constraints in (3.5.15) and maintains this subset as a collection of lists of capacity vectors $\mathcal{R}_t \subseteq \bar{\mathcal{R}}$, $t \in \mathcal{T}$. We refer to this relaxation of ASOCP as the *Master* problem, its optimal value as $V^R(\mathbf{c}, 0)$ and the corresponding bid-price trajectory as $\mathbf{v}_t^R$. In each iteration, the algorithm finds the most violated constraint of the form (3.5.15). Within a one dimensional search over $t \in \mathcal{T}$, the algorithm fixes $\dot{\mathbf{v}}_t^R$ and $\mathbf{v}_t^R$ at their current values and solves the binary optimization problem

$$\max_{\mathbf{r} \in \{0,1\}^m} \dot{v}_{0,t}^R + \sum_{i \in \mathcal{I}} r_i \dot{v}_{i,t}^R + \sum_{j \in \mathcal{J}} \lambda_{j,t} M_j^\epsilon(\mathbf{v}_t^R) \prod_{i \in A_j} r_i. \tag{3.5.17}$$

The value of this objective is equal to the left-hand-side of (3.5.15) assuming that the corresponding constraints of the form (3.5.13) are tight. Since $M_j^\epsilon(\mathbf{v}_t^R) > 0$, we can substitute a new variable q_j for $\prod_{i \in A_j} r_i$ and enforce $q_j = \prod_{i \in A_j} r_i$ through linear constraints which results in the following linear programming problem

$$\pi_t^R = \max_{\mathbf{q},\mathbf{r}} \quad \dot{v}_{0,t}^R + \sum_{i\in\mathcal{I}} r_i \dot{v}_{i,t}^R + \sum_{j\in\mathcal{J}} q_j \lambda_{j,t} M_j^\epsilon(\mathbf{v}_t^R) \tag{3.5.18}$$

$$\text{s.t.} \quad q_j \le r_i, \ j \in \mathcal{J}, \ i \in A_j, \tag{3.5.19}$$

$$0 \le q_j \le 1, \ j \in \mathcal{J}, \tag{3.5.20}$$

$$0 \le r_i \le 1, \ i \in \mathcal{I}. \tag{3.5.21}$$

Constraints (3.5.19)-(3.5.21) in the row generation problem are equivalent to the constraints in a fixed cost selection problem, which has an integral optimal solution ([55]). Thus, it follows

Proposition 8. The linear programming problem (3.5.18)-(3.5.21) is equivalent to the binary optimization problem (3.5.17) and has a binary optimal solution (i.e., all q_j's and r_i's are at their bounds).

The following proposition shows that the problem of finding the most violated constraint in (3.5.15) for each $t \in \mathcal{T}$ can be solved in polynomial time using linear programming. The row generation problem (3.5.18)-(3.5.21) has an advantage of comparative simplicity because it automatically enforces $q_j = \prod_{i\in A_j} r_i$ at optimality and replaces an explicit consideration of acceptance decisions $\mathbf{u}$ by means of approximation $M_j^\epsilon(\mathbf{v}_t^R)$.

The termination condition of the algorithm is based on the error measure obtained from solving the ASOCP with constraint set $\mathcal{R}_t$, $t \in \mathcal{T}$. In particular, we use a nonnegative cubic spline π^A with knot points t_k, $k \in \mathcal{K}$. The error is found by minimizing the function $E_0 = \min_{\pi^A} \int_0^T \pi_\tau^A d\tau$ subject to spline and nonnegativity constraints of the same form as (3.5.3)-(3.5.5) and (3.5.7)-(3.5.12), respectively, in addition to the constraint $\pi_t^A \geq \max(\pi_t^R, 0),\ \forall t \in \mathcal{T}$, which ensures that the spline bounds the true violations from above (please see Online Appendix B.1 for the full formulation of this problem). We let $E_t = \min_{\pi^A} \int_t^T \pi_\tau^A d\tau$. The algorithm proceeds as follows:

1. Initialize sets $\mathcal{R}_t$ to singletons consisting of a vector of all ones for $t_k,\ \forall k \in \mathcal{K}$ and to empty sets for other $t \in \mathcal{T}$. (Other initialization rules are possible.)

2. Solve the Master problem corresponding to constraints indexed by $\mathbf{r} \in \mathcal{R}_t$, $t \in \mathcal{T}$ to find $V^R(\mathbf{c}, 0)$ and the corresponding bid-price trajectory $\mathbf{v}_t^R$.

3. For each $t \in \mathcal{T}$, find the maximum violation π_t^R by solving (3.5.18)-(3.5.21) for the Master bid-price trajectory $\mathbf{v}_t^R$ obtained in step 2. Let $\mathbf{r}_t^R$, $t \in \mathcal{T}$ be the violation-maximizing capacity vectors.

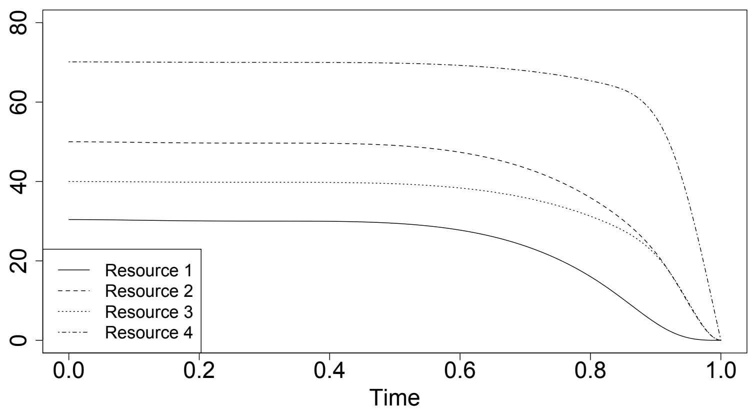
4. Find the error measure E_0 corresponding to the best spline representation π_t^A of the approximation error corresponding to π_t^R. Stop if $\frac{E_0}{V^R(\mathbf{c},0)} \leq \Omega$ where $\Omega > 0$ is a given tolerance parameter. Otherwise proceed to step 5.

5. Let $t^* = \operatorname{argmax}_{t \in \mathcal{T}} \pi_t^R$ be the time corresponding to the most violated constraint. Add the capacity vector $\mathbf{r}_{t^*}^R$ at t^* to $\mathcal{R}_{t^*}$ and proceed to step 2.

Given the Master trajectory we now consider a *modified* bid-price trajectory $\mathbf{v}_t^M$ such that $v_{0,t}^M = v_{0,t}^R + E_t$, $v_{i,t}^M = v_{i,t}^R$, $i \in \mathcal{I}$.

Proposition 9. On termination of the algorithm, the modified bid-price trajectory $\mathbf{v}_t^M$ is feasible for ASOCP, and, consequently, $\frac{E_0}{V^R(\mathbf{c},0)} \leq \Omega$ ensures that this trajectory is within Ω fraction of the optimal solution to ASOCP.

Figure 3.1: ASOCP bid prices trajectories for a 4 resource network.



The initialization step 1 depends on the number and placement of the knot points. The best approach to handling the selection of knot points would be to model the knot points as variables. However, introducing these additional variables creates a highly nonlinear and non-convex optimization problem. Considering the scale of the problems we are trying to solve, this method does not seem appropriate. A practical approach would be to simply adjust the knot points in order to obtain the best performance in terms of expected revenue. We took this approach when constructing our experiments. Another heuristic would be to simulate the arrival process and bid price updates using the dual variables from the LP at various points throughout the booking horizon, selecting the knot points so that the splines can better approximate the expected behavior of the bid prices. The number of knot points can be chosen to balance the computational time with the accuracy of the approximation. The numerical experiments with the proposed computational strategy show that a choice of 3 or 4 knot points is sufficient for achieving effective bid prices.

Figure 3.1 graphs the bid prices produced by the algorithm for a four resource network with $\mathcal{K} = \{0, 0.8, 0.9, 1\}$ and $\Omega = 0.05$. Similar to [1] and in accordance to the discussion in §4, we find that the bid prices remain fairly steady before decreasing at the end of the booking horizon. However, compared to the dynamic bid prices illustrated in [1], the ASOCP bid prices have a smoother shape because, by construction, they are differentiable functions of time.

3.6. Numerical Experiments

In this section we present the results from three sets of numerical experiments. The experiments compared the performance of various bid price control policies used to make accept/reject decisions for a simulated stream of customers. The purpose of these experiments were to provide empirical answers to the following questions:

1. How do discrete time dynamic bid prices compare to continuous time dynamic bid prices in a continuous time setting?
2. How do continuous time dynamic bid prices perform in comparison to LP bid prices that are re-optimized a sufficient number of times?
3. Are continuous time dynamic bid prices scalable to industrial sized networks?

The first two questions were analyzed using network structures similar to the problem instances in [1]. Each spoke had flights traveling to and from the hub, which departed at the end of the booking horizon. In the cases where there were more than one hub, each hub had flights that travel to and from the other hub(s). These flights made up the network's resources. If H was the number of hubs in the network and L was the number of spokes, then the total number of resources in the network was the sum of the number of hub-spoke legs and the number of hub-hub legs. Thus, the network had $m = 2HL + H(H-1)$ single-leg flights. The set of products for each network included all possible OD itineraries. Both high fare and low fare tickets were offered for each OD pair. The load factor was calculated as the expected demand for seats over all itinerary divided by the capacity of the network and was equal to $\int_0^T \sum_j \lambda_{j,t} A_j dt/cm$. The main difference between our experiments and [1]'s was that time was continuous and scaled to the interval $[0, 1]$.

The third set of experiments utilized the Porter Airlines' schedule to create an industrial scale network. Porter Airlines is a regional airline based in Toronto, Canada, with hubs in Toronto, Montreal, Ottawa, and Halifax. The airline provides "short haul" flights to cities in Ontario, Quebec, and Atlantic Canada, as well as cities in the United States, including daily flights to Boston, Newark, and Chicago. Since Porter's business model emphasizes service, speed, and convenience for business and leisure travellers, the airline offers frequent flights between major Canadian and US cities, flying up to 20 times daily from Toronto to Ottawa and 11 times daily from Toronto to Newark-New York. The experiments over the Porter Airlines network demonstrate the potential of the ASOCP as a solution method for problems of industrial size.

All solution methods for each set of experiments were executed through SHARCNET, a high performance Canadian research computing consortium. The first set of experiments was run on the SHARCNET serial throughput cluster Kraken using single threaded AMD Opteron 2.2 GHz processors. The second and third sets of experiments, which have larger networks and re-optimization, were run on the SHARCNET serial throughput cluster ORCA using 8 threaded Intel Xeon 2.7 GHz processors. For further details on the experimental design and additional documentation of the results please see Online Appendix B.2.

Table 3.1: Revenue and run-time in seconds for fixed bid price experiments.

	Upper	Average Revenue			Average Run-Time		
κ	Bound	LP	DD	ASOCP	LP	DD	ASOCP
1	4524	3587	3791	3963	0.0005	0.0195	0.4145
2	9227	7551	7812	8289	0.0003	0.0093	0.4421
3	23323	19503	20075	21488	0.0005	0.0085	0.4992
4	46812	39495	40360	43512	0.0003	0.0130	0.5311
5	93790	79501	81320	87731	0.0008	0.0143	0.5645

3.6.1 Fixed Bid Price Experiments

The first set of experiments examines whether discrete time dynamic bid prices provide a suitable control policy when customer arrivals occur in continuous time. To analyze this question, fixed bid price control policies from the LP, DLP, and ASOCP were used to generate revenues from a simulated stream of customers modeled by a Poisson process. We chose to use the dynamic disaggregation (DD) approach to solve the DLP, since the approach represents state of the art computational efficiency for dynamic bid prices. The numerical experiments in [68] demonstrated that the DD is the fastest method for producing dynamic bid prices in discrete time. The DD approach produces time-dependent bid prices by solving the LP and adding variables and constraints simultaneously to the problem until optimality is obtained. Since the DLP bid prices are constant for the majority of the booking horizon, the dynamics of the bid prices are captured entirely by the columns and rows added in the constraint generation procedure. For brevity, we refer the reader to §3 of [68] for a complete description of the DD algorithm and optimization problems. To approximate the Poisson process in discrete time, we set the probability of an arrival, ρ, to 0.8 producing a discrete booking horizon of $\frac{\lambda}{\rho}$ periods. For reference, [1] fixed the probability of an arrival at 0.8 in his experiments and [68] fixed the probability of an arrival to 0.9

in their experiments. For computing the ASOCP we set $\mathcal{K} = \{0, 0.8, 0.9, 1\}$, $N_k = 50$, and $\Omega = 0.05$. The knot points were selected in this manner to capture the majority of the variation at the end of the booking horizon (see Figure 3.1). The initial set of time-capacity vector pairs before starting the constraint generation procedure was one available seat for each resource at each knot point. For all problem instance $\epsilon = 0.0001$.

The bid price policies were tested on a 1 hub 3 spoke (6 resources and 24 products) network. Similar to [21] and [30] the problems were scaled by the parameter κ, such that the capacity was $c = \kappa c_0$ and the expected arrivals were $\lambda = \kappa\lambda_0$. The expected arrivals and capacity had a base values of $\lambda_0 = 40$ customers and $c_0 = 6$ seats per flight, which corresponded to a load factor of 1.678. The problems were varied by scale factors $\kappa \in \{1, 2, 5, 10, 20\}$. The demand for each OD pair was stationary. The probability that an itinerary request was for a low-fare (high-fare) class ticket was 0.75 (0.25). The prices for low fare products were randomly drawn from the set $[20, 50]$ with equal probability. The high fare products were priced at five times the corresponding low fare. Each pair of network parameters was simulated 1000 times. The experiments only considered fixed bid price controls, implying that each method was solved at $t = 0$ and the resulting bid prices were used to make product acceptance decisions for the entire booking horizon.

Table 3.1 reports the upper bound on the revenue, the simulated revenues for each bid price policy, and the average run-times for fixed bid price experiments. The objective value of the DD was selected as the upper bound, since DD provides a tighter upper bound compared to the LP. For each network, the ASOCP generated the highest revenue, capturing between 87.6-93.5% of the upper bound, while the

Table 3.2: Run-time in seconds for fixed bid price experiments with asymmetric capacities.

	$\nu = 1$		$\nu = 2$		$\nu = 3$		$\nu = 4$	
κ	DD	ASOCP	DD	ASOCP	DD	ASOCP	DD	ASOCP
1	0.0208	0.4986	0.0333	0.5024	0.2054	0.5986	0.9509	0.8887
2	0.0405	0.5353	0.0657	0.5321	0.3028	0.5493	2.5630	0.5983
3	0.0621	0.6010	0.0784	0.5941	1.4724	0.6045	9.7771	0.6031
4	0.0838	0.6792	0.1276	0.6410	2.1020	0.6543	17.8454	0.6814
5	0.1178	0.6541	0.1557	0.7169	1.8977	0.6995	33.8710	0.6697

Table 3.3: Revenue for fixed bid price experiments with asymmetric capacities.

	$\nu = 1$		$\nu = 2$		$\nu = 3$		$\nu = 4$	
κ	DD	ASOCP	DD	ASOCP	DD	ASOCP	DD	ASOCP
1	3786	3926	3767	3869	3674	3761	3501	3563
2	7726	8230	7728	8154	7541	7878	7420	7703
3	20189	21579	20212	21359	19609	20489	19424	20348
4	40351	43370	40509	42732	39977	42064	38790	41015
5	81760	87685	82187	86438	81028	84841	77502	81978

LP and DD captured 79.2-84.8% and 83.8-86.7% of the upper bound, respectively. Although the ASOCP generates greater revenue relative to the DD when arrivals are in continuous time, the run-time of the DD is considerable faster. The ASOCP's average run-time across the various scale factors ranged within 20-60 times the run-time of the DD. From the standpoint of a fixed bid price policy, the DD offers a compromise relative to the LP and ASOCP in terms of the revenue versus run-time performance tradeoff. There is a distinct revenue improvement relative to the LP, with a run substantial run-time improvement relative to the ASOCP.

To account for the stochastic nature of demand and capacity consumption, airlines update capacity vectors and re-optimize bid prices periodically throughout the booking horizon. Consequently, when airlines solve for the optimal bid prices, the capacity of each resources is unlikely to be balanced. For either the DD or ASOCP to be a functional solution method for dynamic bid prices, the run-times should not change significantly when the available capacities across resources vary. To test the DD and ASOCP performance given asymmetric capacities, the first set of experiments was re-simulated with randomized capacities such that

$$\sum_{i \in \mathcal{I}} c_i = 6\kappa c_0 \quad \text{and} \quad c_i \in [\kappa(c_0 - \nu), \ldots, \kappa(c_0 + \nu)], \quad \forall i \in \mathcal{I}. \tag{3.6.1}$$

The parameter ν controls variability in the capacity vectors. For $\nu \in \{1, 2, 3, 4\}$, each set of demands, prices, and arrival paths was simulated for 10 randomized capacity vectors satisfying condition (3.6.1).

Table 3.2 reports the average run-times for the DD and ASOCP across the different values of ν and κ. For each value of κ, the average DD run-time appears to increase substantially. The run-times for $\nu = 4$ increase with κ and range between $45 - 187$ times the run-times for $\nu = 1$. On the other hand, the ASOCP run-times for $\nu = 4$ are $1.003 - 1.782$ times the run-times for $\nu = 1$. These distinct behaviors results from the structure of dynamic bid prices and substantial differences in computing methodology. Theorem 2 of [1] states that the value of the bid price for a given resource is static from the start of the booking horizon until a critical time where the bid price becomes dynamic. As the capacity vectors increase in variability, select flight legs become scarcer at the start of the horizon. Bookings for these resources have a greater impact on the dynamics of the bid price, resulting in the critical times occurring earlier in the booking horizon as ν increases. Therefore, as the capacity vectors increase in variability, the DD requires adding a greater number of rows and columns before the algorithm terminates. On the other hand, the ASOCP considers the entire booking horizon using the spline approximation regardless of the capacity vector. Thus, there is little impact on the algorithm's timing and it is fairly robust to variation in capacity. Finally, despite the dramatic increase in run-time for the DD, Table 3.3 shows that the average revenue is still less than the revenue generated by the ASOCP. Since the run-times increase as the variability in the starting capacity grows, we conclude that the DD is at a disadvantage for situations with frequent updating and unbalanced capacities, when arrivals occur in continuous times.

3.6.2 Bid Price Experiments with Updates

In a study comparing the behavior of LP heuristics, [30] provide theoretical results that advocate using LP bid prices over the asymptotically optimal booking limit heuristics, provided that the LP is resolved sufficiently frequently. For the ASOCP to have value as a bid price solution method, it must be able to provide higher revenues than the LP bid price control that is updated sufficiently often throughout the booking horizon. To ensure adequate updating, we re-optimized the LP at uniformly spaced intervals as many times as the expected number of arrivals over the course of the booking horizon. On the other hand, we limited the number of optimization updates for the ASOCP to 10 (also at evenly spaced intervals) in order to keep the run-time for both bid price controls at a similar order of magnitude. The experiments establish that the ASOCP provides an effective bid price policy and demonstrates the value of time-dependent bid price compared to static bid prices that are updated repeatedly. Since $\mathcal{K} = \{0, 0.8, 0.9, 1\}$, with each update, the middle knot points were scaled so that they represented 80% and 90% of the remaining booking horizon.

The experiment networks consisted of $H \in \{1, 2, 3\}$ hubs connected to $L = 3$ spokes. The initial capacity for each single-leg flight was $c = 150$. The arrival rates

Table 3.4: Revenue and run-time in seconds for multi-hub network with updates.

Network	Load	Upper	Average Revenue		Unused Capacity		Total Run-Time	
$\{h, m, n\}$	Factor	Bound	LP	ASOCP	LP	ASOCP	LP	ASOCP
{1,6,24}	1.636	57534	55728	56879	62.89	65.77	0.132	2.144
	1.796	62136	59988	61485	58.48	62.10	0.149	1.988
	1.969	66250	64225	65880	37.04	41.28	0.162	2.081
{2,14,112}	1.558	94207	92391	93305	334.30	338.56	0.378	6.199
	1.789	98634	97072	97997	294.78	299.66	0.420	6.525
	1.975	107561	105042	105947	263.73	274.47	0.457	6.891
{3,24,264}	1.478	142088	139936	141066	437.33	447.11	0.962	54.394
	1.602	150922	147440	148839	350.85	362.07	1.032	54.976
	1.737	157033	154100	155662	360.93	372.02	1.4979	52.687

λ were varied in order to produce different load factors. For the one hub experiments $\lambda \in \{1000, 1100, 1200\}$, for the two hub experiments $\lambda \in \{1600, 1800, 2000\}$, and for the three hub experiments $\lambda \in \{2400, 2600, 2800\}$. Each network was simulated 500 times with the bid price policies facing identical customer arrivals modeled by a Poisson process with the same parameters as described in §3.6.1.

Similar to the fixed bid price experiments in §3.6.1, we evaluated the performance of the ASOCP by benchmarking revenues against the LP. The average revenues and the upper bound as well as the run-times and remaining inventory are listed in Table 3.4. The percentage of the revenue generated by the LP relative to the LP bound ranged from 96.5% to 98.5%, with an average of 97.6%, while the revenue percentage for the ASOCP ranged from 98.4% to 99.4%, with an average of 99.1%. The low performance gap is due to the asymptotic optimality of the LP Bound; however, the ASOCP is still able to provide a revenue boost compared to LP bid prices. Under a direct comparison of bid price policy performance, the ASOCP provided an average improvement from 0.81% to 2.58% (1.01% across the entire set). Although the total run-time was 25 times larger on average, with the LP being updated 1000-2800 times, the solution time for each network was under a minute, which is reasonable considering the revenue improvements. The experiments also show that ASOCP can achieve good performance with limited updating, unlike the LP, which requires frequent re-optimization. Another interesting observation is that the LP had a higher utilization of resources for each network. This implies that a significant amount of the ASOCP revenue improvement stems from denying product requests in favor of reserving capacity for more profitable itineraries.

3.6.3 Porter Airlines Network

We simulated bid prices using the ASOCP and LP over an entire day in Porter's network, which consisted of 182 single-leg flights and 1098 possible itineraries. Table 3.5 provides a list of Porter's single-leg flights and their frequency of service on December 1, 2011. We consider the network with one fare class and two fare classes for each itinerary. Similar to §3.6.1 and §3.6.2, the booking horizon was scaled to the interval $[0, 1]$ and time was continuous. The capacity of each flight was 70 seats, coinciding with Porter's fleet of Bombardier Q400 aircraft and the expected demand over the entire network for both the single and two-fare cases was given by $\lambda = 12000$. The LP and ASOCP optimization problems were optimized $\phi \in \Phi \equiv \{5, 10, 20, 35\}$ times at uniform intervals over the booking horizon. The Porter network was simulated 200 times, with randomly sampled expected demand and price for each itinerary.

The demand for a product j was determined using a random variable, ψ_j, which was uniform on the interval $[0, m]$. The corresponding probability for itinerary j, conditional on an arrival, was $\psi_j / \sum_{j \in \mathcal{J}} \psi_j$. For the Porter Airline experiments, product demand was described by a probability ϕ_j that the product requested was product j,

Table 3.5: List of single-leg flights for Porter Airlines network (December 1st, 2011).

Departure City	Arrival City	Number of Flights
Boston (BOS)	Toronto (YTZ)	7
Chicago (MDW)	Toronto (YTZ)	6
Halifax (YHZ)	Montreal (YUL)	2
Halifax (YHZ)	Ottawa (YOW)	5
Halifax (YHZ)	St John's (YYT)	4
Moncton (YQM)	Toronto (YTZ)	1
Montreal (YUL)	Halifax (YHZ)	2
Montreal (YUL)	Toronto (YTZ)	18
Newark (EWR)	Toronto (YTZ)	11
Ottawa (YOW)	Halifax (YHZ)	5
Ottawa (YOW)	Moncton (YQM)	1
Ottawa (YOW)	Toronto (YTZ)	20
Quebec City (YQB)	Toronto (YTZ)	3
Sault Ste Marie (YAM)	Toronto (YTZ)	3
St John's (YYT)	Halifax (YHZ)	4
Sudbury (YSB)	Toronto (YTZ)	3
Thunder Bay (YQT)	Toronto (YTZ)	5
Toronto (YTZ)	Boston (BOS)	7
Toronto (YTZ)	Chicago (MDW)	6
Toronto (YTZ)	Montreal (YUL)	18
Toronto (YTZ)	Newark (EWR)	11
Toronto (YTZ)	Ottawa (YOW)	20
Toronto (YTZ)	Quebec City (YQB)	3
Toronto (YTZ)	Sault Ste Marie (YAM)	3
Toronto (YTZ)	Sudbury (YSB)	3
Toronto (YTZ)	Thunder Bay (YQT)	5
Toronto (YTZ)	Windsor (YSG)	3
Windsor (YSG)	Toronto (YTZ)	3

Table 3.6: Average revenue, unused capacity, and run-time in seconds per optimization for the Porter Airlines experiments.

		Revenue		Unused Capacity		Run-time	
Fare Classes	ϕ	LP	ASOCP	LP	ASOCP	LP	ASOCP
1	5	1044910	1080080	3153.11	2707.22	0.0048	11.852
	10	1065150	1080120	2922.04	2714.89	0.0042	18.638
	20	1074400	1080130	2814.73	2717.42	0.0044	16.216
	35	1077800	1080750	2777.49	2710.88	0.0045	17.821
2	5	2901300	2916650	3575.41	3174.25	0.0048	14.356
	10	2919470	2927040	3389.88	3053.13	0.0049	20.449
	20	2928510	2939570	3300.1	3208.57	0.0047	19.225
	35	2931960	2942420	3261.59	3215.68	0.0048	23.383

conditional on an arrival. The probabilities are found using weights, ψ_j, which were uniform random variables between $[\underline{M}_j, \overline{M}_j]$, influenced by the frequency of flights between an OD pair and the time of day the flight leaves. $\underline{M}_j$, was determined by the ratio of the number of flights between each OD pair to the average number of flights between each OD pair, γ_j, and was equaled to $\alpha_1 - \beta_1\sqrt{\gamma_j}$. $\overline{M}_j$ was dictated by the flight's time of day and equaled to $\alpha_1 + \beta_1\delta_j^2$, where δ_j was the absolute time difference from 2:00 pm and the takeoff time of flight j. The probability of a request being for product j was given by $\phi_j = \psi_j / \sum_{j\in\mathcal{J}} \psi_j$. The price for each product was determined by a uniform random variable between $[\underline{N}_j, \overline{N}_j]$. $\underline{N}_j$ was determined by the number of flights and was equaled to $\alpha_2 - \beta_2\gamma_j$. $\overline{N}_j$ was determined by the number of single-leg flights that product j required, $\sum_{i\in\mathcal{I}} a_{ij}$, and was equaled to $\alpha_2 + \beta_2^2 \sum_{i\in\mathcal{I}} a_{ij}$. For the Porter network, the parameters used were $\alpha_1 = 10$, $\beta_1 = \frac{1}{2}$, $\alpha_2 = 200$, and $\beta_2 = 4$. In addition, the set of knot points was reduced to $\mathcal{K} = \{0, 0.8, 1\}$ and the set $\mathcal{T}$ was increased to include 100 time points evenly spaced between adjacent knot points. Once again, the knot points were appropriately scaled with each re-optimization.

Table 3.6 compares the revenues, unused capacity, and run-time for the LP and

ASOCP bid price controls. For both one fare and two fare class problems, the ASOCP produced higher revenues than the LP bid price for each $\phi \in \Phi$. In addition, for a single fare class, the ASOCP-based policy with $\phi = 5$ generated more revenue than each LP-based policy. The improvement offered by the ASOCP ranged between 0.27-3.37% and 0.26-0.53% for the one and two-fare class problems, respectively. The ASOCP also utilized more capacity in comparison to the LP in each experiment. Interestingly, this is inconsistent with the capacity utilization for the randomly generated network experiments. Although the ASOCP run-times are several orders of magnitude greater than the LP run-times, we argue that the ASOCP is a promising approach for solving large-scale capacity control problems. Airlines have access to superior computational power through industrial server farms and work stations, which would increase the optimization speed for computing ASOCP bid prices. On the other hand, airlines are restricted in the number of times that they can re-optimize bid prices. Bid price controls involve extensive scenario analysis and constant monitoring, often leaving bid prices to be computed overnight ([62]). If the booking horizon represents a ten-week period and optimization occur overnight, then the values of $\phi \in \{5, 10, 20, 35\}$ corresponds to re-optimizing the bid prices once every two weeks, once a week, twice a week, and every other day, respectively, over the ten-week period. From this perspective, on the nights when the airline updated the bid prices, the average run-time, regardless of the number of fare classes was under 25 seconds. Finally, SOCP solvers have not evolved to the same extent as LP solvers in terms of solution speed. As SOCP solvers mature, the ASOCP will become even more viable as a solution method for computing bid prices for industrial networks.

3.7. Conclusion and Future Research

In this article we construct a bid price control policy starting from a continuous time network RM framework. After substituting the optimal control policy into the HJB equation and reformulating it as a differential inclusion, we introduce two approximations into the inclusion to establish the AOCP. Using the monotonicity of the bid prices, approximation theory is used to develop the ASOCP, which makes the number of variables independent of the time horizon. Finally, we employ an efficient constraint generation procedure allowing the ASOCP to produce time-dependent bid prices by considering only a select number of time-capacity vectors. The numerical experiments highlight the effectiveness of the proposed approach in generating bid prices as well as its scalability by solving problems on an industrial sized network. Future research could extend the ASOCP structure to incorporate customer choice. It would also be interesting to combine ASOCP with robust optimization to incorporate uncertainty in the arrival rates. The ASOCP not only has the potential to improve revenues for large airlines, but its methodology can be applied to other areas of RM and large-scale approximate optimal control problems.

Chapter 4

Uniqueness in a Multi-Product Non-Cooperative Consumer-to-Consumer Market

4.1. Introduction

With the growth of e-commerce, consumer-to-consumer (C2C) markets have become an increasingly important platform for consumers to purchase and sell commodities. Websites such as eBay, Taobao, Craigslist, Kijiji, and Stub-Hub offer classified and online auctions allowing consumers to exchange goods worldwide. The estimated value of goods traded in C2C markets has grown from \$30 billion in 2003 to over \$80 billion in 2010. The continued growth of global C2C sales in the last decade has largely been driven by the Chinese e-tailing (online retailing) industry. E-tailing sales, which include both business-to-consumer (B2C) and C2C markets, totalled \$120 billion in 2011 and sales have since grown to an estimated \$190 billion. Although the e-tailing figures account for both B2C and C2C trading, more than 70% of sales in 2012 were attributed to C2C transactions. With China's economic rise, growth in consumer culture, and sizeable online population, e-tailing sales are forecasted to reach \$420-\$600 billion by 2020. The global growth in C2C trading implies that exchange markets will have a significant impact on the retailing and business environment.[1]

Although there is a vast literature on general exchange markets, pioneered by [8], the majority of existing research focuses on complete markets within a general equilibrium framework. Since operational and strategic decisions, as well as consumers purchasing decisions, are often made within the context of an industry-specific business environment (as opposed to within the context of the entire economy), we consider the exchange of substitutable products within an incomplete market. We develop a model for a general multi-product exchange market where consumers, constrained by their endowed product set, decide which products to buy and sell in a C2C market. The objective of our research is to connect C2C markets with a general theory of oligopoly competition between firms offering a variety of differentiated products. C2C trading has become more prevalent and is emerging in a variety of new markets. By providing a general framework of C2C trading, we hope to expand secondary market research applications.

The tradition of the oligopoly literature is to model demand using a representative consumer who has a quadratic utility function for the differentiated products that exist in the market ([35]). Quadratic utility is analytically tractable, and with fairly unrestrictive assumptions on the utility parameters, is often sufficient to guarantee a unique outcome for the oligopoly equilibrium (see [25]; [2]; [35]). We follow the quadratic utility approach; however, due to the nature of C2C exchanges, we introduce m different representative consumers. The consumers are differentiated by their willingness to pay and their price sensitivity for the substitutable goods and vary in terms of the population size that they represent. With the exception of [20],

[1] Revenues from the Chinese e-tailing business are taken from [22].

the representative consumer approach has always been restricted to a single consumer comprising the entire market. Since [20] operates under the assumption that each consumer group has the same preference for goods and is differentiated only by their respective incomes, our model of multiple representative consumers is a novel contribution, which simplifies the analysis of multi-product environments including secondary markets. The quadratic utility assumption leads consumers to have linear demand for goods and controls the number of products consumers are willing to sell. Equilibrium is established when goods demanded equals goods supplied at market clearing prices.

As [33] discuss, modeling multi-product oligopolies is complicated by interdependence between product demand and prices, in addition to differentiation amongst products produced by competing firms. For example, linear demand may lead to regions of prices that yield negative demand values for products, while potentially inflating the demand for a substitute to an arbitrarily high level. [58] demonstrate that reformulating the multi-product linear demand as a linear complementarity problem (LCP) allows the demand function to be properly constructed over the entire set of non-negative prices. This approach was utilized by [24], who reduce the LCP to a linear program, and by [26] in the study of sequential Bertrand supply chains.

To ensure that the goods supplied in equilibrium do not exceed the quantity owned by the seller, our model utilizes a LCP formulation similar to [58], [24], and [26]. Unfortunately, the assumptions of the representative consumers' quadratic utility is insufficient to guarantee uniqueness or even existence of the equilibrium solution. A similar complication arises in [26], due to the presence of multiple retailers in the supply chain optimizing their wholesale prices. However, in the case where consumers have symmetric demand sensitivities, the equilibrium quantities traded and the market clearing price are shown to be unique. For the asymmetric demand sensitivities, we remark on conditions that ensure the uniqueness of the solution is maintained.

4.2. Characterizing Consumers

Consider a population of m types of heterogeneous consumers, where each group $i \in \mathcal{I} \equiv \{1 \ldots m\}$ has size k_i. There are n types of durable heterogeneous products, which are gross-substitutes, owned by members of the population. A consumer group i is characterized by product preference and the consumer's willingness to pay for preferred products. These parameters that characterize taste are fixed for each consumer group i, regardless of the population size. In any given period, the total number of products $j \in \mathcal{J} \equiv \{1, ..., n\}$ owned by a population group is denoted by the vector $\mathbf{q}_i$. The combination of $\mathbf{q}_i$ and utility parameters $\tilde{\mathbf{a}}_i$ and $\tilde{\mathbf{B}}_i$ define the utility of consumer group i.

To develop the intuition behind our approach, we start by examining the utility earned by the typical individual within a consumer group i. The value provided by each product is assumed to be a linear function dependent on the type and quantity of other products owned by the average consumer segment. If the total number of products owned by the group is $\mathbf{q}_i$, then the average consumer owns $\mathbf{q}_i/k_i$. The value provided by each product to a type i consumer endowed with quantity $\mathbf{q}_i/k_i$ is

$$v_{ij}(\mathbf{q}_i) = \tilde{a}_{ij} - \frac{1}{2k_i}\tilde{\mathbf{B}}_{ij}\mathbf{q}_i.$$

Since the average consumer owns quantity $\mathbf{q}_i/k_i$, the individual utility for the average consumer i is

$$\tilde{u}_i(\mathbf{q}_i) = \left(\frac{\mathbf{q}_i}{k_i}\right)^T \left(\tilde{\mathbf{a}}_i - \frac{1}{2k_i}\tilde{\mathbf{B}}_i\mathbf{q}_i\right). \tag{4.2.1}$$

To calculate the total utility for the consumer segment i, the individual utility (4.2.1) is multiplied by the population size. The total utility of the representative consumer group i owning product quantities $\mathbf{q}_i$ is the quadratic expression

$$u_i(\mathbf{q}_i) = \mathbf{q}_i^T \left(\tilde{\mathbf{a}}_i - \frac{1}{2k_i}\tilde{\mathbf{B}}_i\mathbf{q}_i\right) \tag{4.2.2}$$

The vector $\tilde{\mathbf{a}}_i = ||\tilde{a}_{ij}||$ is interpreted as a person of type i willingness to pay for each product j provided that the customer owns no products. The matrix $\tilde{\mathbf{B}}_i = ||\tilde{b}_{ijj'}||$, is the decrease in the willingness to pay for products j as quantity of good j' increases. Each representative consumer's willingness to pay is assumed to be positive for all products. Similarly, $\tilde{\mathbf{B}}_i$ is assumed to be positive definite for all i. The assumption on $\tilde{\mathbf{B}}_i$ implies that the utility function (4.2.2) is strictly concave for each representative consumer.

The objective of the representative consumer i is to maximize the surplus

$$CS_i(\mathbf{q}_i) = \mathbf{q}_i^T \big(\tilde{\mathbf{a}}_i - \frac{1}{2k_i}\tilde{\mathbf{B}}_i\mathbf{q}_i\big) - \mathbf{p}\mathbf{q}_i, \tag{4.2.3}$$

given a set of prices for each product $\mathbf{p}$. Differentiating (4.2.3) provides the optimal quantities

$$\mathbf{q}_i = k_i\tilde{\mathbf{B}}_i^{-1}\big(\tilde{\mathbf{a}}_i - \mathbf{p}\big).$$

Introducing the parameters $\mathbf{B}_i = \tilde{\mathbf{B}}_i^{-1}$ and $\mathbf{a}_i = \mathbf{B}_i\tilde{\mathbf{a}}_i$, the optimal quantities are expressed as

$$\mathbf{q}_i = k_i\big(\mathbf{a}_i - \mathbf{B}\mathbf{p}\big).$$

The interpretations of vector $\mathbf{a}_i$ and matrix $\mathbf{B}_i$ are the product preference for any consumer in group i when prices are zero and the demand sensitivities to changes in price. Notice that the vector $\mathbf{B}_i\tilde{\mathbf{a}}_i$ is the per person demand for products and $\mathbf{B}_i$ is the per person price elasticity of demand. Both of these values are independent of the population size of the group, implying that a characterization i is uniquely defined by individual product preference and sensitivities. Since, $\tilde{\mathbf{B}}_i$ is positive definite, the diagonals of $\mathbf{B}_i$ are strictly positive. Matrices $\mathbf{B}_i$ are also assumed to be strictly row and column diagonally dominant, i.e.

$$\begin{aligned} b_{ijj} &> \sum_{j \neq j'} |b_{ijj'}|, \quad \forall (i,j) \\ b_{ijj} &> \sum_{j \neq j'} |b_{ij'j}|, \quad \forall (i,j). \end{aligned}$$

As [26] discuss, the literature modeling gross substitute markets often employ this assumption, since it is intuitive and likely to hold in most applications.

4.3. Consumer-To-Consumer Exchange Market

To develop the intuition for the model, we begin by separating the decisions of the representative consumer to buy and sell products within the C2C market. Define the variables $\mathbf{x}_i^s$ as the number of units of each products available for trade by customer group i and $\mathbf{x}_i^d$ as the demand for each product by customer group i given a price vector $\mathbf{p}$.

Consumer type i's equilibrium quantities for any product after the transaction period is

$$\hat{\mathbf{q}}_i = \mathbf{q}_i + \mathbf{x}_i^d - \mathbf{x}_i^s. \quad (4.3.1)$$

Substituting (4.3.1) into (4.2.2), the representative consumer with preference i earns surplus

$$CS_i = \left(\mathbf{q}_i + \mathbf{x}_i^d - \mathbf{x}_i^s\right)^T \left(\tilde{\mathbf{a}}_i - \tfrac{1}{2k_i}\tilde{\mathbf{B}}_i\left(\mathbf{q}_i + \mathbf{x}_i^d - \mathbf{x}_i^s\right)\right) - \mathbf{p}(\mathbf{x}_i^d - \mathbf{x}_i^s). \quad (4.3.2)$$

Since consumers cannot sell more than what they own and cannot sell negative values, supply is restricted by the constraints

$$\mathbf{0} \leq \mathbf{x}_i^s \leq \mathbf{q}_i. \quad (4.3.3)$$

Similarly, demand is nonnegative, so the representative consumer decisions are constrained by

$$\mathbf{0} \leq \mathbf{x}_i^d. \quad (4.3.4)$$

Given that the objective of each representative consumer is to maximize (4.3.2) such that each consumer group satisfies (4.3.3) and (4.3.4), the Lagrangian for each consumer group is given by:

$$L = CS_i(\mathbf{x}_i, \mathbf{x}) + \underline{\mathbf{s}}_i^T \mathbf{x}_i^s + \overline{\mathbf{s}}_i^T \left(\mathbf{q}_i - \mathbf{x}_i^s\right) + \underline{\mathbf{d}}_i^T \mathbf{x}_i^d.$$

The variables $\underline{\mathbf{s}}$ and $\underline{\mathbf{d}}$ are multipliers on the constraints enforcing nonnegativity on the supply and demand variables for each consumer group, while $\overline{\mathbf{s}}$ is a multiplier on the upper bound for each consumer's product supply.

The market supply and market demand for each product is

$$\mathbf{S}(\mathbf{p}) = \sum_{i \in \mathcal{I}} \mathbf{x}_i^s \qquad \textit{and} \qquad \mathbf{D}(\mathbf{p}) = \sum_{i \in \mathcal{I}} \mathbf{x}_i^d.$$

Since the quantity of goods for sale is limited by consumer supply, the market clearing conditions occurs at the price where market supply is equal to the demand for each product. This implies for any initial allocations of products $\mathbf{q} \in \Re_+^{mn}$, the transfer of goods between consumer groups is given by solving the complementarity problem:

$$\begin{aligned}
&\mathbf{x}_i^d - \mathbf{x}_i^s = \mathbf{a}_i - \mathbf{B}_i(\mathbf{p} + \underline{\mathbf{d}}_i) - \mathbf{q}_i \quad \forall i \in \mathcal{I} \\
&\mathbf{x}_i^d - \mathbf{x}_i^s = \mathbf{a}_i - \mathbf{B}_i(\mathbf{p} - \underline{\mathbf{d}}_i - \underline{\mathbf{s}}_i + \overline{\mathbf{s}}_i) - \mathbf{q}_i \quad \forall i \in \mathcal{I} \\
\mathbf{0} \quad \leq \quad &\underline{\mathbf{s}}_i \perp \mathbf{x}_i^s \quad \geq \quad \mathbf{0} \quad \forall i \in \mathcal{I} \\
\mathbf{0} \quad \leq \quad &\overline{\mathbf{s}}_i \perp \mathbf{q}_i - \mathbf{x}_i^s \quad \geq \quad \mathbf{0} \quad \forall i \in \mathcal{I} \\
\mathbf{0} \quad \leq \quad &\underline{\mathbf{d}}_i \perp \mathbf{x}_i^d \quad \geq \quad \mathbf{0} \quad \forall i \in \mathcal{I} \\
&\sum_{i \in \mathcal{I}} \mathbf{x}_i^s = \sum_{i \in \mathcal{I}} \mathbf{x}_i^d.
\end{aligned}$$

Lemma 5. The supply of product j by consumer i is increasing in price, quantities of product held, and product preference. The demand for product j by consumer i is decreasing in price, quantities of product held, and increasing in product preference.

Under the assumption that there is no opportunity for arbitrage, then in equilibrium no consumer will buy and sell the same product. This implies that the following complementarity condition will always be satisfied in equilibrium

$$\mathbf{0} \quad \leq \quad \mathbf{x}_i^s \;\perp\; \mathbf{x}_i^d \quad \geq \quad \mathbf{0}.$$

Thus, for simplicity we let the variable $\mathbf{x}_i = \mathbf{x}_i^d - \mathbf{x}_i^s$ be the joint supply and demand of consumer i. In this case each consumer groups is trying to maximize their surplus

$$\begin{aligned} CS_i(\mathbf{x}) = \max_{\mathbf{x}_i} \quad & \big(\mathbf{x}_i + \mathbf{q}_i\big)^T \left(\tilde{\mathbf{a}}_i - \frac{1}{2k_i}\tilde{\mathbf{B}}_i\big(\mathbf{x}_i + \mathbf{q}_i\big)\right) - \mathbf{x}_i^T \mathbf{p} \\ \text{s.t.} \quad & \mathbf{x}_i + \mathbf{q}_i \geq \mathbf{0}. \end{aligned}$$

Proposition 10. For any given price $\mathbf{p}$, the optimal quantity decisions for consumer segment i is uniquely determined by the complementarity problem

$$\mathbf{0} \leq \mathbf{s}_i \perp \mathbf{x}_i = k_i\big(\mathbf{a}_i - \mathbf{B}_i\big(\mathbf{p} - \mathbf{s}_i\big)\big) - \mathbf{q}_i \geq \mathbf{0}.$$

For compactness, define $\bar{\mathbf{a}}_i = k_i\mathbf{a}_i$, $\bar{\mathbf{B}}_i = k_i\mathbf{B}_i$, $\bar{\bar{\mathbf{a}}} = \sum_i \bar{\mathbf{a}}_i$, $\bar{\bar{\mathbf{B}}} = \sum_i \bar{\mathbf{B}}_i$, and $\bar{\bar{\mathbf{q}}} = \sum_i \mathbf{q}_i$. With market clearing conditions, the equilibrium solution can be found by solving the complementarity problem:

$$\bar{\mathbf{a}}_i - \bar{\mathbf{B}}_i\big(\mathbf{p} - \mathbf{s}_i\big) - \mathbf{x}_i - \mathbf{q}_i = \mathbf{0}, \quad \forall i \in \mathcal{I} \tag{4.3.5}$$

$$\mathbf{0} \leq \mathbf{s}_i \perp \mathbf{x}_i + \mathbf{q}_i \geq \mathbf{0}, \quad \forall i \in \mathcal{I} \tag{4.3.6}$$

$$\sum_{i \in \mathcal{I}} \mathbf{x}_i = \mathbf{0}. \tag{4.3.7}$$

Corollary 5. Consider a market defined by $\mathbf{a}_i$, $\mathbf{B}_i$, quantities $\mathbf{q}_i$ and population size k_i for $i \in \{1, \ldots, m\}$. For a given set of multipliers $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m]^T > 0$, the equilibrium price is given by

$$\mathbf{p} = \bar{\bar{\mathbf{B}}}^{-1}\big(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}} + \sum_i \bar{\mathbf{B}}_i \mathbf{s}_i\big) \tag{4.3.8}$$

Corollary 6. If $\bar{\bar{\mathbf{q}}} > \mathbf{0}$, then in equilibrium the following complementarity condition holds,

$$\prod_{i=1}^{m} s_{ij} = 0, \quad \forall j \in \mathcal{J}.$$

4.4. Uniqueness for Correlated Demand Elasticities

We demonstrate the uniqueness of the solution to the exchange market for the case of correlated demand cross-elasticities. Defining $\mathbf{B}$ as the market demand sensitivity to prices, consumers are differentiated by their relative sensitivity parameter $\gamma_i > 0$, such that $\mathbf{B}_i = \gamma_i \mathbf{B}$ for all i. Consumer groups are further differentiated by their product preference $\mathbf{a}_i$, their population size k_i, and the existing products currently owned $\mathbf{q}_i$. With the correlated demand elasticities, $\bar{\mathbf{B}}_i = \bar{\gamma}_i \mathbf{B}$, where $\bar{\gamma}_i = \gamma_i k_i$.

For convenience we introduce the normalization parameters $\hat{\gamma}_i = \bar{\gamma}_i / \sum_{i'} \bar{\gamma}_{i'}$ and the parameter $\hat{\mathbf{a}}_i = (1/\gamma_i)\mathbf{a}$. Before proceeding with the statements on uniqueness, it is convenient to reformulate the mixed LCP (4.3.5)-(4.3.7) into a standard LCP. Define the block vectors

$$\mathbf{w} = \begin{bmatrix} \mathbf{x}_1 + \mathbf{q}_1 & \mathbf{x}_2 + \mathbf{q}_2 & \dots & \mathbf{x}_m + \mathbf{q}_m \end{bmatrix}^T \tag{4.4.1}$$

$$\mathbf{r} = \begin{bmatrix} \bar{\mathbf{a}}_1 - \hat{\gamma}_1\left(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}}\right) & \bar{\mathbf{a}}_2 - \hat{\gamma}_2\left(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}}\right) & \dots & \bar{\mathbf{a}}_m - \hat{\gamma}_m\left(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}}\right) \end{bmatrix}^T,$$

and the block matrix

$$\mathbf{M} = \begin{bmatrix} \bar{\gamma}_1\mathbf{B} - \bar{\gamma}_1\hat{\gamma}_1\mathbf{B} & -\bar{\gamma}_1\hat{\gamma}_2\mathbf{B} & \cdots & -\bar{\gamma}_1\hat{\gamma}_m\mathbf{B} \\ -\bar{\gamma}_2\hat{\gamma}_1\mathbf{B} & \bar{\gamma}_2\mathbf{B} - \bar{\gamma}_2\hat{\gamma}_2\mathbf{B} & \cdots & -\bar{\gamma}_2\hat{\gamma}_m\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ -\bar{\gamma}_m\hat{\gamma}_1\mathbf{B} & -\bar{\gamma}_m\hat{\gamma}_2\mathbf{B} & \cdots & \bar{\gamma}_m\mathbf{B} - \bar{\gamma}_m\hat{\gamma}_m\mathbf{B} \end{bmatrix}.$$

Lemma 6. Finding a vector $\mathbf{s}$ which solves the LCP

$$\mathbf{w} = \mathbf{r} + \mathbf{M}\mathbf{s} \tag{4.4.2}$$

$$0 \leq \mathbf{w} \perp \mathbf{s} \geq 0 \tag{4.4.3}$$

is equivalent to solving problem (4.3.5)-(4.3.7) with uniform demand.

Proof. Substituting (4.3.8) into (4.3.5) and defining $\mathbf{w}_i = \mathbf{x}_i + \mathbf{q}_i$

$$\begin{aligned}\mathbf{w}_i &= \bar{\mathbf{a}}_i - \bar{\mathbf{B}}_i\bar{\bar{\mathbf{B}}}^{-1}\big(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}}\big) + \bar{\mathbf{B}}_i\left(\mathbf{s}_i - \sum_{i'}\bar{\bar{\mathbf{B}}}^{-1}\bar{\mathbf{B}}_{i'}\mathbf{s}_{i'}\right)\\ &= \bar{\mathbf{a}}_i - \hat{\gamma}_i\big(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}}\big) + \bar{\gamma}_i\mathbf{B}\left(\mathbf{s}_i - \sum_{i'}\hat{\gamma}_{i'}\mathbf{s}_{i'}\right). \qquad (4.4.4)\end{aligned}$$

Amalgamating (4.4.4) for all i leads to (4.4.2) and the complementarity condition (4.4.3) follows from substituting (4.4.1) into (4.3.6). □

Lemma 7. The complementarity problem (4.4.2)-(4.4.3) is feasible.

Proof. By definition, the LCP is feasible if and only if there exists vector $\hat{\mathbf{s}} \geq 0$ such that $\mathbf{w} \geq 0$. Rearranging (4.4.4) and emphasizing that $\mathbf{w}$ is a direct function of multipliers $\mathbf{s}$, it follows that

$$\begin{aligned}\mathbf{w}_i(\mathbf{s}) &= \bar{\mathbf{a}}_i - \hat{\gamma}_i\bar{\bar{\mathbf{a}}} + k_i\mathbf{B}\left(\mathbf{s}_i - \sum_{i'}\hat{\gamma}_{i'}\mathbf{s}_{i'}\right) + \hat{\gamma}_i\bar{\bar{\mathbf{q}}}\\ &= \bar{\mathbf{a}}_i - \hat{\gamma}_i\left(\sum_{i'}k_{i'}\mathbf{a}_{i'}\right) + k_i\mathbf{B}\left(\mathbf{s}_i - \sum_{i'}\hat{\gamma}_{i'}\mathbf{s}_{i'}\right) + \hat{\gamma}_i\bar{\bar{\mathbf{q}}}\\ &= \bar{\gamma}_i\left(\hat{\mathbf{a}}_i + \mathbf{B}\mathbf{s}_i - \sum_{i'}\hat{\gamma}_{i'}\hat{\mathbf{a}}_{i'} - \mathbf{B}\sum_{i'}\hat{\gamma}_{i'}\mathbf{s}_{i'}\right) + \hat{\gamma}_i\bar{\bar{\mathbf{q}}}\\ &= \bar{\gamma}_i\left(\hat{\mathbf{a}}_i + \mathbf{B}\mathbf{s}_i - \sum_{i'}\hat{\gamma}_{i'}\left(\hat{\mathbf{a}}_{i'} + \mathbf{B}\mathbf{s}_{i'}\right)\right) + \hat{\gamma}_i\bar{\bar{\mathbf{q}}} \qquad (4.4.5)\end{aligned}$$

It is apparent that if there exists a vector $\tilde{\mathbf{s}} \geq \mathbf{0}$ such that (4.4.5) is nonnegative for all i when $\bar{\bar{\mathbf{q}}} = 0$, then the LCP will be feasible for any $\bar{\bar{\mathbf{q}}} \in \Re^n_+$. Consider a vector $\hat{\hat{\mathbf{a}}} = [\hat{\hat{a}}_j]$ such that $\hat{\hat{a}}_j \geq \max_i \hat{a}_{ij}$ for all j and set $\tilde{\mathbf{s}}_i = \mathbf{B}^{-1}\big(\hat{\hat{\mathbf{a}}} - \hat{\mathbf{a}}_i\big)$ for all i. The vector $\hat{\mathbf{s}}_i$ is nonnegative since $\mathbf{B}^{-1} = \tilde{\mathbf{B}}$, which is positive definite, and $\big(\hat{\mathbf{a}} - \mathbf{a}_i\big) \geq \mathbf{0}$.

This implies that for any $\bar{\bar{\mathbf{q}}} \in \Re_{+}^{n}$ and for all i

$$\mathbf{w}_i(\tilde{\mathbf{s}}) \geq \bar{\gamma}_i \Bigg(\hat{\mathbf{a}}_i + \mathbf{B}\tilde{\mathbf{s}}_i - \sum_{i'} \hat{\gamma}_{i'} \Big(\hat{\mathbf{a}}_{i'} + \mathbf{B}\tilde{\mathbf{s}}_{i'} \Big) \Bigg) = \bar{\gamma}_i \Bigg(\hat{\hat{\mathbf{a}}} - \sum_{i'} \hat{\gamma}_{i'} \hat{\hat{\mathbf{a}}} \Bigg) = \mathbf{0}.$$

Thus, there exists a vector $\mathbf{s} \geq \mathbf{0}$ such that $\mathbf{w} \geq 0$ and the LCP is feasible. □

Lemma 8. The matrix $\mathbf{M}$ is symmetric positive semi-definite.

Proof. If a symmetric matrix is diagonally dominant with positive diagonal elements, then the matrix is positive semi-definite. Noting that $\mathbf{B}$ has positive diagonals, the sums of the off diagonals are equal the diagonal element of in each row, and $\hat{\gamma}_i k_j = k_i \hat{k}_j$, the matrix $\mathbf{M}$ is symmetric positive semi-definite. □

Remark 2. By Definition 3.4.6 in [52], a matrix $\tilde{\mathbf{M}} \in \Re^{n \times n}$ where all the principal minors are non-negative, is column adequate if the determinant of $\tilde{\mathbf{M}}_{\alpha\alpha} = 0$ for each $\alpha \subseteq \{1, \ldots, n\}$. The matrix is defined as adequate if the matrix and its transpose are both column adequate. Observing that any symmetric positive semi-definite matrix is adequate, $\mathbf{M}$ belongs to the class of matrices that are also adequate.

Proposition 11. The complementarity problem (4.3.5)-(4.3.7) has a unique solution of the quantities traded $\mathbf{x}^*$ in the exchange market.

Proof. From Corollary 3.5.6 of [52], if $\mathbf{M} \in \Re^{mn \times mn}$ is adequate and $\mathbf{r} \in \Re^{mn}$ is arbitrary, then if the LCP is feasible, there exist a unique vector $\mathbf{w}^*$ and a vector $\mathbf{s}$ satisfying

$$\mathbf{w}^* = \mathbf{r} + \mathbf{M}\mathbf{s} \geq \mathbf{0}, \qquad \mathbf{s} \geq \mathbf{0}, \qquad \mathbf{w}^{*T}\mathbf{s} = 0,$$

with no guarantees that the vector $\mathbf{s}$ is unique. If $\bar{\mathbf{s}}$ and $\hat{\mathbf{s}}$ are two solutions to the LCP, then $\mathbf{w}^* = \mathbf{r} + \mathbf{M}\bar{\mathbf{s}} = \mathbf{r} + \mathbf{M}\hat{\mathbf{s}} = \mathbf{w}^*$. Since $\mathbf{M}$ is adequate, as explained in Remark 2 and Lemma 7 shows that the LCP is feasible, there is a unique solution $\mathbf{x}^* = \mathbf{w}^* - \mathbf{q}$ which produces the equilibrium quantities traded in the exchange market. □

Although the $\mathbf{M}$ matrix belongs to a class which only guarantees w-uniqueness, the following proposition shows that the special structure of the problem leads to an equilibrium solution that is entirely unique.

Proposition 12. If $\bar{\bar{\mathbf{q}}} > \mathbf{0}$, then there is a unique set of market prices $\mathbf{p}^*$, which correspond to an equilibrium solution $\mathbf{x}^*$.

Proof. The null space of $\mathbf{M}$ is any vector $\mathbf{u} \in \Re^{mn}$ that solves $\mathbf{M}\mathbf{u} = \mathbf{0}$. Defining the matrices $\mathbf{\Gamma} = \text{diag}(\bar{\mathbf{B}}_1, \bar{\mathbf{B}}_2, \ldots, \bar{\mathbf{B}}_m)$ and

$$\tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{I} - \hat{\gamma}_1\mathbf{I} & -\hat{\gamma}_2\mathbf{I} & \cdots & -\hat{\gamma}_m\mathbf{I} \\ -\hat{\gamma}_1\mathbf{I} & \mathbf{I} - \hat{\gamma}_2\mathbf{I} & \cdots & -\hat{\gamma}_m\mathbf{I} \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\gamma}_1\mathbf{I} & -\hat{\gamma}_2\mathbf{I} & \cdots & \mathbf{I} - \hat{\gamma}_m\mathbf{I} \end{bmatrix},$$

we can express the matrix $\mathbf{M}$ as $\mathbf{\Gamma}\tilde{\mathbf{M}}$. From the invertibility of $\mathbf{\Gamma}$, the null space is simplified to $\text{Null}(\mathbf{M}) \equiv \{\mathbf{u} \in \Re^{mn} : \tilde{\mathbf{M}}\mathbf{u} = \mathbf{0}\}$. Consider an m-block vector $\mathbf{u} = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_m]$, where $\mathbf{u}_i \in \Re^n$. If $\mathbf{u}$ is in the null space of $\mathbf{M}$, then $\mathbf{u}_i = \sum_i \hat{\gamma}_i \mathbf{u}_i$ for all $i \in \mathcal{I}$. This implies that $\mathbf{u}_1 = \mathbf{u}_2 = \cdots = \mathbf{u}_m$, and that any vector in the null space must have the same value for each product entry across all consumers.

Consider any two solutions $\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}$ to the LCP such that $\mathbf{r} + \mathbf{M}\tilde{\mathbf{s}} = \mathbf{r} + \mathbf{M}\hat{\mathbf{s}} = \mathbf{w}^*$.

It follows that there exists a vector $\tilde{\mathbf{u}} \in \text{Null}(\mathbf{M})$ such that

$$\hat{\mathbf{s}} = \tilde{\mathbf{s}} + \tilde{\mathbf{u}}. \tag{4.4.6}$$

Given a unique solution $\mathbf{w}^*$ from the result of trading, for each product $j \in \mathcal{J}$ there exists at least one consumer i such that $w^*_{ij} > 0$, implying that there is at least one variable $\tilde{s}_{ij} = 0$ and $\hat{s}_{ij} = 0$ for each $j \in \mathcal{J}$. Since $\tilde{\mathbf{u}}$ must have the same entry $\tilde{u}_{ij}$ equalled to some $\tilde{u_j}$ for all consumer groups i, the only Null space that can satisfy (4.4.6) is $\tilde{\mathbf{u}} = \mathbf{0}$. Thus, $\tilde{\mathbf{s}} = \hat{\mathbf{s}}$ equals the unique solution $\mathbf{s}^*$. From Corollary 5, it follows that prices are unique and equal to

$$\mathbf{p}^* = \bar{\bar{\mathbf{B}}}^{-1}\big(\bar{\bar{\mathbf{a}}} - \bar{\bar{\mathbf{q}}} + \sum_i \bar{\mathbf{B}}_i \mathbf{s}^*_i\big).$$

□

4.5. Conclusion

This paper develops an exchange market for an arbitrary number of products and consumer types and establishes conditions that guarantee a unique equilibrium. The generality of the setup can allow future research to incorporate the model into a larger framework with secondary markets applications. For example, models that involve dynamics in consumer behaviour or consumer demographics will impact consumer preferences for products, which in turn lead to exchanges in C2C markets. Although the problem is configured for consumers, adjustment to the framework could allow the traders can be interpreted as firms participating in B2B exchanges. This would further broadens the applications to models involving trade within commodity markets and the movement of goods through supply chains.

Chapter 5

Summary and Conclusions

5.1. Summary

The first essay analyzes the impact of branding on the frequency and profitability of a firm's upgrade strategy. We formulate the timing of product upgrade releases as a Markov Decision Process and prove that the optimal upgrade strategy can be characterized by a threshold policy based on either pent-up demand for a future product release or the lag in technology of the incumbent product. The model is then generalized to account for the possibility that new products may not meet consumer expectations due to unforeseen manufacturing, design, or supply chain issues. When accounting for the potential of a product failure, a threshold policy based on pent-up demand remains optimal, whereas a technology based threshold policy is shown to be sub-optimal. Numerical experiments are utilized to examine market conditions in which the firm should invest in raising brand commitment. We also study how the optimal upgrade decision and the influence of branding changes with the potential of releasing a product failure.

The second essay develops an approximate optimal control problem to produce time-dependent bid prices for the airline network revenue management problem. The main contributions of our paper are the analysis of time-dependent bid prices in continuous time and the use of splines to modify the problem into an approximate second-order cone program (ASOCP). The spline representation of bid prices permits the number of variables to depend solely on the number of resources and not on the size of the selling window. The advantage of this framework is the ASOCP's scalability, which we demonstrate by solving for bid prices on an industrial sized network. The numerical experiments highlight the ASOCP's ability to solve industrial sized problems in seconds.

The third essay develops a general multi-product exchange market for gross substitute products, using multiple representative consumers who are endowed with an initial collection of products. The consumers are differentiated by their relative population size, as well as by their preference and willingness to pay for products. The consumers maximize surplus by participating in the exchange market deciding the quantity of products to purchase and sell. Under the assumption that consumers have correlated demand sensitivities to price, the quantity of products traded and the equilibrium prices is proven to be unique.

5.2. Future Work

My future research looks to expand the scope of these three models by incorporating additional factors faced by firms into the decision models. With regards to technology management, the aggregation of demand for a future product release has important implications for inventory procurement, in terms of volume discounts and greater certainty on customer demand. Factoring improved economies of scale into the decision-making process may provide further insights into the benefits of accumulating demand in anticipation of a future product launch at the expense of present eroding sales.

In terms of the network capacity control problem, the scalability of the ASOCP approach implies that the model will lend itself to efficiently solving bid prices with consumer choice. The inclusion of consumer behavior in the model will make the ASOCP applicable to other areas of revenue management, where consumer preference for products is an important component of the optimal control policy.

The exchange market can be applied to housing markets within a city and utilized to analyze sustainability issues in urban environments. There are environmental consequences resulting from a discrepancy between a city's housing stock and the dwelling needs of the population. This discrepancy can occur due to long-term changes in family demographics, combined with policies that promote the construction of a surplus of specific dwelling units. Thus a city's existing housing stock and long-term demographic trend datasets to the exchange model, in order to enhance the understanding of residential development policies on sustainable urban development. The model can be expanded to optimize public policy with regards to the built environment in order to minimize commuting.

My long-term objectives is to integrate components of the three approaches into unified models to address emergent problems in the areas of technology and revenue management. For example, I plan to develop a model to optimize product transitions in response to shifts in technology and consumer preferences. Further, I plan to analyze how product transitions strategies change, given the presence of a secondary market. The increasing ability of consumers to monetize the residual value of earlier generations of a product through online exchange markets should have a significant impact on the rate of technology adoption and the upgrade strategy of firms. My research in this emerging area will be primarily focused on developing a model that provides insight into allocation strategies over time, in response to shifts in consumer preferences for technological products with overlapping features.

Bibliography

[1] D. Adelman. Dynamic bid-prices in revenue management. *Operations Research*, 55(4):647–661, 2007.

[2] Elodie Adida and Victor DeMiguel. Supply chain competition with multiple manufacturers and retailers. *Operations Research*, 59(1):156–172, 2011.

[3] Ron Adner and Daniel Levinthal. Demand heterogeneity and technology evolution: implications for product and process innovation. *Management Science*, 47(5):611–628, 2001.

[4] Rohini Ahluwalia, Robert E Burnkrant, and H Rao Unnava. Consumer response to negative publicity: the moderating role of commitment. *Journal of marketing research*, 37(2):203–214, 2000.

[5] M. Akan and B. Ata. Bid-price controls for network revenue management: Martingale characterization of optimal bid prices. *Mathematics of Operations Research*, 34(4):912–936, 2009.

[6] F. Alizadeh, J. Eckstein, N. Noyan, and G. Rudolf. Arrival rate approximation by nonnegative cubic splines. *Operations Research*, 56(1):140–156, 2008.

[7] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Programming*, 95(1):3–51, 2003.

[8] Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954.

[9] Yves Balcer and Steven A. Lippman. Technological expectations and adoption of improved technology. *Journal of Economic Theory*, 34(2):292–318, 1984.

[10] Sreekumar R. Bhaskaran and Karthik Ramachandran. Managing technology selection and development risk in competitive environments. *Production and Operations Management*, 20(4):541–555, 2011.

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

[12] P. Bremaud. *Point Processes and Queues.* Springer-Verlag, New York, 1981.

[13] Abram Brown. Apple's new iphone launch was a record-breaker: 9 million sold in 3 days. *Forbes*, 2013. http://www.forbes.com/sites/abrambrown/2013/09/23/apples-new-iphone-launch-was-a-record- breaker-9-million-sold-in-3-days.

[14] Yubo Chen and Jinhong Xie. Third-party product review and firm marketing strategy. *Marketing Science*, 24(2):218–240, 2005.

[15] Soo-Haeng . H. Cho and Kevin F. McCardle. The adoption of multiple dependent technologies. *Operations research*, 57(1):157–169, 2009.

[16] Clayton M Christensen, Fernando F Suárez, and James M Utterback. Strategies for survival in fast-changing industries. *Management science*, 44(12-part-2):S207–S220, 1998.

[17] Wujin Chu, Mee Ryoung Song, and Beomjoon Choi. Post-purchase disadvantages of a less preferred brand and how they can be overcome: an examination of regret and attribution. *Journal of Applied Social Psychology*, 43(4):887–898, 2013.

[18] Cathy J Cobb-Walgren, Cynthia A Ruble, and Naveen Donthu. Brand equity, brand preference, and purchase intent. *Journal of Advertising*, 24(3):25–40, 1995.

[19] Morris A. Cohen, Jehoshua Eliasberg, and Teck-Hua . H. Ho. New product development: The performance and time-to-market tradeoff. *Management Science*, 42(2):173–186, 1996.

[20] Dipankor Coondoo. Multiple representative consumers. *Economics Letters*, 22(2):119–122, 1986.

[21] W. Cooper. Asymptotic behavior of an allocation policy for revenue management. *Operations Research*, 50:720–727, 2002.

[22] Richard Dobbs, Yougang Chen, Gordon Orr, James Manyika, Elisie Chang, and McKinsey Global Institute. *China's E-tail Revolution: Online Shopping as a Catalyst for Growth*. McKinsey Global Institute, 2013.

[23] Franz-Rudolf Esch, Tobias Langner, Bernd H Schmitt, and Patrick Geus. Are brands forever? how brand knowledge and relationships affect current and future purchases. *Journal of Product & Brand Management*, 15(2):98–105, 2006.

[24] Amr Farahat and Georgia Perakis. A nonnegative extension of the affine demand function and equilibrium analysis for multiproduct price competition. *Operations Research Letters*, 38(4):280–286, 2010.

[25] Amr Farahat and Georgia Perakis. Technical note—a comparison of bertrand and cournot profits in oligopolies with differentiated products. *Operations research*, 59(2):507–513, 2011.

[26] Awi Federgruen and Ming Hu. Sequential multi-product price competition in supply chain networks. *Rotman School of Management working paper*, 2013.

[27] Keith Goffin and Colin New. Customer support and new product development-an exploratory study. *International Journal of Operations & Production Management*, 21(3):275–301, 2001.

[28] Neil Hughes. Apple's iphone led 2013 us consumer smartphone sales with 45npd. *appleinsider.com*, 2014. http://appleinsider.com/articles/14/02/20/apples-iphone-led-2013-us-consumer-smartphone-sales-with-45-share—npd.

[29] Kana Inagaki. Sony tries to crack smartphone market with xperia z2: New phone has 4k camera but will it be enough? *Forbes*, 2014. http://online.wsj.com/news/articles/SB10001424052702303636404579397103483908052.

[30] Stefanus Jasin and Sunil Kumar. Analysis of deterministic LP-based booking limit and bid price controls for revenue management. *Operations Research*, 61(6):1312–1320, 2013.

[31] Ralph Jennings. Slower speed stunts sales of htc's supersized smartphones. *Forbes*, 2014. http://www.forbes.com/sites/ralphjennings/2014/02/20/slower-speed-stunts-sales-of-htcs-supersized-smartphones-taipei-consultant/.

[32] Michael D. Johnson, Andreas Herrmann, and Frank Huber. The evolution of loyalty intentions. *Journal of Marketing*, 70(2):122–132, 2006.

[33] Soulaymane Kachani and Kyrylo Shmatov. Competitive pricing in a multi-product multi-attribute environment. *Production and Operations Management*, 20(5):668–680, 2011.

[34] S. Karlin and W. Studden. *Tchebycheff Systems, with Applications in Analysis and Statistics*. Wiley, New York, 1966.

[35] Jonathan Kluberg and Georgia Perakis. Generalized quantity competition for multiple products and loss of efficiency. *Operations Research*, 60(2):335–350, 2012.

[36] R. Mark Krankel, Izak Duenyas, and Roman Kapuscinski. Timing successive product introductions with demand diffusion and stochastic technology improvement. *Manufacturing & Service Operations Management*, 8(2):119–135, 1 2006.

[37] S. Kunnumkal and H. Topaloglu. A refined deterministic linear program for the network revenue management problem with customer choice behavior. *Naval Research Logistics*, 55(6):563–580, 2008.

[38] S. Kunnumkal and H. Topaloglu. Computing time-dependent bid prices in network revenue management problems. *Transportation Science*, 44(1):38–62, 2010.

[39] Son K Lam, Michael Ahearne, Ye Hu, and Niels Schillewaert. Resistance to brand switching when a radically new brand is introduced: A social identity theory perspective. *Journal of Marketing*, 74(6):128–146, 2010.

[40] John Laugesen and Yufei Yuan. What factors contributed to the success of apple's iphone? In *Mobile Business and 2010 Ninth Global Mobility Roundtable (ICMB-GMR), 2010 Ninth International Conference on*, pages 91–99. IEEE, 2010.

[41] Hongmin Li and Stephen C. Graves. Pricing decisions during inter-generational product transition. *Production and Operations Management*, 21(1):14–28, 1 2012.

[42] E. Lieb and M. Loss. *Analysis*. American Mathematical Society, 2001.

[43] Heng Liu and Özalp Özer. Managing a product family under stochastic technological changes. *International Journal of Production Economics*, 122(2):567–580, 12 2009.

[44] Q. Liu and G. Van Ryzin. On the choice-based linear programming model for network revenue management. *Manufacturing and Service Operations Management*, 10(2):288–310, 2008.

[45] Ilan Lobel, Jigar Patel, Gustavo Vulcano, and Jiawei Zhang. Optimizing product launches in the presence of strategic consumers. *Available at SSRN 2222266*, 2014.

[46] P. Loewen and R. Rockafellar. Optimal control of unbounded differential inclusions. *SIAM J. Control Optim.*, 32(2):442–470, 1994.

[47] Marco López and Georg Still. Semi-infinite programming. *European Journal of Operational Research*, 180(2):491–518, 2007.

[48] J. Meissner and A. Strauss. Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research*, 216(2):459–468, 2012.

[49] Marc H Meyer, Peter Tertzakian, and James M Utterback. Metrics for managing research and development in the context of the product family. *Management Science*, 43(1):88–111, 1997.

[50] Leslie Olin Morgan, Ruskin M. Morgan, and William L. Moore. Quality and time-to-market trade-offs when there are multiple product generations. *Manufacturing & Service Operations Management*, 3(2):89–104, 2001.

[51] Rajesh Nellore and R Balachandra. Factors influencing success in integrated product development (ipd) projects. *Engineering Management, IEEE Transactions on*, 48(2):164–174, 2001.

[52] JS Pang, RE Stone, and RW Cottle. The linear complementarity problem. *Computer Science and Scientific Computing. Academic Press*, 1:2, 1992.

[53] Steven S Posavac, David M Sanbonmatsu, Frank R Kardes, and Gavan J Fitzsimons. The brand positivity effect: When evaluation confers preference. *Journal of Consumer Research*, 31(3):643–651, 2004.

[54] Andreas Pufall, Jan C Fransoo, Ad de Jong, and Ton de Kok. The impact of product complexity on ramp-up performance. Technical report, BETA Working Paper, 2012.

[55] JMW Rhys. A selection problem of shared fixed costs and network flows. *Management Science*, 17(3):200–207, 1970.

[56] Michelle L Roehm and Michael K Brady. Consumer responses to performance failures by high-equity brands. *Journal of Consumer Research*, 34(4):537–545, 2007.

[57] Sean Silcoff, Jacquie McNish, and Steve Ladurantaye. Inside the fall of blackberry: How the smartphone inventor failed to adapt. *Globe and Mail*, 10 2013. http://www.theglobeandmail.com/report-on-business/the-inside-story-of-why-blackberry-is-failing/article14563602/?page=all.

[58] WM Soon, GY Zhao, and JP Zhang. Complementarity demand functions and pricing models for multi-product markets. *European Journal of Applied Mathematics*, 20(5):399–430, 2009.

[59] Gilvan C. Souza. Product introduction decisions in a duopoly. *European Journal of Operational Research*, 152(3):745–757, 2004.

[60] G. Still. Discretization in semi-infinite programming: the rate of convergence. *Mathematical Programming*, 91(1):53–69, 2001.

[61] James Surowiecki. Where nokia went wrong. *New Yorker*, 2013. http://www.newyorker.com/business/currency/where-nokia-went-wrong.

[62] K. Talluri and G. van Ryzin. *The Theory and Practice of Revenue Management.* Kluwer Academic Publishers, Norwell, MA, 2004.

[63] Mohan V Tatikonda and Stephen R Rosenthal. Technology novelty, project complexity, and product development project execution success: a deeper look at task

uncertainty in product innovation. *Engineering Management, IEEE Transactions on*, 47(1):74–87, 2000.

[64] C. Tong and H. Topaloglu. On approximate linear programming approach for network revenue management problems. *INFORMS Journal on Computing*, 26:121–134, 2014.

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

[66] R. Vinter. *Optimal control*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, MA, 2000.

[67] Siva Viswanathan. Competing across technology-differentiated channels: The impact of network externalities and switching costs. *Management Science*, 51(3):483–496, 2005.

[68] T. Vossen and D. Zhang. A dynamic disaggregation approach to approximate linear programs for network revenue management. *Production and Operations Management*, page doi 10.1111/poms.12239, 2014.

[69] Oliver E Williamson. The vertical integration of production: market failure considerations. *The American Economic Review*, pages 112–123, 1971.

[70] H. Wolkowicz, L. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*. Kluwer, Dordrecht, The Netherlands, 2000.

[71] Wentian Xie and Helen Liang. A case study: Innovation strategy assessment of the leading smartphone companies. In *Innovation Conference (SIIC), 2013 Suzhou-Silicon Valley-Beijing International*, pages 121–126. IEEE, 2013.

[72] Zhilin Yang and Robin T Peterson. Customer perceived value, satisfaction, and loyalty: the role of switching costs. *Psychology & Marketing*, 21(10):799–822, 2004.

[73] D. Zhang and D Adelman. An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Science*, 42(3):381–394, 2009.

Appendix A

Timing Product Upgrades with Stochastic Technological Advancements, Brand Commitment, and Product Failure

A.1. Technical Appendix

Proof of Lemma 1

To show the monotonicity in the state component n_t, assume that $n_t^1 < n_t^2$ for some $t \in \{t_0 + 1, \dots T\}$. Independent of the upgrade decision, $n_{t+1}^1 = n_t^1 - a(n_t^1) + g(\xi) \leq n_t^2 - a(n_t^2) + g(\xi) = n_{t+1}^2$, where the inequality follows form the inductive assumption and the fact that $\frac{da(n_t)}{dn_t} \leq 1$. To show the monotonicity in the state component d_t for systems $(d_{t_0}^1, n_{t_0}^1, z_{t_0})$ and $(d_{t_0}^2, n_{t_0}^2, z_{t_0})$, consider a policy where both firms make product upgrades in the same periods. In period t_0, the levels of pent-up demand for systems 1 and 2 are $d_{t_0}^1$ and $d_{t_0}^2$, respectively. Since $d_{t_0}^1 \leq d_{t_0}^2$, assume that $d_t^1 \leq d_t^2$ holds for some $t \in \{t_0, \dots, T\}$. Thus, in period $t + 1$, if the firm does not upgrade $d_{t+1}^1 = \theta\big(d_t^1 + a(n_t^1)\bar{\rho}(z_t)\big) \leq \theta\big(d_t^2 + a(n_t^2)\bar{\rho}(z_t)\big) = d_{t+1}^2$, where the inequality holds due to the inductive assumption, the fact that the dynamics of n_t and z_t are independent of pent-up demand, and that $n_t^1 \leq n_t^2$ for all time t. If the firm upgrades in period t then $d_{t+1}^1 = d_{t+1}^2 = 0$ and the result $d_{t+1}^1 = \theta a(n_t^1)\bar{\rho}(z_t) \leq \theta a(n_t^2)\bar{\rho}(z_t) = d_{t+1}^2$ holds. This implies that $d_t^1 \leq d_t^2$ over the interval $t \in \{t_0, \ldots, T+1\}$. The same inductive reasoning is applicable to prove that $z_t^1 \leq z_t^2$, for all $t \in \{t_0, \ldots, T+1\}$. In period t make the inductive assumption that $z_t^1 \leq z_t^2$. If the firms do not upgrade in period $t+1$, then $z_{t+1}^1 = z_t^1 + \xi_t \leq z_t^2 + \xi_t = z_{t+1}^2$. If the firm's do upgrade in period $t+1$, then $z_{t+1}^1 = \xi_t = z_{t+1}^2$ and the result trivially holds.

Proof of Proposition 1

In period t_0, consider technology lag z_{t_0}, serviceable market $n_{t_0}^1$ and $n_{t_0}^2 = n_{t_0}^1 + \epsilon_n$, and the delayed demands $d_{t_0}^1$ and $d_{t_0}^2$, where $d_{t_0}^2 = d_{t_0}^1 + \epsilon_d$, for $\epsilon_n, \epsilon_d \geq 0$. We compare the two system states $(d_{t_0}^1, n_{t_0}^1, z_{t_0})$ and $(d_{t_0}^2, n_{t_0}^2, z_{t_0})$ denoted by superscripts 1 and 2, where both systems experience the same sample path of technology advancement. Let $\tilde{\mu}_1$ denote the optimal upgrade policy for a system starting with state $d_{t_0}^1$ and suppose that the policy of system 2 is to mimic the decisions produced by $\tilde{\mu}_1$ under each realization of technology advancement. Let $V_{t_0,\tilde{\mu}_1}^1$ and $V_{t_0,\tilde{\mu}_1}^2$ be random variables representing the present value of profit at t_0 for a given realization under the upgrade policy $\tilde{\mu}_1$ for systems 1 and 2. Define $\bar{x}_t$ as the binary decision by firm 1 to upgrade, where $\bar{x}_t = 1$ indicates that the firm releases an upgrade in period t. If we consider such a sample path of technological advancement, then

$$\begin{aligned} V^1_{t_0,\tilde{\mu}_1} &= \sum_{t=t_0}^{T} \delta^{t-t_0} \left[\pi a(n^1_t)\rho(z_t) + \bar{x}_t \left(\pi \left(a(n^1_\tau)\bar{\rho}(z_t) + d^1_\tau \right) - K \right) \right] \\ &\leq \sum_{t=t_0}^{T} \delta^{t-t_0} \left[\pi a(n^2_t)\rho(z_t) + \bar{x}_t \left(\pi \left(a(n^2_\tau)\bar{\rho}(z_t) + d^2_\tau \right) - K \right) \right] \\ &= V^2_{t_0,\tilde{\mu}_1} \end{aligned}$$

where the inequality follows from the state dynamics described by Lemma 1. Since $n^1_t \leq n^2_t$ for all t the profit from arrivals is always such that $\pi a(n^1_t) \leq \pi a(n^2_t)$ and since $d^1_t \leq d^2_t$ for all $n^1_t \leq n^2_t$, the profit from delayed demand is greater or equal for the firm with the initial larger market or the initial greater delayed demand. Since $\tilde{\mu}_1$ is optimal for system 1 but suboptimal for system 2, $V_{t_0}(d^1_{t_0}, n_{t_0}, z_{t_0}) = E[V^1_{t_0,\tilde{\mu}_1}] \leq E[V^2_{t_0,\tilde{\mu}_1}] \leq V^2_{t_0}(d^2_{t_0}, n_{t_0}, z_{t_0})$ and monotonicity on d and n holds. A similar argument can be constructed for technology lag z_t using a mimicking policy argument, where the firm with a smaller technology lag copies the upgrade decision of the firm with a greater technology lag. First, the firm with a greater technology lag will lose more sales between between t_0 and the time of the first upgrade or the end of the planning horizon if there is no upgrade. Second, the sales resulting from greater pent-up demand due to a greater technology lag occur after the corresponding sales for the firm with a smaller technology lag and contribute less to present value of the profit. Third, after the first upgrade each firm will have the same profit level going forward since the state dynamics on serviceable market and pent-up demand are independent of the technology lag.

Proof of Proposition 2

Again consider two systems, system 1 with starting state $(d^1_{t_0}, n_{t_0}, z_{t_0})$ and system 2 with starting state $(d^2_{t_0}, n_{t_0}, z_{t_0})$, that experience the same sample path of technological advancement. Define the optimal introduction policy for system 2 as $\tilde{\mu}_2$, the random variable τ as the time of the first product upgrade since t_0 under policy $\tilde{\mu}_2$, and $f(K)$ as is the total cost under policy $\tilde{\mu}_2$. Under $\tilde{\mu}_2$ and a sample path of technology, $\tilde{\Pi}^i_t$ is the profit in period $t \geq t_0$ for system i. Let system 1's upgrade policy be $\tilde{\mu}_2$. i.e. system 1 mimics the upgrade strategy of system 2 irrespective of the state components. Denote $\Delta d_t = d^2_t - d^1_t$, Lemma 1 implies that $0 \leq \Delta d_{t+1} = \theta\big(\Delta d_t\big) \leq \Delta d_t, \quad \forall t \in \{t_0, \ldots, T\}$. Since the systems have the same starting states in terms of n_t and z_t, it is clear that $0 \leq \Delta d_t \leq \Delta d_{t_0}$ for all $t \geq t_0$. Let $V^1_{t_0,\tilde{\mu}_2}$ and $V^2_{t_0,\tilde{\mu}_2}$ be random variables representing the present value of profit at t_0 using policy $\tilde{\mu}_2$. Given that $\tilde{\Pi}^1_\tau = \tilde{\Pi}^2_\tau - \pi\Delta d_\tau$ and $\tilde{\Pi}^1_t = \tilde{\Pi}^2_t$ for all $t \neq \tau$,

$$\begin{aligned}
V^2_{\tilde{\mu}_2,t_0} &= \sum_{t=t_0}^{\tau-1} \delta^{t-t_0}\tilde{\Pi}^2_t + \delta^{\tau-t_0}\tilde{\Pi}^2_\tau + \sum_{t=\tau+1}^{T} \delta^{t-t_0}\tilde{\Pi}^2_t - f(K) \\
&= \sum_{t=t_0}^{\tau-1} \delta^{t-t_0}\tilde{\Pi}^1_t + \delta^{\tau-t_0}\left(\tilde{\Pi}^1_\tau + \pi\Delta d_\tau\right) + \sum_{t=\tau+1}^{T} \delta^{t-t_0}\tilde{\Pi}^1_t - f(K) \\
&\leq V^1_{\tilde{\mu}_2,t_0} + \pi\Delta d_{t_0}
\end{aligned}$$

Since the policy $\tilde{\mu}_2$ was optimal for System 2 and potentially suboptimal for System 1, the inequality $V_{t_0}(d^2_{t_0}, n_{t_0}, z_{t_0}) \leq V_{t_0}(d^1_{t_0}, n_{t_0}, z_{t_0}) + \pi(d^2_{t_0} - d^1_{t_0})$ holds.

Proof of Proposition 3

For states (d_t, n_t, z_t) if the firm invests in a product upgrade in period t, then the firm's profit-to-go is

$$I_t(d_t, n_t, z_t) = \pi\big(d_t + a(n_t)\big) - K + \delta E\big[V_{t+1}(0, n_t - a(n_t) + g(\xi), \xi)\big] \qquad \text{(A.1.1)}$$

If the firm does not upgrade the product, then the firm's profit-to-go is

$$W_t(d_t, n_t, z_t) = \pi a(n_t)\rho(z_t) + \delta E\big[V_{t+1}(\theta(d_t + a(n_t)\bar{\rho}(z_t)), n_t - a(n_t) + g(\xi), z_t + \xi)\big] \qquad \text{(A.1.2)}$$

The proposition holds true if at each time t, $\frac{\partial}{\partial d}\big(I_t(d_t, n_t, z_t) - W_t(d_t, n_t, z_t)\big) > 0$ almost everywhere in d for all n_t and z_t, since the marginal benefit of investing in the next product introduction would be increasing as the number of consumers with delayed decisions increases. Therefore, monotonicity in d would imply that if the firm would delay an upgrade at $\tilde{d}$, then the firm would delay the upgrade for any pent-up demand $d \leq \tilde{d}$.

Indeed, for any feasible tuple (d_t, n_t, z_t) in period t, if the firm upgrades in the current period, then the pent-up demand in period $t+1$ is 0. Thus, d_t does not affect the future expectation of the profit-to-go given an upgrade, and $\frac{\partial}{\partial d_t} I_t(d_t, n_t, z_t) = \pi$ for any value of d_t. Consequently, the condition that $\frac{\partial}{\partial d_t} W_t(d_t, n_t, z_t) < \pi$ is sufficient to prove monotonicity in d, which we demonstrate by induction.

Using the condition $V_{T+1}(d_{T+1}, n_{T+1}, z_{T+1}) = 0$ for all $(d_{T+1}, n_{T+1}, z_{T+1})$, the profit-to-go given the decision to wait is $W_T(d_T, n_T, z_T) = \pi a(n_T)\rho(z_T)$, which trivially implies that $\frac{\partial}{\partial d_T} W_T(d_T, n_T, z_T) < \pi$. Assume for some $t < T$ that $\frac{\partial}{\partial d_t} W_t(d_t, n_t, z_t) <$

π. Since, $V_t(d_t, n_t, z_t) = \max\big(I_t(d_t, n_t, z_t), W_t(d_t, n_t, z_t)\big)$, based on the inductive assumption on the $\frac{\partial}{\partial d_t} W_t(d_t, n_t, z_t) < \pi$, the fact that $\frac{\partial}{\partial d_t} I_t(d_t, n_t, z_t) = \pi$, and the continuity of both functions, there exists at most one point $\tilde{d}$ (a set of measure zero), where $I_t(d_t, n_t, z_t) = W_t(d_t, n_t, z_t)$. Defining the point where $I_t(d_t, n_t, z_t) = W_t(d_t, n_t, z_t)$ as $D_t^*(n_t, z_t)$, implies that it is optimal for the firm to delay the upgrade for $d_t < D_t^*(n_t, z_t)$ and for the firm to upgrade at $d_t > D_t^*(n_t, z_t)$. This implies that $\frac{\partial}{\partial d_t} V_t(d_t, n_t, z_t) < \pi$ for $d_t < D_t^*(n_t, z_t)$ and $\frac{\partial}{\partial d_t} V_t(d_t, n_t, z_t) = \pi$ for $d_t \geq D_t^*(n_t, z_t)$, and thus $\frac{\partial}{\partial d_t} V_t(d_t, n_t, z_t) \leq \pi$ almost everywhere for all d_t. Let the probability of an increase of ξ in the performance of leading-edge technology occur with probability $\lambda(\xi)$, where $\sum_\xi \lambda(\xi) = 1$. Since

$$\begin{aligned} W_{t-1}(d_{t-1}, n_{t-1}, z_{t-1}) &= \pi a(n_{t-1})\rho(z_{t-1}) \\ &\quad + \delta \sum_\xi \lambda(\xi) V_{t-1}(\theta\big(d_{t-1} + a(n_{t-1})\bar{\rho}(z_{t-1})\big), n_{t-1} - a(n_{t-1}) + g(\xi), z_{t-1} + \xi) \end{aligned}$$

it follows that $W_{t-1}(d_{t-1}, n_{t-1}, z_{t-1})$ is piecewise differentiable with a finite number of points of non-differentiability. Thus, using the chain rule,

$$\begin{aligned} \frac{\partial W_{t-1}(d_{t-1}, n_{t-1}, z_{t-1})}{\partial d_{t-1}} &= \delta\theta \sum_\xi \lambda(\xi) \frac{\partial V_t(\theta\big(d_{t-1} + a(n_{t-1})\bar{\rho}(z_{t-1})\big), n_{t-1} - a(n_{t-1}) + g(\xi), z_{t-1} + \xi)}{\partial d_t} \\ &\leq \delta\theta \sum_\xi \lambda(\xi)\pi = \delta\theta\pi \end{aligned}$$

which combined with the fact that $\delta \in [0, 1)$ implies that $\frac{\partial}{\partial d} W_{t-1}(d_t, n_t, z_t) < \pi$. Therefore the induction holds and there exists a threshold value $D_t^*(n_t, z_t)$, which determines the optimal upgrade policy in any period t.

Proof of Proposition 4

We begin by proving monotonicity of the threshold in z_t. Demonstrating $\Delta_{z_t} D_t^*(n_t, z_t) \leq 0$ is equivalent to showing that $\Delta_{z_t} W_t(d_t, n_t, z_t) = W_t(d_t, n_t, z_t+1) - W_t(d_t, n_t, z_t) \leq 0$, since $\Delta_{z_t} I_t(d_t, n_t, z_t) = I_t(d_t, n_t, z_t + 1) - I_t(d_t, n_t, z_t) = 0$. Using the condition $V_{T+1}(d_{T+1}, n_{T+1}, z_{T+1}) = 0$ for all $(d_{T+1}, n_{T+1}, z_{T+1})$, given the decision to wait in period T, $\Delta_{z_T} W_T(d_T, n_T, z_T) = \pi a(n_T)(\rho(z_T+1) - \rho(z_T)) \leq 0$, since sales are nonincreasing in the technological lag (Assumption 1), implying that $\Delta_{z_T} V_T(d_T, n_T, z_T) \leq 0$. Assume that $\Delta_{z_t} W_t(d_t, n_t, z_t) \leq 0$ for some $t \leq T$, which in turn implies that $\Delta_{z_t} V_t(d_t, n_t, z_t) \leq 0$. Since $W_{t-1}(d_{t-1}, n_{t-1}, z_{t-1}) = \pi a(n_{t-1})\rho(z_{t-1}) + \delta \sum_{\xi} \lambda(\xi) V_t(d_t, n_t, z_t + \xi)$ and $\pi(a(n_{t-1})\rho(z_{t-1} + 1) - a(n_{t-1})\rho(z_{t-1})) \leq 0$, using the inductive assumption, $\Delta_{z_{t-1}} W_{t-1}(d_{t-1}, n_{t-1}, z_{t-1}) \leq 0$.

Proof of Proposition 5

The proof for an optimal policy of upgrading if and only if $d \geq D(f_t, n_t, z_t)$ follows the same logic as the proof for Proposition 3. The marginal value of pent-up demand for the upgrade decision is $\frac{\partial}{\partial d_t} I_t(d_t, f_t, n_t, z_t) = \pi \sum_j q_j(z_t) p(j)$ for any value of d_t, so the condition that $\frac{\partial}{\partial d_t} W_t(d_t, f_t, n_t, z_t) < \pi \sum_j q_j(z_t) p(j)$ is sufficient to prove monotonicity in d. Using $\frac{\partial}{\partial d_t} W_{T+1}(d_{T+1}, f_{T+1}, n_{T+1}, z_{T+1}) = 0$ provides the inductive assumption that $\frac{\partial}{\partial d_t} W_t(d_t, f_t, n_t, z_t) \leq \pi \sum_{T+1} q_{T+1}(z_t) p(T + 1)$ which using the argument that there exists at most one point $\tilde{d}$ (a set of measure zero) where $I_t(d_t, f_t, n_t, z_t) = W_t(d_t, f_t, n_t, z_t)$ implies that $\frac{\partial}{\partial d_t} V_t(d_t, f_t, n_t, z_t) \leq \pi \sum_j q_j(z_t) p(j)$. Since $d_t = \theta\big(d_{t-1} +$

$p(f_{t-1})a(n_{t-1})\bar{\rho}(z_{t-1})\big)$, by the chain rule,

$$\begin{aligned}\frac{\partial}{\partial d_{t-1}} W_{t-1}(d_{t-1}, f_{t-1}, n_{t-1}, z_{t-1}) &= \delta\theta \sum_{\xi} \lambda(\xi) \frac{\partial}{\partial d_t} V_t(d_t, f_t, n_t(\xi), z_t(\xi)) \\ &\leq \delta\pi \sum_j q_j(z_t)p(j) < \pi \sum_j q_j(z_t)p(j),\end{aligned}$$

which completes the induction.

A.2. Premium Smartphone and Features

This appendix provides an overview of the timing and specifications of premium Android smartphones released by Samsung (Sam.), Motorola (Mot.), LG, HTC and Sony from October 2011 to September 2013 through one or more of the following US carriers: AT&T; T-Mobile; Sprint; and/or Verizon in direct competition to the iPhone. Each smartphone included in the list has front and rear cameras, 4G LTE support, and was initially priced at or above $199 from at least one of the above carriers based on a two year contract. Data was compiled from OEM and carrier websites.

Table 3: Premium Android smartphones released by Samsung, Motorola, LG, HTC and Sony from Oct-2011 to Sep-2013

OEM	Model	Released	Cores	CPU (GHz)	RAM (Mb)	Size (in.)	Density (Pix/in^2)	Screen (Pix)	Camera (MP)	Battery (mAh)
Sam.	Galaxy S II	Oct-11	dual	1.2	1024	4.3	218	800x480	8	1650
HTC	Amaze 4G	Oct-11	dual	1.5	1024	4.3	256	960x540	8	1730
Sam.	Galaxy Note	Oct-11	dual	1.4	1024	5.3	285	1280x800	8	2500
Sam.	Galaxy S2 Skyrocket HD	Nov-11	dual	1.5	1024	4.7	316	1280x720	8	1850
HTC	Vivid	Nov-11	dual	1.2	1024	4.5	245	960x540	8	1620
Mot.	Droid Razr	Nov-11	dual	1.2	1024	4.3	256	960x540	8	1780
HTC	Rezound	Nov-11	dual	1.5	1024	4.3	342	1280x720	8	1620
LG	Nitro HD	Dec-11	dual	1.5	1024	4.5	326	1280x720	8	1830
Sam.	Galaxy Nexus CDMA	Dec-11	dual	1.2	1024	4.7	316	1280x720	5	1750
LG	Spectrum	Jan-12	dual	1.5	1024	4.5	326	1280x720	8	1830
Mot.	Droid Razr Maxx	Jan-12	dual	1.2	1024	4.3	256	960x540	8	3300
Mot.	Droid 4	Feb-12	dual	1.2	1024	4	275	960x540	8	1785
HTC	One S	Apr-12	dual	1.5	1024	4.3	256	960x540	8	1650
HTC	One X	May-12	quad	1.5	1024	4.7	312	1280x720	8	1800
HTC	Evo 4G LTE	May-12	dual	1.5	1024	4.7	312	1280x720	8	2000
Sam.	Galaxy S3	May-12	quad	1.4	1024	4.8	306	1280x720	8	2100
HTC	Droid Incredible 4G LTE	Jul-12	dual	1.2	1024	4	275	960x540	8	1700
Mot.	Photon Q 4G	Aug-12	dual	1.5	1024	4.3	256	960x540	8	1785
LG	Intuition	Sep-12	dual	1.5	1024	5	256	1024x768	8	2080
Sam.	Galaxy Note II	Sep-12	quad	1.6	2048	5.5	265	1280x720	8	3100
Mot.	Droid Razr Maxx HD	Oct-12	dual	1.5	1024	4.7	312	1280x720	8	3300
Mot.	Droid Razr HD	Oct-12	dual	1.5	1024	4.7	312	1280x720	8	2530
LG	Optimus G	Nov-12	quad	1.5	2048	4.7	318	768x1280	13	2100
HTC	One X+	Nov-12	quad	1.7	1024	4.7	312	1280x720	8	2100
HTC	Droid DNA	Nov-12	quad	1.5	2048	5	441	1920x1080	8	2020
HTC	One	Apr-13	quad	1.7	2048	4.7	468	1920x1080	4*	2300
Sam.	Galaxy S4	Apr-13	quad	1.9	2048	5	441	1920x1080	13	2600
LG	Optimus G Pro	May-13	quad	1.7	2048	5.5	401	1920x1080	13	3140
Sam.	Galaxy S4 Active	Jun-13	quad	1.9	2048	5	443	1920x1080	8	2600
Mot.	Droid Ultra	Aug-13	dual	1.7	2048	5	294	1280x720	10	2130
Mot.	Droid Maxx	Aug-13	dual	1.7	2048	5	294	1280x720	10	3500
Mot.	Moto X	Aug-13	dual	1.7	2048	4.7	316	1280x720	10	2200
LG	G2	Sep-13	quad	2.2	2048	5.2	423	1920x1080	13	3000
Sam.	Galaxy Note 3	Sep-13	quad	2.3	3072	5.7	386	1920x1080	13	3200

*HTC One utilizes a CMOS sensor, which captures light equivalent to an 8-13a MP camera.

Appendix B

Scalable Dynamic Bid Prices for Network Revenue Management in Continuous Time

B.1. Technical Appendix - Network Revenue Management

General Optimality Conditions for Control of Differential Inclusions

Here, we summarize optimality conditions for the optimal control problem (P) in [46]. Our problem is a special case. To streamline application and analysis, we somewhat reduce the generality and omit all technical conditions which become trivial. The problem is to find an absolutely continuous function $y : [0, T] \to \mathbb{R}^k$ (typically referred to as an *arc*) which solves:

$$\min \quad \gamma(y(0)) \tag{B.1.1}$$

$$\text{s.t.} \quad \dot{y}(t) \in F(t, y(t)) \text{ a.e. } t \in [0, T] \tag{B.1.2}$$

$$y(T) \in S. \tag{B.1.3}$$

This formulation uses notation $\dot{y}(t)$ for the derivative of $y(t)$. In our special case, the objective $\gamma(\cdot)$ is linear and $F(t, y)$, S are nonempty, closed and convex. Multifunction $F(t, y)$ is given as an intersection of a finite number of linear constraints whose coefficients are Lipschitz-continuous functions of time. This implies that all required technical conditions are satisfied.

The statement of necessary optimality conditions (Theorem 4.3) uses the Hamiltonian function defined as

$$H(t, y, z) = \sup\{\langle z, d\rangle : d \in F(t, y)\}, \tag{B.1.4}$$

where $\langle z, d\rangle$ denotes the inner product of $z, d \in \mathbb{R}^k$. Suppose that an arc y^* solves the problem above. Then there exists a scalar $z^0 \in \{0, 1\}$ and a function $z : [0, T] \to \mathbb{R}^k$ of bounded variation, not both zero, such that for almost all $t \in [0, T]$, one has

(a) the Hamiltonian inclusion

$$(-\dot{z}(t), \dot{y}^*(t)) \in \bar{\partial} H(t, y^*(t), z(t)),$$

where $\bar{\partial}$ represents Clark's generalized gradient,

(b) the Euler-Lagrange inclusion

$$\dot{z}(t) \in \mathrm{co}\{z' : (z', z(t)) \in N_{\mathrm{gph}\, F(t,\cdot)}(y^*(t), \dot{y}^*(t))\},$$

where $\mathrm{gph}\, F(t, \cdot)$ denotes the graph of $F(t, \cdot)$, and N denotes a normal cone

(c) the Weierstrass-Pontryagin maximum condition

$$\langle z(t), \dot{y}^*(t)\rangle = \max\{\langle z(t), y'\rangle : y' \in F(t, y^*(t))\}.$$

The adjoint arc also satisfies

(d) the transversality inclusion

$$z(0) \in z^0 \partial\gamma(y^*(0)),$$
$$z(T) \in -N_S(y^*(T)).$$

Proof of Theorem 1

Suppose that $z^0 = 1$. We start with the transversality condition (d). Since the objective function $\gamma(\cdot)$ given by (3.3.4) is linear and its subdifferential set consists of a single element (its gradient), the first condition of (d) leads to the boundary conditions of the form (3.4.8)-(3.4.9).

To obtain the adjoint differential inclusion, we observe that the state inclusion (3.3.5) is obtained as an intersection of a set of continuously differentiable convex constraints, and it satisfies Slater's condition. Therefore, the convexification operation in (b) (the Euler-Lagrange inclusion) is redundant, and it takes a simplified form

$$(\dot{\mathbf{z}}_t, \mathbf{z}_t) \in N_{\mathrm{gph}\, F(t,\cdot)}(\mathbf{v}_t^*, \dot{\mathbf{v}}_t^*).$$

The normal cone to the graph of $F(t,\cdot)$ at $\mathbf{v}_t^*, \dot{\mathbf{v}}_t^*$ is characterized as a finitely generated cone spanned by the gradients of active inequalities at $\mathbf{v}_t^*, \dot{\mathbf{v}}_t^*$. We associate a nonnegative multiplier $\mu_{r,t}$ with each constraint. Multipliers of inactive constraints are equal to 0, which leads to condition (3.4.5). Other conditions (3.4.1)-(3.4.4) are obtained immediately by calculating the gradient of $f_{\mathbf{r}}(t, \mathbf{v}_t, \dot{\mathbf{v}}_t)$ with respect to $\mathbf{v}_t$ and $\dot{\mathbf{v}}_t$. Conditions (3.4.1), (3.4.2), (3.4.3) and (3.4.4) match the gradient components $\frac{\partial f_{\mathbf{r}}}{\partial v_{0,t}}$, $\frac{\partial f_{\mathbf{r}}}{\partial v_{i,t}}$, $\frac{\partial f_{\mathbf{r}}}{\partial \dot{v}_{0,t}}$ and $\frac{\partial f_{\mathbf{r}}}{\partial \dot{v}_{i,t}}$, respectively.

The Weierstrass-Pontryagin maximum condition (3.4.6)-(3.4.7) is a direct application of the general condition (c). General condition (a) is redundant.

To rule out the case of $z^0 = 0$, we observe that this implies $\mu_{\mathbf{r},t} = 0$ for all $\mathbf{r}, t$ (from (3.4.3)) and, therefore, $z_{i,t} = 0$ for all i, t. However, the scalar z^0 and the adjoint trajectory cannot be both zero – a contradiction.

The last observation in the theorem is a direct consequence of the fact that $\mu_{\mathbf{r},t} \geq 0$, for all $\mathbf{r} \in \mathcal{R}$ and conditions (3.4.3)-(3.4.5). In particular, nonnegativity of $\mu_{\mathbf{r},t}$'s and (3.4.3)-(3.4.4) means dual feasibility and (3.4.5) is a complimentary slackness condition.

Proof of Corollary 2

Optimality conditions (3.4.8), (3.4.1) and (3.4.3) immediately imply the first part of the statement. Moreover, it follows that at least one of $\mu_{\mathbf{r},t}$ is nonzero and, therefore, the corresponding constraint in (3.3.5) is active. From (3.4.4) it follows that if $z_{i,t} > 0$ then there is at least one $\mathbf{r}$ such that $r_i > 0$ for which $\mu_{\mathbf{r},t} > 0$. Also, it is elementary to check that

$$\frac{\partial M_j^\epsilon}{\partial v_{i,t}} = \begin{cases} 0, & i \notin A_j, \\ -\frac{1}{2}\left\{ \frac{p_j - \sum_{i' \in A_j} v_{i',t}}{\left(\left[p_j - \sum_{i' \in A_j} v_{i',t} \right]^2 + \epsilon^2 \right)^{\frac{1}{2}}} + 1 \right\}, & i \in A_j. \end{cases} \tag{B.1.5}$$

From (B.1.5), it is immediate that $-1 < \frac{\partial M_j^\epsilon}{\partial v_{i,t}} < 0$, $i \in A_j$. Thus, for all $\mathbf{r} \neq \mathbf{0}$, we have

$$\frac{\partial f_{\mathbf{r}}^0}{\partial v_{i,t}}(\mathbf{v}_t^*) = \sum_{j:A_j \leq \mathbf{r}, i \in A_j} \lambda_{j,t} \frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t^*) < 0.$$

Since (3.4.2) represents the weighted sum of $\frac{\partial f_{\mathbf{r}}^0}{\partial v_{i,t}}$'s with $\mu_{\mathbf{r},t}$'s a weights, it follows that $\dot{z}_{i,t} < 0$ unless $\mu_{\mathbf{r},t} = 0$ for all $\mathbf{r}$ such that $r_i > 0$. The latter means that $\dot{z}_{i,t} = 0$ and is only possible if $z_{i,t} = 0$.

Proof of Corollary 3

Consider any $i \in \mathcal{I}$. Since $z_{i,t}$ is strictly decreasing for almost all t except when $z_{i,t} = 0$, it is true that $z_{i,t} < c_i$ for any $t > 0$. Since $r_i \leq c_i$ for all $\mathbf{r} \in \mathcal{R}$, optimality condition (3.4.4) implies that for each t there exists $\mathbf{r}$ such that $r_i < c_i$ and $\mu_{\mathbf{r},t} > 0$. This can only occur if the constraint corresponding to $\mathbf{r}$ is active. Thus, we have

$$\dot{v}_{0,t}^* + r_i \dot{v}_{i,t}^* + \sum_{i' \in \mathcal{I} \setminus \{i\}} r_{i'} \dot{v}_{i',t}^* + f_{\mathbf{r}}^0(t, \mathbf{v}_t^*) = 0.$$

At the same time, $r_i + 1 \leq c_i$. Therefore, we have

$$\dot{v}_{0,t}^* + (r_i + 1)\dot{v}_{i,t}^* + \sum_{i' \in \mathcal{I} \setminus \{i\}} r_{i'} \dot{v}_{i',t}^* + f_{\mathbf{r}+\mathbf{e}_i}^0(t, \mathbf{v}_t^*) \leq 0,$$

where $\mathbf{e}_i$ is the ith unit vector. It follows that

$$\dot{v}_{i,t}^* \leq f_{\mathbf{r}}^0(t, \mathbf{v}_t^*) - f_{\mathbf{r}+\mathbf{e}_i}^0(t, \mathbf{v}_t^*)$$

Since $f^0_{\mathbf{r}+\mathbf{e}_i}(t, \mathbf{v}^*_t) \geq f^0_{\mathbf{r}}(t, \mathbf{v}^*_t)$ it follows that $\dot{v}^*_{i,t} \leq 0$. Nonnegativity follows from the boundary condition $v^*_{i,T} = 0$ in conjunction with monotonicity.

For $i = 0$, the statement is obtained immediately from (3.3.5) with $\mathbf{r} = \mathbf{0}$ and the boundary condition $v^*_{0,T} = 0$.

We also observe that the optimal solution to (3.4.6)-(3.4.7) cannot have $w_i < 0$ if $z_{i,t} > 1$. Indeed, if this were true, we could construct another feasible solution $\mathbf{w}'$ where $w'_0 = w_0 + w_i$ and $w'_i = 0$ with a strictly higher value (because $z_{0,t}(w_0 + w_i) > z_{0,t}w_0 + z_{i,t}w_i$). The new solution is indeed feasible because all constraints with $r_i > 1$ are less restrictive than the ones with $r_i = 1$, and the left-hand-sides of constraints with $r_i = 1$ have the same value at $\mathbf{w}$ and $\mathbf{w}'$. Moreover, since $\dot{\mathbf{v}}^*_t$ is an optimal solution to (3.4.6)-(3.4.7), it must have $\dot{v}^*_{i,t} = 0$. According to Corollary 2, adjoint variable $z_{i,t}$ is strictly decreasing when positive, and, from the transversality condition (3.4.8), we know that it starts at $c_i \geq 1$. Thus, $v^*_{i,t}$ remains constant on an interval $[0, t'_i]$ such that $z_{i,t'_i} = 1$.

Proof of Proposition 6

Let the optimal bid price trajectory of AOCP be $v^*_{i,t}$, $i \in \mathcal{I} \cup \{0\}$. Let the corresponding adjoint trajectory $z_{i,t}$ and multipliers $\mu_{\mathbf{r},t}$ be as described in Theorem 1. We construct a solution to the static LP so that

$$Y_j = \int_0^T \sum_{r:A_j \leq \mathbf{r}} \lambda_{j,t}\mu_{\mathbf{r},t}\frac{1}{2}\left\{\frac{p_j - \sum_{i'\in A_j} v_{i',t}}{\left(\left[p_j - \sum_{i'\in A_j} v_{i',t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + 1\right\} dt, \quad j \in \mathcal{J}. \qquad \text{(B.1.6)}$$

First, we show that (B.1.6) defines a feasible solution to the static LP. Recall that, according to (B.1.5), we have

$$\frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t^*) = -\frac{1}{2}\left\{ \frac{p_j - \sum_{i' \in A_j} v_{i',t}^*}{\left(\left[p_j - \sum_{i' \in A_j} v_{i',t}^*\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + 1 \right\},$$

for each for $i \in A_j$. Thus, Y_j's satisfy constraints (3.4.11) because

$$\begin{aligned}
\sum_{j:i \in A_j} Y_j &= -\int_0^T \sum_{\mathbf{r} \in \mathcal{R}} \mu_{\mathbf{r},t} \sum_{j:i \in A_j, A_j \leq \mathbf{r}} \lambda_{j,t} \frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t^*) dt \\
&= -\int_0^T \dot{z}_{i,t} dt, \quad \forall\, i \in A_j \\
&= z_{i,0} - z_{i,T} \leq c_i, \quad \forall\, i \in A_j.
\end{aligned}$$

The last inequality holds because $z_{i,0} = c_i$ by (3.4.9) and $z_{i,t} \geq 0$ for any t due to (3.4.4). Y_j's also satisfy constraints (3.4.12), because $-1 < \frac{\partial M_j^\epsilon}{\partial v_{i,t}}(\mathbf{v}_t^*) \leq 0$ and, therefore,

$$0 \leq Y_j \leq \int_0^T \sum_{r:A_j \leq \mathbf{r}} \mu_{\mathbf{r},t} \lambda_{j,t} dt \leq \int_0^T \lambda_{j,t} dt.$$

It remains to relate the value of the objective function (3.4.10) at Y_j's defined by (B.1.6) to the optimal value of AOCP. Multiply (3.3.5) by $\mu_{\mathbf{r},t}$, sum over all $\mathbf{r}$ and integrate over t to get

$$\int_0^T \sum_{\mathbf{r} \in \mathcal{R}} \mu_{\mathbf{r},t} \left(\dot{v}_{0,t}^* + \sum_{i \in \mathcal{I}} r_i \dot{v}_{i,t}^* + f_{\mathbf{r}}^0(t, \mathbf{v}_t^*) \right) dt = 0.$$

The equality holds because of (3.4.5). The first part of this integral can be expressed in terms of $z_{i,t}$'s using (3.4.3)-(3.4.4) and integrated by parts to obtain

$$\begin{aligned}\int_0^T \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t}\left(\dot{v}^*_{0,t} + \sum_{i\in\mathcal{I}} r_i \dot{v}^*_{i,t}\right) dt &= \int_0^T \sum_{i\in\mathcal{I}\cup\{0\}} z_{i,t}\dot{v}^*_{i,t} dt \\ &= \sum_{i\in\mathcal{I}\cup\{0\}} z_{i,t} v^*_{i,t}\Big|_0^T - \int_0^T \sum_{i\in\mathcal{I}\cup\{0\}} \dot{z}_{i,t} v^*_{i,t} dt \\ &= -\Big(v^*_{0,0} + \sum_{i\in\mathcal{I}} c_i v^*_{i,0}\Big) - \int_0^T \sum_{i\in\mathcal{I}} \dot{z}_{i,t} v^*_{i,t} dt,\end{aligned}$$

where the last equality holds because of (3.3.6), (3.4.1) and (3.4.8)-(3.4.9). It follows that we can express the AOCP optimal objective value as

$$\begin{aligned}v^*_{0,0} + \sum_{i\in\mathcal{I}} c_i v^*_{i,0} &= \int_0^T \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} f^0_{\mathbf{r}}(t, \mathbf{v}^*_t) dt - \int_0^T \sum_{i\in\mathcal{I}} \dot{z}_{i,t} v^*_{i,t} dt \\ &= \int_0^T \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} \sum_{j\in\mathcal{J}:A_j\le\mathbf{r}} \lambda_{j,t}\left(M^\epsilon_j(\mathbf{v}^*_t) - \sum_{i\in\mathcal{I}} v^*_{i,t}\frac{\partial M^\epsilon_j}{\partial v_{i,t}}(\mathbf{v}^*_t)\right) dt\end{aligned}$$

For each j we can write

$$\begin{aligned}M^\epsilon_j(\mathbf{v}^*_t) - \sum_{i\in\mathcal{I}} v^*_{i,t}\frac{\partial M^\epsilon_j}{\partial v_{i,t}}(\mathbf{v}^*_t) &= \frac{1}{2}\Bigg\{\left(\left[p_j - \textstyle\sum_{i\in A_j} v^*_{i,t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}} + p_j - \textstyle\sum_{i\in A_j} v^*_{i,t} \\ &\quad + \textstyle\sum_{i\in A_j} v^*_{i,t}\left(\dfrac{p_j - \sum_{i\in A_j} v^*_{i,t}}{\left(\left[p_j - \sum_{i\in A_j} v^*_{i,t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + 1\right)\Bigg\} \\ &= \frac{1}{2}\left\{\frac{p_j\left[p_j - \sum_{i\in A_j} v^*_{i,t}\right] + \epsilon^2}{\left(\left[p_j - \sum_{i\in A_j} v^*_{i,t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + p_j\right\} \\ &= \frac{p_j}{2}\left\{\frac{p_j - \sum_{i\in A_j} v^*_{i,t}}{\left(\left[p_j - \sum_{i\in A_j} v^*_{i,t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + 1\right\} \\ &\quad + \frac{1}{2}\frac{\epsilon^2}{\left(\left[p_j - \sum_{i\in A_j} v^*_{i,t}\right]^2 + \epsilon^2\right)^{\frac{1}{2}}}\end{aligned}$$

The last term in this expression is bounded by $\frac{\epsilon}{2}$. Therefore,

$$\begin{aligned} v_{0,0}^* + \sum_{i\in\mathcal{I}} c_i v_{i,0}^* &\leq \int_0^T \sum_{\mathbf{r}\in\mathcal{R}} \mu_{\mathbf{r},t} \sum_{j\in\mathcal{J}:A_j\leq\mathbf{r}} \lambda_{j,t} \left[\frac{p_j}{2} \left\{ \frac{p_j - \sum_{i\in A_j} v_{i,t}^*}{\left(\left[p_j - \sum_{i\in A_j} v_{i,t}^*\right]^2 + \epsilon^2\right)^{\frac{1}{2}}} + 1 \right\} + \frac{\epsilon}{2} \right] dt \\ &\leq \sum_{j\in\mathcal{J}} p_j Y_j + \frac{\epsilon}{2} \int_0^T \sum_{j\in\mathcal{J}} \lambda_{j,t} dt. \end{aligned}$$

The optimal value of the static LP is $V^{LP}(\mathbf{c}, 0) \geq \sum_{j\in\mathcal{J}} p_j Y_j$. The claim of the proposition follows.

Proof of Proposition 7

We use the results of [60] to demonstrate that the discretization error over each spline piece is bounded by $L_k d_k^2$. The first step of the proof is to demonstrate that problem (3.5.1)-(3.5.12) and (3.3.5) satisfies the assumptions from [60] outlined above. Fortunately, our SIP is a simplified case of the general SIP because our infinite index set is one dimensional over $[0, T]$. Assumption 1 implies that g is Lipschitz continuous near the solution $\bar{x}$ with respect to Y. Since (3.3.5) is continuously differentiable, Assumption 1 is satisfied. The inequalities in Assumption 2 for our problem can be expressed as $\zeta_1(t) = t - T \leq 0$ and $\zeta_2(t) = -t \leq 0$. The derivatives of the inequalities are $\frac{\partial \zeta_1}{\partial t} = 1$ and $\frac{\partial \zeta_2}{\partial t} = -1$, so they satisfy the MFCQ condition (Assumption 3) that $\frac{\partial \zeta_i(\bar{y})}{\partial y}\eta < 0,\ i \in I(\bar{y})$ with $\eta(\bar{y}|\zeta_1(\bar{y}) = 0) < 0$ and $\eta(\bar{y}|\zeta_2(\bar{y}) = 0) > 0$. The sets S_l represents the boundaries to the infinite set of constraints, since Y and $S_l \cap Y \neq \emptyset$. Since we include $t = 0$ and $t = T$ in our set $\mathcal{T}$, as well as the set of spline knot points t_k, $k \in \mathcal{K}$, our problem satisfies Assumption 4. This implies that all assumptions are satisfied.

Using the fact that all assumptions are satisfied, there exists $||\hat{t}_d - t|| \leq d$ and

$\frac{\partial^2 \eta_1}{\partial t^2} = \frac{\partial^2 \eta_2}{\partial t^2} = 0$, for some value L

$$f_{\mathbf{r}}(t, \mathbf{v}_t^M, \dot{\mathbf{v}}_t^M) \leq \frac{1}{2}\frac{d^2}{dt^2}\big(f_{\mathbf{r}}(t, \mathbf{v}_t^M, \dot{\mathbf{v}}_t^M)\big)d^2 \leq Ld^2,$$

where, for $\mathbf{r} \in \mathcal{R}$ and $t \in [0, T]$,

$$\frac{d^2}{dt^2}\big(f_{\mathbf{r}}(t, \mathbf{v}_t^M, \dot{\mathbf{v}}_t^M)\big) = \dddot{v}_{0,t}^M + \sum_{i \in \mathcal{I}} r_i \dddot{v}_{i,t}^M + \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \frac{\partial^2\big(\lambda_{j,t} M_j^\epsilon(\mathbf{v}_t^M)\big)}{\partial t^2}.$$

To bound $\frac{d^2}{dt^2}\big(f_{\mathbf{r}}(t, \mathbf{v}_t^M, \dot{\mathbf{v}}_t^M)\big)$ for any $\mathbf{r} \in \mathcal{R}$ by L, we first substitute the spline parameters $a_{i,l}^k$ of the discretized SIP into the first three derivatives of the bid prices. For $i \in \mathcal{I} \cup \{0\}$ and $t \in [t_k + t_{k+1}]$,

$$-\dot{v}_{i,t}^M = \sum_{l=0}^{3} a_{i,l}^k \left(\frac{t - t_k}{t_{k+1} - t_k}\right)^l, \quad -\ddot{v}_{i,t}^M = \sum_{l=1}^{3} a_{i,l}^k \frac{l(t-t_k)^{l-1}}{(t_{k+1} - t_k)^l}, \quad -\dddot{v}_{i,t}^M = \sum_{l=2}^{3} a_{i,l}^k \frac{l(l-1)(t-t_k)^{l-2}}{(t_{k+1} - t_k)^l}.$$

This implies that the derivatives of the bid prices are bounded as follows:

$$|\dot{v}_{i,t}^M| \leq \sum_{l=0}^{3} |a_{i,l}^k|, \qquad |\ddot{v}_{i,t}^M| \leq \sum_{l=1}^{3} |a_{i,l}^k| \frac{l}{t_{k+1} - t_k}, \qquad |\dddot{v}_{i,t}^M| \leq \sum_{l=2}^{3} |a_{i,l}^k| \frac{l(l-1)}{(t_{k+1} - t_k)^2},$$

which by definitions is equivalent to $\dot{v}_{i,k}^B$, $\ddot{v}_{i,k}^B$, and $\dddot{v}_{i,k}^B$. For some product $j \in \mathcal{J}$ such that $A_j \leq \mathbf{r}$,

$$\frac{\partial^2\big(\lambda_{j,t} M_j^\epsilon(\mathbf{v}_t^M)\big)}{\partial t^2} = \lambda_{t,j}\frac{d^2}{dt^2}\big(M_j^\epsilon(\mathbf{v}_t^M)\big) + 2\dot{\lambda}_{t,j}\frac{d}{dt}\big(M_j^\epsilon(\mathbf{v}_t^M)\big) + M_j^\epsilon(\mathbf{v}_t^M)\ddot{\lambda}_{t,j}.$$

We can now bound each of the M terms as follows:

$$
\begin{aligned}
M_j^\epsilon(\mathbf{v}_t^M) &\leq p_j \\
\frac{d}{dt}\big(M_j^\epsilon(\mathbf{v}_t^M)\big) &= \frac{-\sum_{i\in A_j}\dot{v}_{i,t}^M}{2}\left(\frac{p_j-\sum_{i\in A_j}v_{i,t}^M}{\sqrt{(p_j-\sum_{i\in A_j}v_{i,t}^M)^2+\epsilon^2}}+1\right) \quad\leq\quad -\sum_{i\in A_j}\dot{v}_{i,t}^M \\
\frac{d^2}{dt^2}\big(M_j^\epsilon(\mathbf{v}_t^M)\big) &= \frac{\sum_{i\in A_j}\big(\dot{v}_{i,t}^M\big)^2}{2}\frac{\Big((p_j-\sum_{i\in A_j}v_{i,t}^M)^2+\epsilon^2-\big(p_j-\sum_{i\in A_j}v_{i,t}^M\big)^2\Big)}{\Big(\big(p_j-\sum_{i\in A_j}v_{i,t}^M\big)^2+\epsilon^2\Big)^{3/2}} \\
&\quad -\frac{\sum_{i\in A_j}\ddot{v}_{i,t}^M}{2}\left(\frac{p_j-\sum_{i\in A_j}v_{i,t}^M}{\sqrt{(p_j-\sum_{i\in A_j}v_{i,t}^M)^2+\epsilon^2}}+1\right) \quad\leq\quad \sum_{i\in A_j}\frac{\big(\dot{v}_{i,t}^M\big)^2}{2\epsilon}-\sum_{i\in A_j}\ddot{v}_{i,t}^M.
\end{aligned}
$$

Since we have expressed the terms of $\frac{d^2}{dt^2}\big(f_\mathbf{r}(t,\mathbf{v}_t^M,\dot{\mathbf{v}}_t^M)\big)$ in terms of spline coefficients, the bound L and mesh distance d can be varied over each mesh interval $[t_k,t_{k+1}]$. Given that the absolute values of $\lambda_{t,j}$ and its first and second derivatives over each spline interval are bounded by $\lambda_{j,k}^B$, $\dot{\lambda}_{j,k}^B$, and $\ddot{\lambda}_{j,k}^B$, then we can bound $\frac{1}{2}\frac{d^2}{dt^2}\big(f_\mathbf{r}(t,\mathbf{v}_t^M,\dot{\mathbf{v}}_t^M)\big)$ over each interval $k\in\mathcal{K}$ by

$$
L_k=\frac{1}{2}\left[\sum_{i\in\mathcal{I}\cup\{0\}}\dddot{v}_{i,k}^B+\sum_{j\in\mathcal{J}}\left(\lambda_{j,k}^B\sum_{i\in A_j}\left(\frac{\big(\dot{v}_{i,k}^B\big)^2}{2\epsilon}-\ddot{v}_{i,k}^B\right)+2\dot{\lambda}_{j,k}^B\sum_{i\in A_j}\dot{v}_{i,k}^B+\ddot{\lambda}_{j,k}^B p_j\right)\right].
$$

Note that the value L_k corresponds to the resource vector where the airline has sufficient capacity available to fulfill all product requests, since this vector of $\mathbf{r}$ will lead to the largest possible absolute value of $\frac{d^2}{dt^2}\big(f_\mathbf{r}(t,\mathbf{v}_t^M,\dot{\mathbf{v}}_t^M)\big)$. Since $f_\mathbf{r}(t,\mathbf{v}_t^M,\dot{\mathbf{v}}_t^M)\leq \frac{1}{2}\frac{d^2}{dt^2}\big(f_\mathbf{r}(t,\mathbf{v}_t^M,\dot{\mathbf{v}}_t^M)\big)d^2\leq L_k d_k^2$, for all $t\in[t_k,t_{k+1}]$, it follows that

$$
V^M(\mathbf{c},0)\leq V^{A*}(\mathbf{c},0)+\sum_{k=1}^{K}L_k d_k^2,
$$

Proof of Proposition 9

Consider a Master bid-price trajectory $\mathbf{v}_t^R = \{\mathbf{v}_0^R, \mathbf{v}_1^R, \ldots, \mathbf{v}_m^R\}$ and the corresponding violations π_t^R. To bound the error over the entire booking horizon we approximate π_t^R with the following spline

$$\pi_t^A = \sum_{l=0}^{3} b_{i,l}^k \left(\frac{t - t_k}{t_{k+1} - t_k}\right)^l, \forall\ t \in [t_k, t_{k+1}],$$

where $t_k,\ k \in \mathcal{K} \equiv \{1, \ldots, K\}$ is the same collection of knot points as used in the ASOCP. The continuity of this approximation and its derivatives is ensured by the constraints:

$$\begin{aligned}
b_{i,0}^{k+1} &= \sum_{l=0}^{3} b_{i,l}^k, \\
\frac{1}{t_{k+2} - t_{k+1}} b_{i,1}^{k+1} &= \sum_{l=1}^{3} \frac{l}{t_{k+1} - t_k} b_{i,l}^k, \\
\frac{2}{(t_{k+2} - t_{k+1})^2} b_{i,2}^{k+1} &= \sum_{l=2}^{3} \frac{l(l-1)}{(t_{k+1} - t_k)^2} b_{i,l}^k,
\end{aligned}$$

which apply to every $k = 1, \ldots, K-1$ and $i \in \mathcal{I}$. We ensure that this approximation is nonnegative upper bound on π_t^R we also include the constraints

$$\begin{aligned}
\pi_t^A &\geq \max(\pi_t^*, 0) \\
b_{i,0}^k &= y_{i,0}^k, \\
b_{i,1}^k &= 2y_{i,1}^k + x_{i,0}^k - y_{i,0}^k, \\
b_{i,2}^k &= y_{i,2}^k + 2x_{i,1}^k - 2y_{i,1}^k, \\
b_{i,3}^k &= x_{i,2}^k - y_{i,2}^k, \\
\left(\frac{x_{i,0} + x_{i,2}}{2}\right)^2 &\geq \left(\frac{x_{i,0} - x_{i,2}}{2}\right)^2 + x_{i,1}^2, \\
\left(\frac{y_{i,0} + y_{i,2}}{2}\right)^2 &\geq \left(\frac{y_{i,0} - y_{i,2}}{2}\right)^2 + y_{i,1}^2.
\end{aligned}$$

The total error from time t to the end of the booking horizon, E_t, is found by integrating π_t^A over the interval $[t, T]$. Thus, the objective is to find a minimum upper bound on the approximation of the violation of (3.5.15)

$$E_0 = \min_{\pi^A} \int_{t=0}^{T} \pi_t^A dt.$$

If π_t^A is a solution to the error spline approximation problem, then for all $\mathbf{r} \in \mathcal{R}_t$ and $t \in \mathcal{T}$ the following constraint holds:

$$\dot{v}_{0,t}^R - \pi_t^A + \sum_{i \in \mathcal{I}} r_i \dot{v}_{i,t}^R + \sum_{j \in \mathcal{J}: A_j \leq \mathbf{r}} \lambda_{j,t} \frac{1}{2}(h_{j,t} + g_{j,t}) \leq 0.$$

Let $\dot{v}_{0,t}^M = \dot{v}_{0,t}^R - \pi_t^A$ for all $t \in \mathcal{T}$ and $\dot{v}_{i,t}^M = \dot{v}_{i,t}^R$ for all $i \in \mathcal{I}$ and $t \in \mathcal{T}$. The feasible values of $v_{i,t}^M$ can be found by integrating over $[t, T]$. At time t, $v_{i,t}^M$ is trivially equal to $v^R_{i,t}$ for all $i \in \mathcal{I}$ and

$$\begin{aligned} v^M_{0,t} &= -\int_t^T (\dot{v}^R_{0,t} - \pi^A_t)dt \\ &= v^R_{0,t} + E_t. \end{aligned}$$

The bid-price trajectory $\mathbf{v}^M_t$ is feasible for ASOCP and provides an upper bound on its optimal value

$$V^M = v^R_{0,0} + E_0 + \sum_{i \in \mathcal{I}} v^R_{i,0} c_i = V^R(\mathbf{c}, 0) + E_0,$$

This bound ensures that the current solution based on $\mathcal{R}_t$, $t \in \mathcal{T}$ is within fraction Ω of the optimal solution if $\frac{E_0}{V^R(\mathbf{c},0)} \leq \Omega$.

B.2. Numerical Experiments

The standard errors for comparing the fixed bid price controls of the LP, DD, and ASOCP for the fixed bid price experiments are Table B.1. The networks are indexed by the scaling parameter κ and by the variability in the initial capacity vector $\mathbf{u}$. Tables B.2 reports the run-time for the master and row problem for the ASOCP in the fixed bid price experiments. The standard errors and ASOCP run-times for the multi-hub experiments with updating are presented in Table B.3. The networks are identified by lowest to highest load factor at each number of hubs. Table B.4 shows the standard errors and total run-time, rather than the average run-time per optimization, for the master and row problems across the porter network.

Table B.1: LP and ASOCP standard errors for fixed bid price experiments.

		$\nu=0$		$\nu=1$		$\nu=2$		$\nu=3$		$\nu=4$	
κ	LP	DD	ASOCP	DD	ASOCP	DD	ASOCP	DD	ASOCP	DD	ASOCP
1	17	23	20	16	14	16	14	15	14	15	14
2	28	43	30	28	20	28	20	27	20	28	21
5	56	80	56	55	40	56	41	55	42	55	42
10	102	138	97	97	70	98	71	100	75	104	79
20	193	245	179	171	126	168	128	176	133	191	143

Table B.2: ASOCP run-time breakdown for fixed bid price experiments.

	$\nu=0$		$\nu=1$		$\nu=2$		$\nu=3$		$\nu=4$	
κ	Master	Row	Master	Row	Master	Row	Master	Row	Master	Row
1	0.2849	0.1296	0.3329	0.1657	0.3455	0.1568	0.4129	0.1856	0.6158	0.2729
2	0.3128	0.1292	0.3702	0.1651	0.3631	0.1690	0.3779	0.1713	0.4134	0.1849
5	0.3705	0.1287	0.4316	0.1693	0.4253	0.1688	0.4388	0.1658	0.4355	0.1676
10	0.4025	0.1285	0.5091	0.1702	0.4749	0.1662	0.4866	0.1677	0.5071	0.1743
20	0.4400	0.1286	0.5110	0.1431	0.5448	0.1720	0.5333	0.1662	0.5162	0.1535

Table B.3: Standard errors and ASOCP run-time breakdown for multi-hub networks.

	1 Hub				2 Hubs				3 Hubs			
Load	Std. Errors		Run Times		Std. Errors		Run Times		Std. Errors		Run Times	
Factor	LP	ASOCP	Master	Row	LP	ASOCP	Master	Row	LP	ASOCP	Master	Row
Low	691	710	1.441	0.703	402	423	4.917	1.282	630	642	53.582	0.812
Mid	468	480	1.368	0.620	591	602	5.361	1.164	504	530	54.165	0.811
High	541	551	1.447	0.633	605	630	5.654	1.237	556	578	51.880	0.808

Table B.4: Standard errors and ASOCP run-time breakdown for Porter Network.

	Std. Errors				Total Run Time			
ϕ	LP - 1F	ASOCP - 1F	LP - 2F	ASOCP - 2F	Master - 1F	Row - 1F	Master - 2F	Row - 2F
5	938	811	4308	4015	56.745	2.518	64.969	6.810
10	903	813	4214	3976	181.309	5.071	198.504	5.987
20	859	792	4143	4201	316.765	7.548	358.596	25.908
35	880	810	4130	4158	609.511	14.240	761.892	56.496