Summer 9-1-2014

A Study of Upgrading in Capacity Management

Yueshan Yu
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:

Fuqiang Zhang, Chair

Lingxiu Dong

Amr Farahat

John Nachbar

Nan Yang

A Study of Upgrading in Capacity Management

by

Yueshan Yu

A dissertation presented to the Graduate School of Arts and Sciences

of Washington University in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

August 2014

St. Louis, Missouri

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Fuqiang Zhang, for his continuous support of my doctoral studies, and his encouragement, motivation, and immense knowledge. His guidance helped me through the process of researching and writing this dissertation. I am extremely grateful for his advice, which has been invaluable both professionally and personally. I cannot imagine having a better advisor and mentor.

I would also like to thank Professor Amr Farahat and Professor Nan Yang for their constant support and valuable comments. I am also deeply grateful to Professor Lingxiu Dong and Professor John Nachbar for serving as my committee members and for their valuable comments, which have greatly improved this dissertation. I would like to extend my thanks to other members of the Operations & Manufacturing Management faculty from whom I have learned or with whom I have worked over the last six years for showing me what it means to be a good researcher, each in their own unique way: Professor Panagiotis (Panos) Kouvelis, Professor Yossi Aviv, Professor Sergio Chayet, and Professor Danko Turcic.

Chapter 2 of this dissertation has benefited from the insightful advice and comments from Professor Xin Chen at the University of Illinois at Urbana-Champaign. I appreciate his time and help with my dissertation.

I have greatly benefited from inspirational discussions with Yixuan Xiao, Guang Xiao, and Dashan Huang. And I would like to thank Xiaole Wu, Li Wang, Fan Zhang, Wuchao Sun, and Xiaofei Huang for their generous help and all the joyful moments. My thanks also go to other peer students whose names are not listed.

This journey would not have been possible without the support of my family. I am immensely indebted to my parents for the unequivocal support that encouraged me to strive towards my goal. Finally, I would like express appreciation to my beloved wife Yunwen to whom I own the greatest debt for her unconditional love, faithful support, and endless patience. Thank you!

Chapter 1

Introduction

A central theme in operations management is matching supply with demand.

On the supply side, in order to achieve better market coverage and higher profit, many firms have expanded their product lines and started to offer differentiated products to various demand segments. However, the complexity of the supply-demand matching problem increases dramatically as more classes of products are introduced. To mitigate the risk of the mismatches between supply and demand, many firms use capacity or inventory flexibility to satisfy uncertain demand from multiple classes of consumers. In particular, consumers whose first-choice product is no longer available might be upgraded by the firm using a superior product. Such a practice takes advantage of risk pooling and results in several immediate benefits: First, it generates additional revenue by serving more demand; second, it increases service level by reducing lost sales; third, it may lead to a lower inventory investment by hedging against demand uncertainty. However, for many applications, the capacity/product allocation has to overcome two obstacles: First, firms usually assign capacity before demand is fully known as demand arrives over time; second, firms may not be able to replenish capacity before the end of the selling season due to a long lead time.

On the demand side, consumers are often heterogeneous in terms of product preference, taste of quality, and sophistication in decision-making. When firms offer multiple differentiated products to market, consumers will make their product selection decisions based on their individual preferences as well as the product prices. Moreover, when firms use product upgrade as an operational strategy, different consumers may have different perceptions of such a practice. For instance, a naive consumer may ignore the potential upgrading opportunity, while an opportunistic consumer may intentionally choose the product that is more likely to be sold out, hoping to receive a free upgrade. Such a heterogeneity clearly has an impact on firms' optimal selling strategies, including product differentiation, pricing, and upgrading policy.

This dissertation is devoted to studying the upgrading practice in capacity management, while taking the aforementioned factors into account. There are three parts in this dissertation: Chapter 2 studies firms' capacity management with upgrading. In this chapter, we focus on the firms' capacity investment and allocation decisions in a dynamic setting. The demand is assumed to be exogenously given. Chapter 3 examines firms' product differentiation and pricing strategies when consumer heterogeneity is taken into account. In this chapter, consumer decisions are endogenized. Chapter 4 extends the model in Chapter 3 by considering an upgrading problem where the firm controls the product differentiation by varying the upgrading probability.

The dynamic capacity management problem in Chapter 2 is modelled as follows. Consider a firm selling N products with differentiated quality in a fixed horizon consisting of T periods. Consumers, who randomly arrive in each period, are divided into N classes based on their first-choice products. In case of stock-out, the consumer could be upgraded to a higher quality product at no extra charge. Unsatisfied demand is backlogged and the firm incurs a backlog penalty cost. The firm makes two decisions: capacity investment at the beginning of the horizon and capacity allocation in each period throughout the horizon. We characterize the structure of the optimal allocation policy, which represents a Parallel and Sequential Rationing (PSR) policy. Then we propose a heuristic that adapts certainty equivalence control (CEC) to exploit the PSR properties and overcomes the curse of dimensionality in this dynamic programming model. Numerical studies demonstrate that the heuristic is efficient and yields a close-to-optimum performance. With the help of the heuristic, we derive several insights into the dynamic capacity management problem using extensive numerical experiments.

Chapter 3 and Chapter 4 complement Chapter 2 by shifting the focus to firms' product differentiation and pricing decisions. Specifically, these two chapters study a firm's optimal strategies when selling two differentiated products, the regular product and the premium product, to heterogeneous consumers. A single-period model is proposed where the firm is a monopolist but faces certain capacity constraints. Consumers arrive to the market and choose which product to purchase (or not to purchase anything). The firm may adjust product differentiation by varying the add-on services attached to the regular product. There are two types of consumers, naive and opportunistic, who may place different valuations on the add-on services. For example, when the firm offers the free product upgrade as an add-on service, a naive consumer may be unaware of or unable to evaluate such a potential benefit; in contrast, an opportunistic consumer may change her product selection decision after incorporating the potential upgrading opportunity. In Chapter 3, we examine how product differentiation caused by add-on services affects firms' pricing decisions and profits. We find that depending on problem situations, the firm's profit can be improved by either reducing or increasing product differentiation. Similarly, altering the consumer mix (the fraction of naive consumers) may influence the firm's profit in both directions. In addition, the firm's optimal product differentiation and consumer mix decisions hinge upon its capacity limits. In fact, the capacity constraint on the premium product can serve as an effective device in segmenting the two types of consumers.

Chapter 4 takes a step forward by developing a more sophisticated model to endogenize the upgrading probability. In reality, an opportunistic consumer makes her purchase decision in anticipation of the potential product upgrade, whose probability depends on various factors including firms' available capacities and other consumers' decisions. In Chapter 4, we consider a random market size; in addition, we follow the rational expectations paradigm by assuming the opportunistic consumers can rationally predict the upgrading probability. It has been found that the opportunistic behavior can be either beneficial or detrimental to the firm depending on model parameters. For a given fraction of the opportunistic consumers, we also investigate the firm's optimal percentage of the leftover premium capacity to be used for upgrading. Numerical analysis shows that the optimal percentage may vary from 0 to 100% depending on the problem situations. This indicates that the firm needs to be careful when determining the upgrading frequency.

Chapter 2

Dynamic Capacity Management with Upgrading

2.1 Introduction

Driven by intensified market competition and rapidly-changing consumer trends, many firms have expanded their product lines to cater to different customer segments. On the one hand, by offering products with a wide range of quality, design and characteristics, firms can reach more consumers, generate additional sales, and extract higher profit margins. On the other hand, it has caused significant difficulties in matching supply with demand because the demand is less predictable at the individual segment level than at the aggregate level. Various operational strategies (e.g., postponement, component commonality, modular design) have been proposed for firms to enjoy the benefit of product differentiation while mitigating the risk of mismatches between supply and demand. This chapter studies the influential practice of upgrading, where products with higher ranks can be used to satisfy demand for a lower product that is sold out. Such practice takes the advantage of risk pooling (product substitution essentially allows product/demand pooling), which results in several immediate benefits: first, it increases revenue by serving more demand; second, it enhances customer service by reducing lost sales; third, it may lead to lower inventory investment by hedging against demand uncertainty.

The practice of upgrading or substitution has been widely adopted in the business world. In the automobile industry, firms may shift demand for a dedicated capacity to a flexible capacity when the dedicated capacity is constrained (Wall 2003).

In the semiconductor industry, faster memory chips can substitute for slower chips when the latter are no longer available (Leachman 1987). More examples in production/inventory control settings can be found in Bassok et al. (1999) and Shumsky and Zhang (2009). Similar practice is ubiquitous in the service industries as well. For instance, airlines may assign business-class seats to economy-class passengers, car rental companies may upgrade customers to more expensive cars, and hotels may use luxury rooms to satisfy demand for standard rooms.

Both practitioners and academics surely understand the importance of the upgrading practice. Substantial research has been conducted on how to manage upgrading in a variety of problem settings. Here we contribute to this large body of literature by studying a dynamic capacity management problem under general upgrading structure. For convenience, we use the terms "product" and "capacity" exchangeably throughout the chapter, and similarly for "upgrading" and "substitution" (strictly speaking, upgrading is one-way substitution). A brief description of our problem is as follows. Consider a firm selling N products with differentiated quality in a fixed horizon consisting of T periods. There are N classes of customers who arrive randomly in each period. Each customer requests one unit of the product; in the case of stock-out, the customer can be satisfied with a higher quality product at no extra charge. Unsatisfied demand is backlogged and the firm incurs a backlog penalty cost. The firm needs to first determine the procurement quantity of each product at the beginning of the horizon, and then decide how to distribute the products among incoming customers. Due to long ordering lead time, the firm cannot replenish inventory before the end of the horizon; as a result, the firm must dynamically allocate the products over time, before observing future demand.

This chapter represents an extension of the recent work by Shumsky and Zhang (2009, referred to as SZ hereafter). As one of the first studies that incorporate dynamic allocation into substitution models, SZ make a simplifying assumption to maintain tractability. Specifically, they consider single-step upgrading, i.e., a demand can only be upgraded by the adjacent product. Clearly, this is a restrictive assumption because in many practical situations firms may have incentives to use multi-step upgrading to satisfy demand. Thus there is a need for a theoretical model that captures the realistic upgrading structure. The purpose of this chapter is to fill such a gap in the literature. While relaxing the single-step upgrading assumption, we attempt to address the following questions as in SZ: What is the optimal initial capacity? How should the products be allocated among customers over time? Are there any effective and efficient heuristics for solving the capacity management problem? The main findings from this chapter are summarized as follows.

We start with the dynamic capacity allocation problem. In each period, the firm needs to use the available products to satisfy the realized demand. When a product is depleted while there is still demand for that product, the firm may use upgrading to satisfy the customers. How to make such upgrading decisions is a key in substitution models. With the general upgrading structure, the optimal allocation policy is complicated by the fact that the upgrading decisions within a period are interdependent. Under the backlog assumption, we show that a Parallel and Sequential Rationing (PSR) policy is optimal among all possible policies. The PSR policy consists of two stages: In stage 1, the firm uses parallel allocation (i.e., demand is satisfied by the same-class capacity) to satisfy demand as much as possible. Then in stage 2, the firm sequentially upgrades leftover demand, starting from the highest demand class; when upgrading a given demand class, the firm starts with the lowest capacity class. The optimality of such a sequential rationing scheme depends on an important property. That is, when using a particular class of capacity to upgrade demand, the upgrading decision does not depend on the status of the portion of the system below that class. The PSR can greatly reduce the computational complexity because the upgrading decisions do not have to be solved simultaneously. As an extension, we also consider the multi-horizon model with capacity replenishment and show that the PSR policy remains optimal. Our theoretical results, though intuitive, turn out to be very challenging to prove. Indeed, our proofs rely on intricate arguments and fully exploit the special structure of the upgrading problem.

Despite the simplified solution procedure given by the PSR, solving the problem is challenging due to the curse of dimensionality. We search for fast heuristics that perform well for the firm. We present a heuristic that adapts certainty equivalence control (CEC) to exploit the PSR properties. Such a heuristic is more appealing than the commonly used CEC heuristic, and we call it refined certainty equivalence control (RCEC) heuristic. Through extensive numerical experiments, we find that the RCEC heuristic delivers nearly optimal profit for the firm: the average profit gap is less than 0.8% among all the experiments and the number is 2.76% at the 90^{th} percentile.

The RCEC heuristic enables us to solve large problems effectively. Thus we can use numerical studies to derive several insights into the dynamic capacity management problem. First, compared to single-step upgrading, general upgrading (multi-step upgrading) can be highly valuable, especially when the initial capacities are severely imbalanced. Second, given that the optimal upgrading policy is used, the firm's profit is not sensitive in the initial capacity. For instance, either the newsvendor capacities (calculated assuming no upgrading) or the static capacities (calculated assuming complete demand information) provide nearly optimal profit for the firm. However, the negative impact of using suboptimal allocation policies could be quite significant. These findings suggest that from the practical perspective, deriving the optimal allocation policy should receive a higher priority than calculating the optimal initial capacity.

The remainder of the chapter is organized as follows. Section 2.2 reviews the related literature. Section 2.3 describes the model setting. The optimal allocation policy is characterized by Sections 2.4 and 2.5. Section 2.6 extends the base model to multiple horizons with capacity replenishment. Section 2.7 proposes the RCEC heuristic and Section 2.8 presents the findings from numerical studies. The chapter concludes with Section 2.9. All proofs are given in the Appendices.

2.2 Literature Review

This chapter falls in the vast literature on how to match supply with demand when there are multiple classes of uncertain demand. To facilitate the review, we may divide this literature into four major categories using the following criteria: (1) whether there are multiple capacity types or a single capacity type; and (2) whether the nature of capacity allocation is static or dynamic. A problem is called static if capacity allocation can be made after observing full demand information. The category that involves the single capacity and static allocation essentially reduces to the newsvendor model that is less relevant. Thus, below our review focuses on the representative studies from the other three categories.

The first category of studies involves multiple capacity types and static capacity allocation. In these studies, firms invest in capacities before demand is realized and then allocate capacities to customers after observing all demand. Due to the existence of multiple capacity types, the issue of substitution naturally arises. Van Mieghem (2003) and Yao and Zheng (2003) provide comprehensive surveys of this category of studies, which can be further divided into two groups. One group of papers studies the optimal capacity investment and/or allocation decisions under substitution. Parlar and Goyal (1984) and Pasternack and Drezner (1991) are among the first to consider the simplest substitution structure with two products. Bassok et al. (1999) extend the problem to the general multi-product case. Hsu and Bassok (1999) introduce random yield into the substitution problem. By assuming single-level substitution, Netessine et al. (2002) study the impact of demand correlation on the optimal capacity levels. Van Mieghem and Rudi (2002) propose the notion of newsvendor networks that consist of multiple newsvendors and multiple periods of demand. Similar settings can be found in the studies on multi-period inventory models with transshipment, including Robinson (1990), Archibald et al. (1997), and Axsäter (2003). Although these studies involve multiple periods, replenishment is allowed and capacity allocation in each period is made with full demand information. The other group of studies focuses on the value of capacity flexibility. Fine and Freund (1990) and Van Mieghem (1998) consider two types of capacities (dedicated and flexible) and study the optimal investment in flexibility. Bish and Wang (2004) and Chod and Rudi (2005) incorporate pricing decisions when studying the value of resource flexibility. Jordan and Graves (1995) investigate a manufacturing flexibility design problem and discover the well-known chaining rule: Limited capacity flexibility, configured in a chaining structure, almost delivers the benefit of full flexibility. Their classic work on the design of flexibility has inspired numerous follow-up studies. For example, recently, Chou et al. (2010, 2011) have provided analytical evaluations of the chaining structure for both symmetric and asymmetric problem settings with large scales.

The second category of related literature studies the allocation of a single capacity to multi-class demand in a dynamic environment. This category dates back to the early work by Topkis (1968), who characterizes the optimal rationing policy that assigns capacity to different customer classes over time. Since then similar rationing policies have been applied to various industry settings. For instance, many revenue management studies focus on how to maximize firms' revenue through capacity rationing when there are multiple fare classes for a single seat type; see Talluri and van Ryzin (2004b) for a review of this literature. A stream of studies on production and inventory control has also derived threshold policies when serving multiple customer classes; see Ha (1997, 2000), de Véricourt et al. (2001, 2002), Deshpande et al. (2003), Savin et al. (2005), Ding et al. (2006) and the references therein.

The third category of studies involves multiple capacity types and dynamic capacity allocation. It differs from the first category mainly in that firms need to allocate capacities to customers without full demand information. There are relatively few papers in this category. Shumsky and Zhang (2009) consider a dynamic capacity management problem with single-step upgrading. They characterize the optimal upgrading policy and provide easy-to-compute bounds for the optimal protection limits that can help solve large problems. Xu et al. (2011) consider a two-product dynamic substitution problem where customers may or may not accept the substitution choice offered by seller. This chapter extends Shumsky and Zhang (2009) to allow general upgrading. We show that a sequential upgrading policy is optimal for such a problem and provide fast heuristics that can effectively solve the optimal capacity investment and allocation decisions. Our problem can be framed as a network revenue management model with full upgrading, where the fares are fixed and the incidence matrix is the identity matrix (see Gallego and van Ryzin 1997). Gallego and Stefanescu (2009) introduce two continuous optimal control formulations for capacity allocation but concentrate on the analysis of deterministic cases. Steinhardt and Gönsch (2012) study a similar network revenue management problem but allow at most one buying request in each period. In contrast, this chapter considers stochastic and batch demand arrivals in each period. Our work is also related to the studies on airline revenue management that involve multiple fare products. Talluri and van Ryzin (2004a) study revenue management under a general customer choice model. Zhang and Cooper (2005) consider the selling of parallel flights with dynamic customer choice among the flights. More recent developments include Liu and van Ryzin (2008a) and Zhang (2011). In these studies, firms need to decide the subset of products from which a customer can choose from; while in this chapter, firms decide how to allocate capacities to realized demand. Therefore, both the model settings and results are quite different between these studies and this chapter.

2.3 Model

Consider a firm managing N types of products to satisfy customer demand. The products are indexed in decreasing quality so that product 1 has the highest quality while product N has the lowest quality. There are N corresponding classes of customer demand, i.e., a customer is called class j if she requests product j ($1 \leq j \leq N$). The sales horizon consists of T discrete periods. The initial capacities of the products must be determined prior to the first period and no capacity replenishment is allowed during the sales horizon. (In Section 2.6, we extend the model to consider multiple horizons and allow for replenishment.) Customers arrive over time and the demand in each period is random. Let $\mathbf{D}^t = (d_1^t, d_2^t, \cdots, d_N^t)^\intercal \in \Re_+^N$ denote the demand vector for period t ($1 \leq t \leq T$), where superscript $\intercal$ stands for the transpose operation. Throughout the chapter we use bold letters for vectors and matrices, and use $(\mathbf{Z})_i$ for the i-th component of vector $\mathbf{Z}$ (or $(\mathbf{Z})_{ij}$ for the corresponding element in matrix $\mathbf{Z}$). For instance, $(\mathbf{D}^t)_i = d_i^t$ is the demand for product i in period t. We assume demand is independent across periods; however, demands for different products within a period can be correlated.

Let r_j be the revenue the firm collects from satisfying a class j customer. If product j is out of stock, then a class j customer could be upgraded at no extra charge by any product i as long as $i < j$. If a class j demand cannot be satisfied in period t, then it will be backlogged to the next period and the firm has to incur a goodwill cost g_j[1]. Define $\mathbf{G} = (g_1, \cdots, g_N) \in \Re_+^N$. To incorporate service settings like the car rental industry, we include a usage cost denoted by u_i for product i. We make the following assumptions:

Assumption 2.3.1 (A1) $r_1 > r_2 > \cdots > r_N$.

Assumption 2.3.2 (A2) $g_1 > g_2 > \cdots > g_N$.

Assumption 2.3.3 (A3) $u_1 > u_2 > \cdots > u_N$.

We may define $\alpha_{ij} = r_j + g_j - u_i$ ($i \leq j$) as the profit margin for satisfying a class j customer using product i. Based on the above assumptions, we know $\alpha_{ij} > \alpha_{ik}$ and $\alpha_{jk} > \alpha_{ik}$ ($i < j < k$). In other words, for a given capacity, it is more profitable to satisfy a higher class of demand; for a given demand, it is more profitable to use a lower class of capacity. These assumptions are similar to but more general than those made in SZ: we have relaxed the single step upgrading assumption in SZ ($\alpha_{ij} > 0$ only if $j = i + 1$) and added Assumption (A2) about the backorder costs.

[1] The backorder assumption is used mainly for tractability. Notice that an unmet demand could be upgraded in any subsequent periods, so it is reasonable to assume that the customers are willing to wait for potential upgrades, i.e., unsatisfied demands can be backlogged.

Note that the above assumptions do not require all α_{ij} to be positive. Specifically, if $\alpha_{ij} < 0$ for some i and j, then the assumptions imply that $\alpha_{1j} < \cdots < \alpha_{ij} < 0$ and $\alpha_{iN} < \cdots < \alpha_{ij} < 0$, which are reasonable in practice.

The firm's objective is to maximize the expected profit over the sales horizon. There are two major decisions for the firm. First, the firm needs to determine the initial capacity before the start of the selling season; second, the firm needs to allocate the available capacities to satisfy demands in each period. Let $\mathbf{C} = (c_1, \cdots, c_N) \in \Re^N_+$ denote the capacity cost vector, $\mathbf{X}^t = (x_1^t, x_2^t, \cdots, x_N^t)^\intercal \in \Re^N_+$ the starting capacities in period t, and $\tilde{\mathbf{D}}^t = (\tilde{d}_1^t, \tilde{d}_2^t, \cdots, \tilde{d}_N^t)^\intercal \in \Re^N_+$ the backordered demand at the beginning of period t. We use $\mathbf{Y}^t$ for the capacity allocation matrix in period t, i.e., $(\mathbf{Y}^t)_{ij} = y_{ij}^t$ is the amount of product i offered to satisfy class j demand ($y_{ij} = 0$ if $i > j$). Define $\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ as the optimal revenue-to-go function in period t given the state variable $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$. Then the buyer's problem can be formulated as follows:

$$\max_{\mathbf{X}^1 \in \Re^N_+} \Pi(\mathbf{X}^1) = \max_{\mathbf{X}^1 \in \Re^N_+} \left\{ \Theta^1(\mathbf{X}^1, \mathbf{0}) - \mathbf{C}\mathbf{X}^1 \right\}, \tag{2.1}$$

and for each period t $(1 \le t \le T)$:

$$\begin{aligned} \Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t) &= \mathop{\mathbb{E}}_{\mathbf{D}^t} \left\{ \Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t | \mathbf{D}^t) \right\} \\ &= \mathop{\mathbb{E}}_{\mathbf{D}^t} \left\{ \max_{\mathbf{Y}^t} \left[H(\mathbf{Y}^t | \tilde{\mathbf{D}}^t; \mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1}) \right] \right\}, \end{aligned} \tag{2.2}$$

where

$$H(\mathbf{Y}^t | \tilde{\mathbf{D}}^t; \mathbf{D}^t) = \sum_{1 \le i \le j \le N} \alpha_{ij} y_{ij}^t - \mathbf{G}(\tilde{\mathbf{D}}^t + \mathbf{D}^t), \tag{2.3}$$

$$\mathbf{X}^{t+1} = \mathbf{X}^t - \mathbf{Y}^t \mathbf{1} \ge \mathbf{0}, \tag{2.4}$$

$$\tilde{\mathbf{D}}^{t+1} = \tilde{\mathbf{D}}^t + \mathbf{D}^t - (\mathbf{Y}^t)^\intercal \mathbf{1} \ge \mathbf{0}, \tag{2.5}$$

$$\mathbf{Y}^t \ge 0, \qquad \mathbf{1} = (1, 1, \cdots, 1)^\intercal.$$

We assume the leftover products have zero value at the end of the selling season, so $\Theta^{T+1} \equiv 0$. Note that the optimal revenue-to-go function $\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ is recursively defined in (2.2). Given the allocation decision $\mathbf{Y}^t$, $H(\mathbf{Y}^t | \tilde{\mathbf{D}}^t; \mathbf{D}^t)$ in (2.3) denotes the single period revenue, which is the difference between the upgrading revenue and the goodwill cost. The state transition between two consecutive periods is governed by (2.4) and (2.5), which represent two constraints for the allocation decision $\mathbf{Y}^t$ in period t.

2.4 Parallel and Sequential Rationing (PSR)

This section starts analyzing the upgrading problem given in (2.1). First we introduce several useful definitions and qualitatively characterize the optimal allocation policy. The formal optimality proof will be presented in the next section. As the first step, since

$$\begin{aligned} \Pi(\mathbf{0}) &= -\mathbf{G}\sum_{t=1}^{T}(T+1-t)\,\mathbb{E}[\mathbf{D}^t] > -\infty, \\ \lim_{\mathbf{X}^1\to\infty}\Pi(\mathbf{X}^1) &= \sum_{t=1}^{T}\sum_{i=1}^{N}(r_i-u_i)\,\mathbb{E}[d_i^t] - \lim_{\mathbf{X}^1\to\infty}\mathbf{C}\mathbf{X}^1 = -\infty, \end{aligned} \tag{2.6}$$

and the fact that $\Pi(\mathbf{X}^1)$ is continuous in $\mathbf{X}^1 \in \Re_+^N$, we know there exists a finite $\mathbf{X}^* \in \Re_+^N$ that solves the optimization problem in (2.1).

From Murty (1983) and Rockafellar (1996), for any demand realization $\mathbf{D}^T$ in period T, it is straightforward to see $\Theta^T(\mathbf{X}^T, \tilde{\mathbf{D}}^T|\mathbf{D}^T)$ is concave in the state variable $(\mathbf{X}^T, \tilde{\mathbf{D}}^T)$, which are the right-hand side variables in the linear program defined by (2.2). Since concavity is preserved under the expectation operation on $\mathbf{D}^t$ $(1 \le t \le T)$ and the maximization operation with respect to the allocation decision $\mathbf{Y}^t$ (see, for example, Simchi-Levi et al. 2014, Proposition 2.1.3 and 2.1.15), Θ^t is again concave in $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ in each period t. Clearly, the function

$$\hat{\Theta}^t(\mathbf{Y}^t|\mathbf{X}^t, \tilde{\mathbf{D}}^t; \mathbf{D}^t) = H(\mathbf{Y}^t|\tilde{\mathbf{D}}^t; \mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1}), \tag{2.7}$$

representing the revenue function in period t given state $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ and demand realization $\mathbf{D}^t$, is also concave in the allocation decision $\mathbf{Y}^t$. The concavity property is summarized in the following proposition whose formal proof is omitted.

Proposition 2.4.1 *In period t, $\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ is concave in $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$, and $\hat{\Theta}^t(\mathbf{Y}^t|\mathbf{X}^t, \tilde{\mathbf{D}}^t; \mathbf{D}^t)$ is concave in $\mathbf{Y}^t$.*

Notice that the allocation decision $\mathbf{Y}^t$ is constrained by a bounded polyhedron defined by (2.4-2.5) and $\hat{\Theta}^t$ in (2.7) is continuous in $\mathbf{Y}^t$. Thus, there always exists an optimal allocation to the general upgrading problem in each period t. For a given state $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ and demand realization $\mathbf{D}^t$, there are two types of decisions: parallel allocations y_{ii}^t for all i $(1 \leq i \leq N)$ and upgrading decisions y_{ij}^t for classes i and j $(1 \leq i < j \leq N)$. These are dynamic decisions because they will not only determine the revenue H in the current period but also affect the future revenue $\Theta^{t+1}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1})$.

It is straightforward to solve the parallel allocation problem. In our model, the maximum revenue we can get from a unit of capacity i is α_{ii} through the parallel allocation, i.e., capacity i is used to fulfill demand class i. It is suboptimal to satisfy demand from lower classes using capacity i when there is still unmet demand i. Further, the expected value of carrying over capacity i to the next period will not exceed α_{ii}, either. Hence the optimal strategy is to use the parallel allocation as much as possible. That is, $y_{ii}^t = \min(d_i^t + \tilde{d}_i^t,\ x_i^t)$. Another implication is that in the state variable $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$, class i $(1 \leq i \leq N)$ cannot be positive in both $\mathbf{X}^t$ and $\tilde{\mathbf{D}}^t$. Thus, we can use a single variable $\mathbf{M}^t = (\mathbf{X}^t - \tilde{\mathbf{D}}^t) = (m_1^t, m_2^t, \cdots, m_N^t)^\intercal$ to represent the state at the beginning of period t (before the parallel allocation): $m_i^t > 0$ means there is positive capacity for i while $m_i^t < 0$ means there is backordered demand for i. In the rest of the chapter we will use $\mathbf{M}^t$ and $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ exchangeably.

The more challenging question is how to make the upgrading decisions after the parallel allocation. The state after the parallel allocation in period t is $(m_1^t - d_1^t, m_2^t - d_2^t, \cdots, m_N^t - d_N^t)^\intercal$. Note that $m_i^t - d_i^t > 0$ means that there is leftover capacity i, while $m_i^t - d_i^t < 0$ implies that there is unsatisfied demand i and capacity i must have been depleted. The firm needs to decide how much demand should be upgraded using higher capacities. This is equivalent to a rationing problem, i.e., how much capacity should be protected to satisfy future demand. The upgrading problem in our model is different from the one studied in SZ. Particularly, with the single-step assumption in SZ, when capacity i is depleted, classes above i and those below i become independent of each other in future periods, and thus the upgrading problem is greatly simplified because all the upgrading decisions can be solved independently. However, with the general upgrading structure in our model, the upgrading decisions after parallel allocation are no longer isolated. In this case, we may have to solve all decisions simultaneously, which could be computationally intensive. Fortunately, close scrutiny shows that the following two observations can greatly reduce the complexity of the upgrading problem.

First, the upgrading decision y_{ij}^t of using capacity i to upgrade demand j is independent of the demands and the capacities below class j. To explain, consider the last unit of capacity i that could be used to upgrade an unmet demand j. If this unit is used for upgrading, the immediate value obtained is α_{ij}. If such unit is carried over to the next period, it means that there is a corresponding unsatisfied demand j left to the next period. Notice that due to the existence of the backlogged demand j, the specific unit of capacity i will never be used to upgrade the demand below class j in any future period. This implies that we can solve the upgrading problem sequentially by starting from the highest class j with $m_j^t - d_j^t < 0$.

Second, for demand class j with $m_j^t - d_j^t < 0$, the upgrading decisions y_{ij}^t, $i = 1, ..., j-1$ can also be solved sequentially in i. Consider two capacity classes i and k $(i < k < j)$ with positive capacities after the parallel allocation. Since $\alpha_{ij} < \alpha_{kj}$ by assumption, we should first evaluate the possibility of using capacity k to upgrade demand j. After that, we consider using capacity i to satisfy demand j. Interestingly, we do not need to consider capacity i anymore if all demand in class j is satisfied by capacity k or we do not use full capacity k to upgrade demand j.

Based on these observations, the upgrading problem can be sequentially solved as follows:

Step 1: Identify the smallest j $(1 \leq j \leq N)$ with $m_j^t - d_j^t < 0$ (the highest class with unmet demand);

Step 2: For the largest i (the lowest capacity class) less than j with $m_i^t - d_i^t > 0$, determine the upgrading quantity y_{ij}^t in period t (or equivalently, the quantity of capacity i to be protected for the next period). When solving y_{ij}^t, we can ignore the classes lower than j;

Step 3: Repeat *Step 2* until all capacity classes available for upgrading demand j have been considered;

Step 4: Repeat *Step 1* until all unmet demand classes have been considered.

To summarize, the firm may allocate capacity using the so-called Parallel and Sequential Rationing (PSR) policy. Under such a policy, the firm first performs the parallel allocation on each class to satisfy new demands, and then sequentially decides upgrading quantities for classes with unmet demand.

The most crucial decision in the sequential upgrading procedure is to determine y_{ij}^t in Step 2. Consider the decision about how much capacity i should be used to upgrade demand j. It is clear that as long as the current upgrade revenue α_{ij} is greater than the expected marginal value in the future, capacity i should be used to upgrade demand j. Such an upgrading or rationing decision essentially specifies the protection levels for the capacities. Let p_{ij} be the optimal protection level of capacity i with respect to demand j, i.e., the firm should stop upgrading demand j by capacity i when the capacity level of i drops to p_{ij}. Since $\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ is concave in $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ by Proposition 2.4.1, the expected marginal value of capacity i is monotonically increasing as capacity i decreases. Hence, the protection level p_{ij} in period t is the unique lower bound above which using capacity i to upgrade demand j is profitable. Define $\frac{\partial}{\partial p}\Theta^t = \left[\frac{\partial}{\partial p^+}\Theta^t, \frac{\partial}{\partial p^-}\Theta^t\right]$ as the subdifferential of Θ^t with respect to some variable p, where $\frac{\partial}{\partial p^-}\Theta^t$ and $\frac{\partial}{\partial p^+}\Theta^t$ are the left and right derivatives, respectively. Let $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)^\intercal$ denote the state of the system immediately before the epoch of determining y_{ij}^t. The optimal protection levels can be defined as follows.

Definition 2.4.1 *The optimal protection level* $p_{ij} \geq 0$ *under state* $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)^\intercal$ *is defined as*

$$p_{ij} = \begin{cases} p & \text{if } \alpha_{ij} \in \frac{\partial}{\partial p}\Theta^{t+1}(n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t), \\ 0 & \text{if } \alpha_{ij} > \frac{\partial}{\partial p^+}\Theta^{t+1}(n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t)|_{p=0}. \end{cases} \tag{2.8}$$

With the protection levels p_{ij} and $\mathbf{N}^t$, the optimal upgrading decision y_{ij}^t is simply given by $\min\left((n_i^t - p_{ij})^+,\ (-n_j^t)^+\right)$ where $(x)^+ = \max(x, 0)$. Notice that there are 0's between classes i and j since the PSR algorithm does not consider y_{ij}^t if there exists a class s ($i < s < j$) with positive capacity or unmet demand. When class s has positive capacity, it is more profitable to upgrade demand j with capacity s instead of capacity i, and it is unnecessary for us to consider y_{ij}^t if there is capacity s remaining after solving y_{sj}^t. When there is unmet demand for class s, capacity i should upgrade demand s first, and it would be suboptimal to upgrade demand j if class s still has unmet demand after upgrading y_{is}^t.

In the next section, we will show that

$$\frac{\partial}{\partial p}\Theta^{t+1}(n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t)$$

is independent of the values of $(n_{j+1}^t, \cdots, n_N^t)$. This implies that the upgrading decision y_{ij}^t is independent of the demands and the capacities below class j. Later we can see that when solving p_{ij}, it is sufficient to use the first $i-1$ components of $\mathbf{M}^t - \mathbf{D}^t$ (i.e., the state of the system in period t after the parallel allocation) instead of $\mathbf{N}^t$ (i.e., the state of the system prior to deciding y_{ij}^t) in the PSR algorithm. This is a unique and interesting property of the general upgrading problem, allowing us to simultaneously and independently solve all protection levels based on $\mathbf{M}^t - \mathbf{D}^t$.

Before presenting the main results, we wish to further reduce the computation in the general upgrading problem by exploring its structure. With the single-step upgrading rule, SZ shows that whenever a capacity (say, i) is depleted, then the entire problem decoupled into two independent subproblems, where the first subproblem consists of products above i and the second consists of products below i (see Lemma

4 in SZ). Under the full-upgrading rule, such a property in SZ clearly does not hold. However, it can be shown that under certain conditions, our problem can also be separated into independent subproblems, as stated in the next lemma.

Lemma 2.4.1 *Consider an N-class general upgrading problem with state $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)^\intercal$ in period t. If $\sum_{s=k}^{i} n_s^t \leq 0$ for all class $k \leq i$, then the problem can be separated into two independent subproblems: an upper part consisting of classes $(1, \cdots, i)$, and a lower part consisting of classes $(i+1, \cdots, N)$.*

For convenience, we say class i is separable if it satisfies the condition stated in Lemma 2.4.1. Notice that $n_i^t \leq 0$ is not enough to split the N-class general upgrading problem since there may be class k ($k < i$) which can upgrade demands in classes $(i+1, \cdots, N)$. However, the condition in Lemma 2.4.1 determines that none of classes $(1, \cdots, i)$ has enough capacity to upgrade the demand in $(i+1, \cdots, N)$ when optimal upgrading is performed. Specifically, there may exist class $k < i$ with positive capacity which can upgrade the demand in $(i+1, \cdots, N)$, but it is more profitable for capacity k to satisfy the demand in classes $(k+1, \cdots, i)$ first, which will consume all of class k's capacity. Therefore, Lemma 2.4.1 asserts that the entire upgrading problem can be simplified by decomposition under certain conditions. That is, the profit of the N-class problem can be written as the sum of the profits from independent subproblems $(1, \cdots, i)$ and $(i+1, \cdots, N)$ whenever class i is separable. The next section presents the optimality proof and some useful properties of the PSR policy. These results apply to all the subproblems as well as to the entire upgrading problem.

2.5 Optimality and Properties of PSR

2.5.1 Optimality

The optimality proof of the PSR policy is by induction. We begin with the last period T. In the last period, since leftover capacities have no salvage value, the optimal protection levels must be zero. Specifically, for a given demand realization, the upgrading problem in the last period can be viewed as a standard transportation problem. In addition, the objective function has a special cost structure, i.e., $\alpha_{ij} > \alpha_{i,j+1}$, $\alpha_{ij} > \alpha_{i-1,j}$ for $i \leq j$, and $\alpha_{ij} + \alpha_{i'j'} = \alpha_{ij'} + \alpha_{i'j}$ if $\max(i, i') \leq \min(j, j')$. The optimal solution can be readily obtained from the following lemma.

Lemma 2.5.1 *The PSR algorithm solves the general upgrading problem* (2.2) *in period T with all protection levels being* 0.

The zero protection levels in the final period imply greedy upgrading. That is, after the parallel allocation, the sequential rationing proceeds from class 1 to N and upgrades the unmet demand by the lowest capacity classes as much as possible. Later we will show that the PSR algorithm also solves the upgrading problem (2.2) in any period t; however, the optimal protection levels are not necessarily zero.

To gain more understanding of the general upgrading problem, let us consider the protection level p_{ij} $(1 \leq i < j \leq N)$ in period $T-1$. Since the optimal p_{ij} is determined by the expected marginal value of Θ^T in (2.8), we focus on how the marginal value depends on the current state $\mathbf{N}^{T-1}$ in period $T-1$. By Lemma 2.5.1, Θ^T can be evaluated in the following three steps. First, we solve the upgrading decisions within classes $(1, \cdots, i-1)$; second, we satisfy the upgrading need that arises within classes $(i, \cdots, j)$ (note we may use capacity $k < i$ to upgrade demand); finally, we use upgrading to satisfy the unmet demand within classes $(j+1, \cdots, N)$. Lemma 2.5.2 below characterizes the relation between the optimal protection level p_{ij} in period $T-1$ and the state $\mathbf{N}^{T-1}$. As a preparation, we first introduce the concept of effective state.

Definition 2.5.1 *Consider a state vector $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)$ in period t $(1 \leq t \leq T)$. For class r $(1 \leq r \leq N)$, the effective state $\hat{\mathbf{N}}_r^t = (\hat{n}_1^t, \cdots, \hat{n}_r^t, n_{r+1}^t, \cdots, n_N^t)$ is defined as the resulting state after applying the greedy upgrading for classes* $(1, \cdots, r)$.

In fact, given any state $\mathbf{N}^t$ and its effective state $\hat{\mathbf{N}}_r^t$, if we use h $(1 \leq h \leq r)$ to denote the highest class with $\hat{n}_h^t > 0$, then class $h-1$ is separable in $\mathbf{N}^t$. To see this, note that given $\hat{n}_h^t > 0$, there is no upgrade between classes $(1, \cdots, h-1)$ and $(h, \cdots, r)$ when using the greedy upgrading. Thus, for all class $k < h$, we have $\sum_{s=k}^{h-1} n_s^t \leq \sum_{s=k}^{h-1} \hat{n}_s^t \leq 0$ since there may be upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, h-1)$ when performing the greedy upgrading. Hence, $h-1$ is separable, and classes $(1, \cdots, h-1)$ can be ignored in the subsequent allocation decisions.

Consider a state vector $\mathbf{N}^t = (n_1^t, \cdots, n_N^t)$ in period t. For $1 \leq i < j \leq N$, define

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}^t) = \frac{\partial}{\partial n_i^+}\Theta^t(\mathbf{N}^t) - \frac{\partial}{\partial n_j^-}\Theta^t(\mathbf{N}^t), \qquad \Delta_{ij}^{-+}\Theta^t(\mathbf{N}^t) = \frac{\partial}{\partial n_i^-}\Theta^t(\mathbf{N}^t) - \frac{\partial}{\partial n_j^+}\Theta^t(\mathbf{N}^t).$$

Then we have the following lemma.

Lemma 2.5.2 *Consider an N-class general upgrading problem in period $T-1$ with state vector $\mathbf{N}^{T-1}$, where $(n_{i+1}^{T-1}, \cdots, n_{j-1}^{T-1}) \leq 0$ and $n_j^{T-1} < 0$. Then,*

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{+-}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}), \qquad \Delta_{ij}^{-+}\Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{-+}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}). \qquad (2.9)$$

In addition, they are independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$.

Notice that the protection level p_{ij} in (2.8) can be equivalently defined

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}) \leq \alpha_{ij} \leq \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}),$$

where $\mathbf{N} = (n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t)$. Thus, Lemma 2.5.2 states that the optimal protection level p_{ij} in period $T-1$ is independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$, while it is affected by the classes above i through the effective state $(\hat{n}_1^{T-1}, \cdots, \hat{n}_{i-1}^{T-1})$. These results provide the rationale behind the sequential rationing in the PSR algorithm. Clearly, they will significantly simplify the optimal solution to the upgrading problem. We offer the following intuitive explanation of these results. First, we explain why $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1})$ and $\Delta_{ij}^{-+}\Theta^T(\mathbf{N}^{T-1})$ are independent of $(n_j^{T-1}, \cdots, n_N^{T-1})$. Before deciding p_{ij} or y_{ij}^{T-1}, without losing generality, we may label all units of capacity i in an increasing order of importance, with the first unit having the least importance (i.e., it must be used first in any subsequent period). Meanwhile, the unsatisfied demand in class j can be treated as a waiting line, which will be satisfied in the first-come first-served sequence. Note that deciding p_{ij} is equivalent to comparing α_{ij} with the expected value of capacity unit 1 in class i. Given the backorder assumption, capacity unit 1 can only satisfy either a future demand in classes $(i, \cdots, j-1)$ or the first unit in the waiting line of class j. Hence, the expected value of capacity unit 1 in class i is independent of states $(n_{j+1}^{T-1}, \cdots, n_N^{T-1})$. Furthermore, the above argument only relies on the fact that there exists unmet demand j. Thus, the expected value of capacity unit 1 is also independent of n_j^{T-1}, the length of the waiting line in class j.

Next, we explain the equalities in (2.9). For any class k $(1 < k < i)$ with positive capacity, it would not upgrade demand i in any optimal policy if there exists backordered demand r $(k < r < i)$, which is more valuable for capacity k than demand i. The remaining capacity of class k after upgrading all backordered demands in classes $(k+1, \cdots, i-1)$ equals $\hat{n}_k^{T-1}$ as defined in the effective state. Therefore, the expected future value of capacity i in period $T-1$ should equivalently depend on the effective state $(\hat{n}_1^{T-1}, \cdots, \hat{n}_{i-1}^{T-1})$, which are non-negative when classes $(1, \cdots, N)$ are not separable. Note that this argument applies to any period t.

Now we are in the position to use induction to prove the optimality of the PSR.

Proposition 2.5.1 *1. The PSR algorithm solves the general upgrading problem in period t;*

2. For a state vector $\mathbf{N}^t$ with $(n_{i+1}^t, \cdots, n_{j-1}^t) \leq 0$ and $n_j^t < 0$, we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t).$$

In addition, they are independent of the values of $(n_j^t, \cdots, n_N^t)$.

For any given period t under the PSR algorithm, the effective states of all intermediate states for classes $(1, 2, \ldots, i-1)$ are the same before we exhaust the capacity of class i. Thus, Proposition 2.5.1 implies that when solving p_{ij}, it is sufficient to use the first $i-1$ components of $\mathbf{M}^t - \mathbf{D}^t$, the state of the system in period t after the parallel allocation. Specifically, for any classes i and j $(1 \leq i < j \leq N)$

with $n_i^t > 0$ and $n_j^t < 0$, the protection level p_{ij} can be immediately determined by $\frac{\partial}{\partial p}\Theta^{t+1}(m_1^t - d_1^t, \cdots, m_{i-1}^t - d_{i-1}^t, p, 0, \cdots, 0, -p, 0, \cdots, 0)$.

2.5.2 Properties of Protection Levels

After establishing the optimality of the PSR algorithm, we explore some important properties related to the optimal protection levels from the PSR algorithm.

First, if both the initial capacity $\mathbf{X}^1$ and all demands are integer valued, similar to SZ, we can prove that there exists an integer valued optimal policy generated by the PSR algorithm.

Proposition 2.5.2 *If initial capacity $\mathbf{X}^1$ and demand $\mathbf{D}^1, \cdots, \mathbf{D}^T$ are integer valued, there exists an integer valued optimal policy $\mathbf{Y}^1, \cdots, \mathbf{Y}^T$ derived by the PSR algorithm.*

To further characterize the protection level p_{ij} defined in (2.8), we need to deal with the marginal value of Θ^t with respect to each capacity level and unmet demand level. Intuitively, one may think that the profit will be higher if there is an additional unit of capacity $i-1$ $(1 < i \leq N)$ rather than capacity i. But this is not necessarily true. When making upgrading decisions, one more unit of capacity from the higher class $i-1$ always provides more flexibility, but such a flexibility does not necessarily mean higher profit since $\alpha_{ij} > \alpha_{i-1,j}$ $(i < j)$ by our model assumption. Similarly, one more unit of demand in a lower class, which can be upgraded by more classes of capacities, has similar advantage but can not guarantee greater profit because $\alpha_{ij} < \alpha_{i,j+1}$ $(i \leq j)$. However, we can provide some bounds on such profit differences. With these bounds, we show two different monotone properties of the protection levels. First, since lower demand has less value for any capacity, the protection level should increase in the class index of demand.

Proposition 2.5.3 *For the same $(n_1^t, \cdots, n_{i-1}^t)$ in period t $(1 \leq t \leq T)$, $p_{ij} \leq p_{i,j+1}$ when $i < j$.*

Because the general upgrading problem in period T is a transportation problem, $\Theta^T(\mathbf{X}^T, \tilde{\mathbf{D}}^T)$ is submodular in $(\mathbf{X}^T, -\tilde{\mathbf{D}}^T)$ (see Topkis 1998). This implies the protection level p_{ij} in period $T-1$ under state $\mathbf{N}^{T-1}$ is decreasing in $(n_1^{T-1}, \cdots, n_{i-1}^{T-1})$. In fact, the same monotonicity holds in earlier periods.

Proposition 2.5.4 *The optimal protection level p_{ij} $(1 \leq i < j \leq N)$ in period t $(1 \leq t \leq T)$ are decreasing in $(n_1^t, \cdots, n_{i-1}^t)$.*

For any class i $(1 \leq i \leq N)$, this result assures that the more capacities (or less back-ordered demands) in classes higher than i, the more upgrades can be offered by class i. Note that larger $(n_1^t, \cdots, n_{i-1}^t)$ means higher probability of demand i being upgraded in remaining periods, which decreases the expected marginal value of capacity i and gives class i a greater incentive to upgrade lower demands in the current period.

It is noteworthy that although the result for the last period can be proved using lattice programming in Topkis (1998), the commonly used preservation property of supermodularity under optimization operations, Theorem 2.7.6 in Topkis (1998), does not apply. Therefore, our proof relies heavily on the structure of the general upgrading problem and fully utilizes the optimality of the PSR algorithm.

One may ask whether the optimal protection levels are decreasing over time, i.e, the protection level would be lower if there are fewer periods to go. Interestingly, though this is true in SZ, it does not hold in our upgrading problem. This is mainly due to the existence of the backorder cost. Note that the purpose of the protection levels is to balance the goodwill loss of carrying backorders and the revenue loss of losing future demand from the same class. For early periods that are still far away from the end of the horizon, because a backorder causes the goodwill loss in each period until it is upgraded, the protection levels may be lower to avoid high backorder costs; in contrast, when it is close to the end of the horizon, the protection levels may come back up because carrying backorders will be less costly.

We may use a 2-product 3-period example to explain this counter-intuitive result. Let $(2, -2)$ be the state after the parallel allocation, $\mathbf{D}^2 = (0, 0)$ and $\mathbf{D}^3 = (1, 0)$ with probability 1. Working backwardly to solve the p_{12} in period 2, since

$$\Theta^3(2, -2) - \Theta^3(1, -1) = \alpha_{12} - g_2 < \alpha_{12}, \quad \Theta^3(1, -1) - \Theta^3(0, 0) = \alpha_{11} - g_2,$$

we have $p_{12} = 1$ in period 2 if $\alpha_{11} - g_2 > \alpha_{12}$. Since $\mathbf{D}^2 = (0, 0)$, there is

$$\Theta^2(2, -2) - \Theta^2(1, -1) = \alpha_{12} - g_2 < \alpha_{12}, \quad \Theta^2(1, -1) - \Theta^2(0, 0) = \alpha_{11} - 2g_2.$$

Therefore, if $\alpha_{11} - g_2 > \alpha_{12} > \alpha_{11} - 2g_2$, the optimal protection level p_{12} increases from 0 in period 1 to 1 in period 2. That is, the protection level does not necessarily decrease over time in our general upgrading problem[2].

2.6 Multiple Horizons with Capacity Replenishment

Now we extend our model to multiple horizons with capacity replenishment. Specifically, there are K ($K \geq 1$) horizons, each consisting of T periods. Demands across horizons are independent and identically distributed. At the beginning of each horizon k ($1 \leq k \leq K$), the firm observes the leftover capacity $\mathbf{X}$ and unmet demand $\tilde{\mathbf{D}}$ carried over from the previous horizon. There are two decisions for the firm in each horizon: First, the firm decides how much capacity to replenish; second, it allocates capacity to satisfy demand as formulated in (2.2). For completeness, we assume unmet demand after the K-th horizon can also be satisfied by purchasing additional capacity. There is a unit cost vector $\mathbf{C} = (c_1, \cdots, c_N) \in \Re_+^N$ for capacity replenishment. The remaining capacity at the end of each horizon incurs a holding cost $\mathbf{h} = (h_1, \cdots, h_N) \in \Re_+^N$. The leftover capacity after the K-th horizon can be sold at the initial capacity cost, i.e., it has salvage value $\mathbf{C}$. Revenues and costs are discounted at a rate γ ($0 < \gamma \leq 1$) for each horizon. The rest of the model setting remains the same as in Section 2.3.

[2]This counter-intuitive example remains valid for any goodwill cost g_2 if the length T satisfies $\alpha_{11} - (T-2)g_2 > \alpha_{12} > \alpha_{11} - (T-1)g_2$ and $\mathbf{D}^2 = \cdots = \mathbf{D}^{T-1} = (0, 0)$ and $\mathbf{D}^T = (1, 0)$.

In the replenishment model, at the end of the last horizon, leftover capacity and unmet demand have a positive end-value given by

$$\Theta^{T+1}(\mathbf{X}^{T+1}, \tilde{\mathbf{D}}^{T+1}) = (\gamma\mathbf{C} - \mathbf{h})\mathbf{X}^{T+1} + \gamma(\alpha - \mathbf{C})\tilde{\mathbf{D}}^{T+1}, \tag{2.10}$$

where $\alpha = (\alpha_{11}, \cdots, \alpha_{NN})$ is the revenue from parallel allocation. This end-value is different from the single-horizon model with $\Theta^{T+1}(\mathbf{X}^{T+1}, \tilde{\mathbf{D}}^{T+1}) \equiv \mathbf{0}$ in Section 2.3. Let $\Pi(\mathbf{X}; \gamma\mathbf{C} - \mathbf{h}; \gamma(\alpha - \mathbf{C}))$ denote the optimal profit in the replenishment model with initial capacity $\mathbf{X}$ and $K = 1$.

From the proof of Proposition 2.4.1, $\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$, which is similarly defined as (2.2) with $\Theta^{T+1} \equiv \mathbf{0}$ being replaced by Θ^{T+1} in (2.10), is still concave in $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$. In particular, $\Pi(\mathbf{X}; \gamma\mathbf{C} - \mathbf{h}; \gamma(\alpha - \mathbf{C}))$ is concave in $\mathbf{X}$ from the concavity of $\Theta^1(\mathbf{X}, \mathbf{0})$. Furthermore, similarly as (2.6), we can show that there exists an optimizer $\mathbf{X}^*$ for the concave function $\Pi(\mathbf{X}; \gamma\mathbf{C} - \mathbf{h}; \gamma(\alpha - \mathbf{C}))$:

$$\mathbf{X}^* \in \arg\max_{\mathbf{X} \in \Re^N_+} \Pi(\mathbf{X}; \gamma\mathbf{C} - \mathbf{h}; \gamma(\alpha - \mathbf{C})). \tag{2.11}$$

Note that $\mathbf{X}^*$ is the optimal capacity level for the replenishment model with $K = 1$.

The next proposition characterizes the optimal capacity replenishment and allocation policies in the multi-horizon model, given that the firm starts with an initial capacity $\mathbf{X} \leq \mathbf{X}^*$. It shows that the structural results from the base model in Section 2.3 remain valid in the multi-horizon model, thus we will focus on the base model in the rest of the chapter.

Proposition 2.6.1 *Suppose the firm starts with an initial capacity* $\mathbf{X} \leq \mathbf{X}^*$*. The firm's optimal replenishment policy in the multi-horizon model is a base stock policy with the optimal base stock level* $\mathbf{X}^*$ *in* (2.11)*. Furthermore, the PSR algorithm solves the optimal allocation decisions within each horizon.*

2.7 Heuristics and Benchmark Models

So far we have characterized the structure of the optimal allocation policy for our dynamic capacity management problem. In this section, we propose an effective heuristic for solving the optimal allocation policy. For future comparison, we also present two benchmark models that are simplified versions of the general upgrading problem.

2.7.1 Heuristics

We have shown that the PSR algorithm yields the optimal allocation decisions $\mathbf{Y}^t$ for the firm in period t, which essentially consists of the optimal protection levels for each capacity. These optimal protection levels are defined by (2.8) and can be solved by backward induction. For instance, the optimal protection levels in period t depend on the revenue-to-go function Θ^{t+1}, which is determined by the protection levels used in period $t+1$. To evaluate Θ^{t+1}, one needs to derive the optimal protection levels for all possible states in period $t+1$ (note that these protection levels, though possessing the appealing properties established earlier, are still state-dependent). Due to the curse of dimensionality, solving the exact optimal upgrading decisions is quite difficult for large problems[3]. Therefore, we need to search for heuristics that can solve the problem effectively.

Since solving the allocation decision is equivalent to solving the Bellman equation (2.2) in period t, in order to develop efficient heuristics, we focus on the one-step lookahead policy which hinges upon reasonable approximations to Θ^{t+1}. The basic idea is as follows. Suppose $\bar{\Theta}^{t+1}_{\text{approx}}$ is an easy-to-compute and acceptable approximation to Θ^{t+1}. Given the initial state $(\mathbf{X}^t, \tilde{\mathbf{D}}^t)$ and the realized demand $\mathbf{D}^t$ in period t, we solve the following optimization program

$$\max_{\mathbf{Y}^t} \left[H(\mathbf{Y}^t | \tilde{\mathbf{D}}^t; \mathbf{D}^t) + \bar{\Theta}^{t+1}_{\text{approx}}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1}) \right], \tag{2.12}$$

[3]To deal with the dimensionality issue, SZ propose a series of bounds to approximate the optimal protection levels. For instance, when computing the protection level for product i, one may consider only the capacity for $i-1$, while assuming the products above $i-1$ to be either ∞ (this gives a lower bound of the protection level) or 0 (this gives an upper bound). It has been found that under the single-step upgrading assumption, these bounds are very tight and yield nearly optimal revenue for the firm. However, such bounds do not work well in our model, where general upgrading is allowed.

and obtain the corresponding allocation decision $\mathbf{Y}^t_{\text{approx}}(\mathbf{X}^t, \tilde{\mathbf{D}}^t|\mathbf{D}^t)$ in period t. Let Θ^t_{approx} be the revenue collected by applying the policy $(\mathbf{Y}^t_{\text{approx}}, \cdots, \mathbf{Y}^T_{\text{approx}})$ from period t to T. For simplicity, we do not distinguish between the policy and the decision (e.g., $\mathbf{Y}^t_{\text{approx}}$ and $\mathbf{Y}^t_{\text{approx}}(\mathbf{X}^t, \tilde{\mathbf{D}}^t|\mathbf{D}^t)$), since the proper interpretation is usually clear from the context. Note that $\mathbf{Y}^t_{\text{approx}}$ is a suboptimal policy in the general upgrading problem and $\Theta^t_{\text{approx}} \neq \bar{\Theta}^t_{\text{approx}}$ in general. Moreover, $\Theta^t_{\text{approx}}(\mathbf{N}^t) \leq \Theta^t(\mathbf{N}^t)$ for any state $\mathbf{N}^t$ in period t since $\Theta^t(\mathbf{N}^t)$ adopts the optimal policy from period t to T.

As pointed out by Bertsekas (2005b), even with readily available revenue-to-go approximations, Θ^t_{approx} may still involve substantial computational efforts. A number of simplifications of the optimization in (2.12), including different $\bar{\Theta}^{t+1}_{\text{approx}}$ functions, have been considered. Here we present two of them that stand out both in terms of computational time and in terms of revenue performance. Because of the linearity in the upgrading problem, the first natural candidate is the traditional Certainty Equivalence Control (CEC) heuristic in the literature (see Bertsekas 2005a, for example). The CEC is a suboptimal control which treats the uncertain quantities as fixed typical values in the stochastic dynamic programming. In our case, we use demand means as typical values in evaluating the function $\bar{\Theta}^{t+1}_{\text{approx}}$. Thus, under the CEC, expectation calculations are no longer relevant, which can alleviate the computational burden in our problem. Specifically, the optimal allocation policy in period t is solved together with all future periods where the mean demand is used as approximation. That is, the optimal allocation decision $\mathbf{Y}^t_{\mathrm{CEC}}$ in the CEC heuristic will be obtained by solving the following linear program:

$$\begin{aligned}
\max_{(\mathbf{Y}^t_{\mathrm{CEC}},\bar{\mathbf{Y}}^{t+1},\cdots,\bar{\mathbf{Y}}^T)\geq 0} \quad & \left\{ H(\mathbf{Y}^t_{\mathrm{CEC}}|\tilde{\mathbf{D}}^t;\mathbf{D}^t) + \sum_{l=t+1}^{T} H(\bar{\mathbf{Y}}^l|\tilde{\mathbf{D}}^l;\mu^l) \right\} \\
\text{s.t.} \quad & \tilde{\mathbf{D}}^{t+1} = \tilde{\mathbf{D}}^t + \mathbf{D}^t - (\mathbf{Y}^t_{\mathrm{CEC}})^\intercal \mathbf{1}, \\
& \tilde{\mathbf{D}}^{l+1} = \tilde{\mathbf{D}}^l + \mu^l - (\bar{\mathbf{Y}}^l)^\intercal \mathbf{1}, \qquad l = t+1,\cdots,T \\
& \left(\mathbf{Y}^t_{\mathrm{CEC}} + \sum_{l=t+1}^{T} \bar{\mathbf{Y}}^l \right) \mathbf{1} \leq \mathbf{X}^t, \\
& (\mathbf{Y}^t_{\mathrm{CEC}})^\intercal \mathbf{1} \leq \tilde{\mathbf{D}}^t + \mathbf{D}^t, \\
& \left(\mathbf{Y}^t_{\mathrm{CEC}} + \sum_{l=t+1}^{k} \bar{\mathbf{Y}}^l \right)^\intercal \mathbf{1} \leq \tilde{\mathbf{D}}^t + \mathbf{D}^t + \sum_{l=t+1}^{k} \mu^l, \qquad k = t+1,\cdots,T,
\end{aligned} \tag{2.13}$$

where $\mathbf{X}^t$, $\tilde{\mathbf{D}}^t$, and $\mathbf{D}^t$ are the capacities, backorders, and realized demand in period t, respectively, and $(\mu^1, \mu^2, \cdots, \mu^T)$ denote the mean demand vectors.

The solution to (2.13) yields the allocation decisions $(\mathbf{Y}^t_{\mathrm{CEC}}, \bar{\mathbf{Y}}^{t+1}, \cdots, \bar{\mathbf{Y}}^T)$ for periods from t to T, where $(\bar{\mathbf{Y}}^{t+1}, \cdots, \bar{\mathbf{Y}}^T)$ are discarded in the subsequent periods. We implement $\mathbf{Y}^t_{\mathrm{CEC}}$ as the allocation decision for period t and then move on to solve problem (2.13) in period $t+1$. Let Θ^t_{CEC} be the revenue collected by applying the policy $(\mathbf{Y}^t_{\mathrm{CEC}}, \cdots, \mathbf{Y}^T_{\mathrm{CEC}})$ in periods from t to T. Define $\Pi_{\mathrm{CEC}}(\mathbf{X}) = \Theta^1_{\mathrm{CEC}}(\mathbf{X}, \mathbf{0})$ as the firm's total revenue given initial capacity $\mathbf{X}$ under the CEC heuristic.

Although the above CEC heuristic can simplify our problem, its computational time is still quite long. Consider an N-product general upgrading problem with t periods remaining, the CEC heuristic solves the allocation decisions in the current period as a transportation problem with N classes of capacities and tN classes of demands, whose running time is $O(tN^3(\log(tN) + N\log N))$ (see Brenner 2008). In addition, the optimal allocation is derived from the linear program in (2.13), which does not use the PSR procedure and the marginal analysis in (2.8). This means that the CEC might be further improved by exploiting the special properties inherited in our upgrading problem.

To this end, we further simplify the revenue-to-go function by applying greedy upgrading. So the approximation to Θ^{t+1} consists of two components: certainty equivalence control (CEC) and greedy upgrading. Under the CEC, again the mean demand is used as an approximation in all future periods. At the same time, $\bar{\Theta}^{t+1}_{\text{approx}}$ is simplified by adopting greedy upgrading from periods $t+1$ to T rather than solving the linear program as in the CEC heuristic. Such simplification, though suboptimal, is much easier to compute than the linear program[4]. Given these characteristics of the approximation, we call it refined certainty equivalence control (RCEC) and write $\bar{\Theta}^{t+1}_{\text{approx}}$ as $\bar{\Theta}^{t+1}_{\text{RCEC}}$. In addition to the above approximation, the RCEC heuristic then calculates the protection levels in (2.8) by replacing Θ^{t+1} with $\bar{\Theta}^{t+1}_{\text{RCEC}}$, and determines the allocation decision $\mathbf{Y}^t_{\text{RCEC}}$ in period t by performing the PSR algorithm to solve the following program

$$\max_{\mathbf{Y}^t} \left[H(\mathbf{Y}^t | \tilde{\mathbf{D}}^t; \mathbf{D}^t) + \bar{\Theta}^{t+1}_{\text{RCEC}}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1}) \right].$$

Note that $\bar{\Theta}^s_{\text{RCEC}}$ $(s \geq t+1)$ can be defined recursively as follow:

$$\bar{\Theta}^s_{\text{RCEC}}(\mathbf{X}^s, \tilde{\mathbf{D}}^s) = H(\mathbf{Y}^s_\mu | \tilde{\mathbf{D}}^s; \mu^s) + \bar{\Theta}^{s+1}_{\text{RCEC}}(\mathbf{X}^{s+1}; \tilde{\mathbf{D}}^{s+1}), \tag{2.14}$$

where $\mathbf{X}^{s+1} = \mathbf{X}^s - \mathbf{Y}^s_\mu \mathbf{1}$, $\tilde{\mathbf{D}}^{s+1} = \tilde{\mathbf{D}}^s + \mu^s - (\mathbf{Y}^s_\mu)^\intercal \mathbf{1}$, $\bar{\Theta}^{T+1}_{\text{RCEC}} \equiv 0$, and $\mathbf{Y}^s_\mu = \big(y^s_{ij}(\mu)\big)_{N\times N}$ is the solution to the following linear program:

$$\max_{\mathbf{Y}^s_\mu \geq 0} \left\{ \sum_{1 \leq i \leq j \leq N} \alpha_{ij} y^s_{ij}(\mu) \middle| (\mathbf{Y}^s_\mu)^\intercal \mathbf{1} \leq \mu^s + \tilde{\mathbf{D}}^s,\ \mathbf{Y}^s_\mu \mathbf{1} \leq \mathbf{X}^s \right\}.$$

Given the protection levels derived from $\bar{\Theta}^{t+1}_{\text{RCEC}}$, $\mathbf{Y}^t_{\text{RCEC}}$ is the allocation policy in period t solved by the PSR algorithm, and Θ^t_{RCEC} is the revenue collected by applying policy $(\mathbf{Y}^t_{\text{RCEC}}, \cdots, \mathbf{Y}^T_{\text{RCEC}})$ in period t to T. Define $\Pi_{\text{RCEC}}(\mathbf{X}) = \Theta^1_{\text{RCEC}}(\mathbf{X}, \mathbf{0})$ as the firm's total revenue given initial capacity $\mathbf{X}$ under the RCEC heuristic, and $\mathbf{X}_{\text{RCEC}}$ as the optimal capacity that maximizes $\Pi_{\text{RCEC}}(\mathbf{X})$.

[4]We have tested the heuristic without the greedy upgrading and found that the performance is almost identical. That is, the use of greedy upgrading in this heuristic can significantly reduce the computational complexity but has a negligible impact on the revenue performance.

Now we analyze the running time of the RCEC. Although greedy upgrading (rather the optimal allocation) is used in $\bar{\Theta}^{t+1}_{\text{RCEC}}$, it can be shown that for any state $\mathbf{N}^t = (n_1^t, \cdots, n_N^t)$,

$$\frac{\partial}{\partial p}\bar{\Theta}^{t+1}_{\text{RCEC}}(n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t) \tag{2.15}$$

is decreasing in p[5], so the protection levels can be solved by the binary search, and it suffices to examine whether the protection level p_{ij} is between $\max(n_i^t + n_j^t, 0)$ and n_i^t. If the binary search calls the greedy upgrading more than twice, then it implies the case that there remain both surplus capacity i and unmet demand j after performing the y_{ij} allocation. Thus, the number of calls of the greedy upgrading is at most two when solving each p_{sr} ($i \leq s < r \leq j$); otherwise there exists either surplus capacity s or unmet demand r, and the upgrade quantity y_{ij} must be zero by the PSR. Furthermore, from the sequential procedure defined in PSR, there is no upgrade between classes $(1, \cdots, i-1)$ and $(j, \cdots, N)$ in this case, and it is unnecessary to compute the protection levels between these two sets. Consequently, the N classes can be partitioned into several blocks, say K blocks, and in each block there exists at most one pair of i and j such that the greedy upgrading is called more than twice to determine p_{ij}. For the block k ($1 \leq k \leq K$) with size n_k ($2 \leq n_k \leq N$), the number of greedy upgradings is no more than $O(n_k^2 + \log|X|)$, where $|X|$ is the upper bound of the initial capacity in each class. Since there is no upgrade between blocks, to solve the allocation decision in each period, the total number of calls of the greedy upgrading would be bounded by $O(N^2 + N\log|X|)$.

Consider an N-product general upgrading problem with t periods remaining. Since greedy upgrading can be solved in the running time of $O(tN^2)$, from the above analysis, the RCEC has a running time of $O(tN^3(N+\log|X|))$ in the worst scenario, which is significantly shorter than the CEC when $|X|$ is moderate. More appealingly, the

[5]Since future demands are known, there exists a period s ($t+1 \leq s \leq T$) in which capacity i will be depleted. From the expression in (2.15), a marginal change of p only affects the greedy upgrading in period s because both capacity i and backorder demand j change simultaneously in p. In particular, capacity i is used to sequentially satisfy demands from class i to j in period s. As p increases, the additional units of capacity i will be used to satisfy demands from lower classes that have smaller profit margins. Thus, the partial derivative is a decreasing step function of p.

PSR algorithm can further reduce the computational complexity in practice. Recall the discussion right after Proposition 2.5.1, the protection level p_{ij} $(1 \le i < j \le N)$ in period t only depends on the effective capacities above i, which are decided by $\mathbf{M}^t - \mathbf{D}^t$. Thus, we can use parallel computing technique and solve all protection levels independently based on $\mathbf{M}^t - \mathbf{D}^t$.

A common feature of the RCEC and CEC heuristics is that both use mean demand in future periods as an approximation. However, there is a critical difference between these two heuristics. In the RCEC, the PSR procedure is used; in particular, the optimal protection level is determined using (2.8) (i.e., by comparing the upgrading value to the future marginal value). By contrast, in the CEC, the optimal allocation is derived from the linear program in (2.13), which utilizes neither the PSR procedure nor (2.8). From our observations in the numerical study, the adoption of the PSR algorithm in the RCEC plays an important role in both reducing the computational complexity and improving the approximation performance, which will be discussed in Section 2.8.1.

2.7.2 Benchmark Models

For future comparison, we introduce two benchmark models in this subsection. The first one is called the crystal ball (CB) model. In this model, the firm has perfect demand forecast when allocating the capacities in each period. Such a benchmark has been widely adopted in the literature because it offers the "perfect hindsight" upper bound of the firm's optimal profit. For instance, it has been used in SZ but is called static model because the firm essentially faces a static capacity allocation problem given complete demand information. Let ω represent a sample path of demand $(\mathbf{D}^1, \cdots, \mathbf{D}^T)$ over the sales horizon, and $\mathbf{D}^t(\omega)$ the demand in period t on sample path ω. Then, the firm's expected profit from period t to T is defined as $\mathbb{E}_\omega[\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t; \omega)]$, where

$$
\begin{aligned}
\Theta^t(\mathbf{X}^t, \tilde{\mathbf{D}}^t; \omega) = \max_{\mathbf{Y}^t,\cdots,\mathbf{Y}^T} & \sum_{l=t}^{T} H(\mathbf{Y}^l | \tilde{\mathbf{D}}^l; \mathbf{D}^l(\omega)) \\
\text{s.t.} \quad & \tilde{\mathbf{D}}^{l+1} = \tilde{\mathbf{D}}^l + \mathbf{D}^l(\omega) - (\mathbf{Y}^l)^\intercal \mathbf{1} \qquad l = t, \cdots, T \\
& \sum_{l=t}^{T} \mathbf{Y}^l \mathbf{1} \leq \mathbf{X}^t, \\
& \sum_{l=t}^{k} (\mathbf{Y}^l)^\intercal \mathbf{1} \leq \tilde{\mathbf{D}}^t + \sum_{l=t}^{k} \mathbf{D}^l(\omega), \qquad k = t, \cdots, T, \\
& \mathbf{Y}^l \geq 0, \qquad l = t, \cdots, T.
\end{aligned}
$$

The firm's optimal profit in the crystal ball model is given by

$$
\max_{\mathbf{X}^1 \in \Re_+^N} \Pi_{\mathrm{CB}}(\mathbf{X}^1) = \max_{\mathbf{X}^1 \in \Re_+^N} \left\{ \mathbb{E}_{\omega}[\Theta^1(\mathbf{X}^1, \mathbf{0}; \omega)] - \mathbf{C}\mathbf{X}^1 \right\}, \tag{2.16}
$$

which can be used to benchmark the performance of our heuristic in the dynamic upgrading problem.

The second benchmark is the model without product upgrading. In this case, the firm's problem reduces to N independent newsvendors (NV) with backorders. The firm's expected profit can be written as

$$
\begin{aligned}
\max_{\mathbf{X}^1 \in \Re_+^N} \Pi_{\mathrm{NV}}(\mathbf{X}^1) = \max_{\mathbf{X}^1 \in \Re_+^N} & \left\{ \mathop{\mathbb{E}}_{\{\mathbf{D}^1,\cdots,\mathbf{D}^T\}} \sum_{s=1}^{N} \sum_{t=1}^{T} \left[\alpha_{ss} \min(x_s^t, d_s^t) - g_s(\tilde{d}_s^t + d_s^t) \right] - \mathbf{C}\mathbf{X}^1 \right\} \\
\text{s.t.} \quad & x_s^{t+1} = (x_s^t - d_s^t)^+, \ \tilde{d}_s^{t+1} = \tilde{d}_s^t + (d_s^t - x_s^t)^+, \\
& x_s^1 = (\mathbf{X}^1)_s, \ d_s^t = (\mathbf{D}^t)_s, \qquad s = 1, \cdots, N, \ t = 1, \cdots, T.
\end{aligned} \tag{2.17}
$$

Note that although the two benchmark models (CB and NV) are similar to the static and independent newsvendor models used in SZ, due to the backlog assumption, the firm has to allocate capacity in each period in our model, rather than accumulate the demand for the entire selling season and then allocate the capacity as in SZ.

2.8 Numerical Studies

In this section, we conduct numerical studies to derive insights into the capacity management problem. First, we test the performance of the RCEC heuristic proposed in the previous section. After that, by using the heuristic and benchmark models, we investigate the importance of the allocation mechanism and the capacity sizing decision. For simplicity, we focus on integral demands.

2.8.1 Performance of RCEC

Due to the complexity of the problem, we use extensive numerical experiments to test the performance of the heuristics. These experiments are conducted using MATLAB R2013a on an Intel Core i7-2600 desktop with 12G RAM. We focus on the RCEC heuristic because it will be used later for further numerical investigation.

The first set of experiments we consider has $N = 4$ and $T = 3$. For this problem size, we are able to use backward induction to evaluate the firm's optimal profit $\Pi(\mathbf{X})$ given in (2.1). Later we will also discuss the performance of the RCEC for large problem sizes where it is difficult to evaluate $\Pi(\mathbf{X})$ directly. Given an initial capacity $\mathbf{X} \in \Re_+^N$, define the performance measure as

$$\Delta_{\text{opt}} = \left| \frac{\Pi_{\text{RCEC}}(\mathbf{X}) - \Pi(\mathbf{X})}{\Pi(\mathbf{X})} \right| * 100\%, \tag{2.18}$$

i.e., the percentage of profit loss by using $\Pi_{\text{RCEC}}(\mathbf{X})$ rather than $\Pi(\mathbf{X})$.

To calculate $\Pi(\mathbf{X}) = \Theta^1(\mathbf{X}, \mathbf{0})$, we use the Monte Carlo method and consider a comprehensive range of scenarios, which capture different fluctuation patterns of demand means along the selling horizon (i.e., variation of $\mathbb{E}[\mathbf{D}^t]$ from $t = 1$ to T), different correlations between classes of demands in each period (i.e., $\text{Corr}(d_i^t, d_j^t)$ for all $1 \le i \le j \le N$), different demand distributions (i.e., Normal distribution and Poisson distribution), and various economic parameters (i.e., revenue $(r_1, \cdots, r_N)$, goodwill cost $(g_1, \cdots, g_N)$, usage cost $(u_1, \cdots, u_N)$ and capacity cost $(c_1, \cdots, c_N)$). Furthermore, to ensure the robustness of the results, we also test a number of different initial capacities $\mathbf{X}$ used in (2.18), which consist of both realistic and extreme scenarios. In total there are 4212 experiments in this numerical study. A full description of the setup of the numerical study is lengthy and thus given in the appendix.

The statistics for the Δ_{opt} value are reported in Table 2.1. It can be seen that the

Mean	Std.	Median	90%-percentile	Max.
0.40336	1.13279	0.14540	0.77343	17.98673

Table 2.1.: The percentage profit loss (Δ_{opt}) of RCEC relative to the optimal solution.

RCEC performs very well in this numerical study. Among all the experiments tested, the 90^{th} percentile of the profit loss is 0.77%, and the average is 0.40%.

Next we test the performance of the RCEC in larger problems. Specifically, we consider problems with $N = 5$ products and up to $T = 30$ periods. Given such sizes, it is extremely time-consuming to evaluate the optimal revenue function $\Pi(\mathbf{X})$. Instead, we use $\Pi_{\text{CB}}(\mathbf{X})$ from the crystal ball (CB) model defined in (2.16) as the benchmark for comparison. Recall that $\Pi_{\text{CB}}(\mathbf{X})$ is an upper bound of the optimal revenue $\Pi(\mathbf{X})$ for any $\mathbf{X}$, and the following relationship holds: $\Pi_{\text{CB}}(\mathbf{X}) \geq \Pi(\mathbf{X}) \geq \Pi_{\text{RCEC}}(\mathbf{X})$. Define

$$\Delta_{\text{CB}} = \left| \frac{\Pi_{\text{RCEC}}(\mathbf{X}) - \Pi_{\text{CB}}(\mathbf{X})}{\Pi_{\text{CB}}(\mathbf{X})} \right| * 100\%.$$

Then Δ_{CB} is an upper bound of Δ_{opt}, the percentage profit loss of the RCEC (i.e., $\Pi_{\text{RCEC}}(\mathbf{X})$) relative to the optimal revenue (i.e., $\Pi(\mathbf{X})$).

Similar experiment design has been used as Table 2.1 except that now we consider 5 products with several different T values. This allows us to examine up to 4 levels of upgrading. Also by varying T we can study the impact of the number of periods (or the frequency of upgrading decisions) on the problem. Specifically, T takes values from a set $\{3, 15, 30\}$. For each T, there are 13260 experiments in total in this numerical study. To save space, we provide a detailed description in the appendix.

We summarize the statistics of Δ_{CB} for different T's in Table 2.2. It shows that the value of Δ_{CB} is increasing in the number of periods, T. The RCEC ignores the randomness of the demand in future periods (recall that the mean demand is used). Thus, compared to $\Pi_{\text{CB}}(\mathbf{X})$, more demand information is lost as T increases. Table 2.2 also indicates that the value of Δ_{CB} is very small in general: Even for $T = 30$, Δ_{CB} is 5.37% at the 90^{th} percentile, and the average is about 2.37%. This observation has two implications. First, since Δ_{CB} is the upper bound of Δ_{opt}, we know that Δ_{opt} is also very small in the tested examples. This means that for the 5-product numerical experiments, the RCEC also performs well. Second, the observation implies that the difference between $\Pi_{\text{CB}}(\mathbf{X})$ and $\Pi(\mathbf{X})$ is small. In other words, the value of advance demand information is generally small. Such a result is consistent with some of the findings reported in the literature. For instance, SZ finds from numerical study that when the optimal upgrading policy is used, the firm's expected revenue is consistently within 1% of the revenue in a static model (i.e., the crystal ball model). Similarly, Acimovic and Graves (2013) find in a dynamic order fulfillment setting that the crystal ball model improves the performance of the proposed heuristic by 2%, i.e., the performance difference between the crystal ball model and the true optimum is smaller than 2%.

T	Mean	Std.	Median	90%-percentile	Max.
3	0.14000	0.38286	0.00428	0.33580	6.73835
15	1.51822	2.51158	0.23127	4.82826	12.05775
30	2.37289	3.35659	0.42136	5.36783	23.37090

Table 2.2.: The percentage profit loss (Δ_{CB}) of RCEC relative to the CB solution.

We now compare the performances of the RCEC and the CEC. Define the ratio

$$\gamma = \frac{\Pi_{\text{RCEC}}(\mathbf{X})}{\Pi_{\text{CEC}}(\mathbf{X})}$$

to measure the relative performances of the two heuristics. So a ratio higher (lower) than 1 implies that the RCEC outperforms (underperforms) the CEC. We calculate the ratio for the problem instances used in the numerical study underlying Table 2.2

(i.e., $N = 5$ and $T = \{3, 15, 30\}$). The statistics of the ratio values are summarized in Table 2.3 since the results are consistent across different T's. Meanwhile, as we mentioned earlier, we also compare the actual computation time of the CEC and the RCEC heuristics in these instances. Specifically, we use MOSEK toolbox for MATLAB version 7 to solve the linear program in (2.13) in the CEC heuristic, and we apply the binary search to solve the protection levels in (2.8) while replacing Θ^{t+1} by $\bar{\Theta}^{t+1}_{\mathrm{RCEC}}$ in (2.14). Similarly, we define

$$\gamma_{\mathrm{time}} = \frac{\text{Time for solving } \Pi_{\mathrm{RCEC}}(\mathbf{X})}{\text{Time for solving } \Pi_{\mathrm{CEC}}(\mathbf{X})},$$

whose statistics are also reported in Table 2.3.

We observe that the CEC may outperform the RCEC in some instances (e.g., the ratio can be as low as 30.58%); however, for the majority of the examples, the RCEC performs better than the CEC (see, e.g., the 25^{th} percentile column), although the differences are insignificant. More importantly, the reduction of computation time from CEC to RCEC is substantial: all else being equal, the average time for solving a test instance using the RCEC is only 9.64% of that using the CEC.

	Mean	Std.	Min.	25%-percentile	Median	Max.
γ	1.00118	0.02718	0.30584	1.00001	1.00008	3.14509
γ_{time}	0.09636	0.05614	0.00416	0.06084	0.08606	1.12851

Table 2.3.: Comparison of RCEC and CEC.

Why does the RCEC exhibit a better overall performance? We offer the following plausible explanation. In both the CEC and RCEC heuristics, we replace the future random demands by their means in each period. Such an approximation clearly will change our original problem and result in suboptimal solutions. In the RCEC, the optimal protection level is determined by comparing two values: The first is the upgrading value from using the product in the current period; the second is the expected marginal value of the product if it is saved to the next period. For illustration, consider the upgrading of demand j using capacity i in period t. The latter value is defined as $\bar{\Theta}^{t+1}_{\text{RCEC}}(\mathbf{X}^{t+1} + \mathbf{e}_i, \tilde{\mathbf{D}}^{t+1} + \mathbf{e}_j | \mu^{t+1}, \cdots, \mu^T) - \bar{\Theta}^{t+1}_{\text{RCEC}}(\mathbf{X}^{t+1}, \tilde{\mathbf{D}}^{t+1} | \mu^{t+1}, \cdots, \mu^T)$, where $\mathbf{e}_s$ $(s = i, j)$ is the unit vector with 1 in position s. The mean demand approximation may introduce biases into the two revenue functions. However, since the expected marginal revenue is defined as the difference of the two revenue functions, these biases may be cancelled out to some degree. In other words, the inaccuracies introduced by certainty equivalence control might be reduced in the RCEC heuristic. Note that such a cancellation effect does not exist in the traditional CEC heuristic. Therefore, the RCEC generally outperforms the CEC. In addition, the RCEC is more attractive than the CEC in terms of computational time in our numerical study.

It is worth mentioning that one may also use the deflected linear decision rule (DLDR) method proposed in Chen et al. (2008) to approximate Θ^t in the PSR algorithm. Let Θ^t_{DLDR} be the revenue collected by using $\mathbf{Y}^t_{\text{DLDR}}$'s in the remaining sales horizon, and denote $\Pi_{\text{DLDR}}(\mathbf{X}) = \Theta^1_{\text{DLDR}}(\mathbf{X}, \mathbf{0})$ as the expected revenue under the DLDR heuristic. We evaluate $\Pi_{\text{DLDR}}(\mathbf{X})$ in the numerical study described above and find that $\Pi_{\text{DLDR}}(\mathbf{X})$ and $\Pi_{\text{RCEC}}(\mathbf{X})$ are almost identical in all the problem instances.

In summary, based on the results in Tables 2.1 and 2.2, we conclude that the RCEC is able to deliver close-to-optimal revenues for the firm in a wide range of problem situations. In addition, the RCEC greatly reduces the computational complexity of the original problem. Therefore, in the rest of the chapter, we will use the RCEC to solve the dynamic capacity management problem.

2.8.2 Value of Optimal Upgrading

Given the efficiency and effectiveness of the RCEC heuristic, we are ready to derive more insights into the problem using numerical studies. There are a couple of natural questions we would like to address. First, what is the value of using multi-step upgrading? Second, what is the value of using the optimal capacity? Both questions are important from a practical standpoint because managers need to know how complex an upgrading structure should be used and how to determine the initial capacity. This subsection focuses on the first question and the second will be addressed in the next subsection.

Let $\Pi^k_{\mathrm{RCEC}}(\mathbf{X})$ be the revenue function given initial capacity $\mathbf{X}$ and k-level upgrading (i.e., product i can be used to satisfy class j demand only if $i \le j \le i+k$). Note that when $k = 0$, no upgrading is allowed, and $\Pi^0_{\mathrm{RCEC}}(\mathbf{X}) = \Pi_{\mathrm{NV}}(\mathbf{X})$, where $\Pi_{\mathrm{NV}}(\mathbf{X})$ is the optimal revenue in the newsvendor model in (2.17). Define

$$\Delta^k_{\mathrm{RCEC}} = \frac{\Pi^k_{\mathrm{RCEC}}(\mathbf{X}) - \Pi^{k-1}_{\mathrm{RCEC}}(\mathbf{X})}{\Pi^{k-1}_{\mathrm{RCEC}}(\mathbf{X})} * 100\%, \qquad k = 1, 2, 3, 4,$$

which measures the percentage profit gain from one additional level of upgrading under the RCEC.

We evaluate the values of Δ^k_{RCEC} using the same parameters as those for Table 2.2 except the initial capacities. Intuitively, upgrade is more valuable when the capacity is unbalanced, i.e., there is excess capacity for some products while there is shortage for the others. Such unbalance may occur even if the initial capacities are optimally set, because demand may fluctuate due to seasonality and trend while capacities are determined for the long term. Thus, when choosing the initial capacity we use the following procedure. Start with the optimal capacity under the RCEC, i.e., $\mathbf{X}_{\mathrm{RCEC}}$; then set the capacity for one product (say, product j) to 0 while adding capacity $(\mathbf{X}_{\mathrm{RCEC}})_j$ to a higher-quality product; finally, scale the entire capacity vector by different multipliers. Mathematically, for $1 \le i < j \le 5$, we consider all initial capacity $\mathbf{X}$, whose components are given by

$$(\mathbf{X})_i = \lambda((\mathbf{X}_{\mathrm{RCEC}})_i + (\mathbf{X}_{\mathrm{RCEC}})_j),\ (\mathbf{X})_j = 0,\ (\mathbf{X})_s = \lambda(\mathbf{X}_{\mathrm{RCEC}})_s, \forall\, s \in \{1, \cdots, 5\} \backslash \{i, j\},$$

where $\lambda \in \{0.9,\ 1,\ 1.1\}$. There are 10 combinations of the initial capacities for each λ and parameter set; one example is

$$\mathbf{X} = ((X_{\mathrm{RCEC}})_1 + (X_{\mathrm{RCEC}})_2, 0, (X_{\mathrm{RCEC}})_3, (X_{\mathrm{RCEC}})_4, (X_{\mathrm{RCEC}})_5).$$

A full list of the initial capacities are given in the appendix. We believe such a design captures the possible capacity scenarios that may happen over time as the firm allocates products to satisfy realized demand, especially those with unbalanced capacities. Moreover, the mean of total demand over the selling horizon remains the same for different $T \in \{3, 15, 30\}$, which implies that less demand information is available within each period for larger T. The numerical results are given in Table 2.4.

T	Upgrading Level k	Mean	Median	90%-percentile
3	1	29.75	20.64	51.63
	2	5.71	2.09	15.21
	3	1.45	0.11	4.99
	4	0.25	0.01	0.28
15	1	25.96	20.25	47.44
	2	4.86	2.99	12.88
	3	0.79	0.04	2.70
	4	0.07	0	0.09
30	1	20.38	19.88	45.63
	2	3.89	1.63	2.70
	3	0.67	0.02	2.02
	4	0.05	0	0.07

Table 2.4.: The value of using multi-step upgrading (Δ^k_{RCEC}).

There are several observations from Table 2.4. First, we can see that the value of multi-step upgrading can be highly valuable. For instance, with $T = 3$, the benefit of moving from one-step upgrading to two-step upgrading can be as high as 15.21% at the 90^{th} percentile (i.e., for at least 10% of the scenarios, the value is more than 15.21%). The number becomes 4.99% if we move from two-step upgrading to three-step upgrading. This result implies that single-step upgrading may not capture the full benefit of upgrading and multi-step upgrading is critically needed in many cases.

In particular, Table 2.4 suggests that the firm's profit increases in the upgrading level k and the marginal value decreases in k, both of which are quite intuitive.

Second, Table 2.4 indicates that the value of multi-step upgrading decreases in T. That is, using more upgrading levels will be less beneficial when there are more time periods in the selling horizon. Close scrutiny reveals that there is a key contributing factor to this interesting observation. A large T value means there are more time periods, which allows "chain allocation" to be more likely to happen. To see this, first consider $T = 1$. In this case, under single-step upgrading, product 1 cannot be used to satisfy demand 3. However, with $T = 2$, it is possible that product 2 is used to satisfy demand 3 in period 1; and then in the second period, product 1 is used to satisfy demand 2. These two allocations essentially mean that product 1 is used to satisfy demand 3. The chain allocation is analogous to multi-step upgrading; the only difference is that it can be better executed when there are more time periods. Therefore, multi-step upgrading is less valuable since it can be implemented even under single-step upgrading, but in a different way.

Finally, the numerical experiments suggest that the multi-step upgrading is most valuable when the initial capacity is unbalanced. For example, for $T = 3$, when the optimal initial capacity $\mathbf{X}_{\text{RCEC}}$ is used, the incremental value of moving from 2-level to 3-level upgrading is 0.04% on average; however, for initial capacity $\mathbf{X} = ((X_{\text{RCEC}})_1, (X_{\text{RCEC}})_2 + (X_{\text{RCEC}})_5, (X_{\text{RCEC}})_3, (X_{\text{RCEC}})_4, 0)$, the counterpart value is 5.10%. This indicates that the multi-step upgrading is quite important because unbalanced capacity may arise over time, even if the problem starts with the optimal initial capacity.

What is the benefit of using more upgrading levels if the optimal initial capacities are used? To answer this question, let $\mathbf{X}_{\text{RCEC}}(k)$ $(k = 0, 1, \cdots, 4)$ be the optimal initial capacities obtained from the RCEC heuristic with k-level upgrading, and redefine

$$\Delta^k_{\text{RCEC}} = \frac{\Pi^k_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}(k)) - \Pi^{k-1}_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}(k-1))}{\Pi^{k-1}_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}(k-1))} * 100\%, \qquad k = 1, 2, 3, 4$$

which is the percentage profit gain from one additional level of upgrading under the RCEC if the corresponding optimal initial capacities are used. Using the same set of parameters as in Table 2.4, we obtain the numerical results given in Table 2.5.

Upgrading Level k	Mean	Median	90%-percentile
1	2.80	2.67	4.37
2	0.92	0.81	1.64
3	0.55	0.49	1.35
4	0.50	0.35	1.03

Table 2.5.: The value of using multi-step upgrading (Δ^k_{RCEC}) under optimal initial capacity.

As one may expect, the values of using multi-step upgrading are much smaller in Table 2.5 because the initial capacities have been accordingly adjusted, and this lowers the benefit of using more levels of upgrading. However, the value of multi-step upgrading should not be overlooked either: the profit gain by moving from one-step to two-step upgrading is 0.92% on average and 1.64% at the 90^{th} percentile[6].

2.8.3 Capacity Decision vs. Allocation Mechanism

The profit of the upgrading problem hinges upon both the initial capacity and the allocation mechanism. This raises an interesting question: which decision is more important, capacity sizing or allocation mechanism? This is a practical question because the firm may wish to focus limited resources on improving the decision that has a bigger impact on profit. To shed some light on this question, we measure the

[6]In our numerical study, upgrade constitutes 2.78% of the total satisfied demands on average when the optimal initial capacity is used, and 29.47% when the suboptimal initial capacities are adopted. If the firm uses frequent upgrading to satisfy customer demand (e.g., the initial capacity is poorly decided), customers may learn about the upgrading pattern and become opportunistic. That is, a class i customer may intentionally ask for product j ($i < j$), hoping that she will be upgraded when product j is out of stock. Incorporating such a behavior is out of the scope of this chapter and therefore left for future research.

importance of each decision using the profit loss when a suboptimal decision is applied rather than the optimal one. Next, we describe the suboptimal decisions that will be used.

In our problem, it is time-consuming to derive the optimal initial capacity even if we can efficiently solve the optimal allocation decision by the RCEC heuristic. So we consider two simple alternatives. The first alternative is to use the optimal capacity $\mathbf{X}_{\mathrm{CB}}$ in the crystal ball model. The crystal ball model is called static model in SZ, who find that $\mathbf{X}_{\mathrm{CB}}$ yields nearly optimal revenue for the firm in their single-step upgrading model. To check whether the result carries over to our general upgrading model, define

$$\Delta_{\mathbf{X}_{\mathrm{CB}}} = \left|\frac{\Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{CB}}) - \Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{RCEC}})}{\Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{RCEC}})}\right| * 100\%$$

to measure the performance of the crystal-ball capacity $\mathbf{X}_{\mathrm{CB}}$. Since the true optimal capacity is unknown, we use $\mathbf{X}_{\mathrm{RCEC}}$ as the benchmark for the comparison. With the same parameters used for Tables 2.2, 2.3 and 2.4, we evaluate $\Delta_{\mathbf{X}_{\mathrm{CB}}}$ for 780 examples and summarize the results in Table 2.6 (the first row). It can be seen that $\Delta_{\mathbf{X}_{\mathrm{CB}}}$ is generally negligible in the numerical study: The average revenue difference is 0.017% and the maximum is 1.062%.

	Mean	Std.	Median	90%-percentile	Max.
$\Delta_{\mathbf{X}_{\mathrm{CB}}}$	0.01735	$5.62378 * 10^{-2}$	0	0.043287	1.06237
$\Delta_{\mathbf{X}_{\mathrm{NV}}}$	0.33278	$2.91287 * 10^{-1}$	0.27123	0.72231	1.62893
Δ_{greedy}	5.19543	5.69987	8.22994	12.28855	12.70996

Table 2.6.: Capacity decision vs. allocation mechanism.

An even simpler alternative is to use the newsvendor capacity $\mathbf{X}_{\mathrm{NV}}$, i.e., the optimal capacity under no upgrading. Similarly, in the same numerical study, we define

$$\Delta_{\mathbf{X}_{\mathrm{NV}}} = \left|\frac{\Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{NV}}) - \Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{RCEC}})}{\Pi_{\mathrm{RCEC}}(\mathbf{X}_{\mathrm{RCEC}})}\right| * 100\%$$

and present the statistics of $\Delta_{\mathbf{X}_{\text{NV}}}$ in Table 2.6 (the second row). We can see that $\Delta_{\mathbf{X}_{\text{NV}}}$ is greater than $\Delta_{\mathbf{X}_{\text{CB}}}$ in general, but it offers reasonably good performance as well. The average and maximum revenue differences are 0.333% and 1.629%, respectively. In particular, the number at the 90^{th} percentile is 0.722%, which means that the newsvendor capacity performs quite well for the majority of the scenarios. From the above observations, one can see that these simple alternatives to the optimal capacity perform reasonably well. Therefore, as long as the optimal upgrading policy is used, the value of using the optimal capacity seems to be very small in our problem setting.

Next, we consider the impact of using suboptimal allocation policy. We first use greedy upgrading as the suboptimal policy, which myopically upgrades all unmet demands by surplus capacities. It serves as a reasonable suboptimal policy because it is intuitive and straightforward to implement in practice. Furthermore, the RCEC heuristic incorporates greedy upgrading to simplify its computation. Specifically, let $\Pi_{\text{greedy}}(\mathbf{X})$ be the expected profit using greedy upgrading given initial capacity $\mathbf{X}$. We define

$$\Delta_{\text{greedy}} = \left| \frac{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}) - \Pi_{\text{greedy}}(\mathbf{X}_{\text{RCEC}})}{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})} \right| * 100\%$$

as the profit loss due to greedy upgrading. The same parameters for $\Delta_{\mathbf{X}_{\text{CB}}}$ and $\Delta_{\mathbf{X}_{\text{NV}}}$ have been used, and the statistics of Δ_{greedy} are presented in Table 2.6 (the third row). The average profit loss due to greedy upgrading is 5.195%, which is much larger than those for $\Delta_{\mathbf{X}_{\text{CB}}}$ and $\Delta_{\mathbf{X}_{\text{NV}}}$. In addition to greedy upgrading, we also test suboptimal allocation policies that involve only k-step ($k = 0, \cdots, N-2$) upgrading. The magnitudes of profit losses are still generally much larger than those for $\Delta_{\mathbf{X}_{\text{CB}}}$ and $\Delta_{\mathbf{X}_{\text{NV}}}$. To save space, the detailed results are presented in the appendix.

The above numerical results suggest that the benefit of choosing an effective allocation mechanism outweighs that of choosing an accurate initial capacity. Based on these observations, in practice, the firm may decide the initial capacity by using simple approximations (e.g., either the NV or CB model) and focus on optimally allocating the capacity during the sales horizon.

2.9 Conclusion

This chapter studies a firm's capacity investment and allocation problem in a dynamic setting with stochastic demand. There are multiple demand classes, which can be satisfied by multiple classes of capacities. Demand arrives in discrete time periods, and the firm needs to make capacity allocation decisions in each period before observing future demand. A general upgrading structure is considered, which is broad enough to cover a wide range of practical upgrading situations. One may also view this as an inventory management problem with one-way dynamic substitution.

We first show that for any given initial capacity, a Parallel and Sequential Rationing (PSR) policy is optimal for the firm. Under the PSR policy, the firm can make upgrading decisions in each period sequentially rather than simultaneously, which greatly reduces the complexity of the capacity allocation problem. Despite the well-structured PSR policy, the dynamic allocation problem is still subject to the curse of dimensionality. Thus we propose a Refined Certainty Equivalence Control (RCEC) heuristic that improves over the traditional CEC methodology by exploiting the property of the PSR policy. Through extensive numerical experiments, we find that the RCEC heuristic is highly efficient and yields nearly optimal revenue for the firm. With the help of the RCEC heuristic, we conduct numerical studies to derive managerial insights about the dynamic capacity management problem. Our numerical studies indicate that the multi-step upgrading could be significantly valuable, especially when the capacities are not balanced (either due to suboptimal initial investment or unexpected demand realizations over time). We find that using simple approximations (e.g., the NV and CB models) for the initial capacities leads to negligible profit loss, while the negative impact of using a suboptimal allocation (e.g., greedy upgrading) could be quite significant. In this sense, the allocation mechanism plays a more important role in our problem than the capacity sizing decision.

Chapter 3

Upgrading, Product Differentiation, and Heterogeneous Consumers

3.1 Introduction

In order to achieve better market segmentation and extract higher profit margins, firms often offer more than a single product (or service) to consumers. Examples include business class and economy class tickets in the airline industry, full-size and compact-size cars provided by car rental companies, and deluxe and standard rooms in the hotel industry. In many cases, consumers can choose between a regular product and a premium product. While both products provide the basic function desired by consumers, a premium product is bundled with additional services and features that can yield a higher utility for a consumer. For instance, the basic function of an air ticket is to fly a passenger to her destination; however, a business class ticket offers more comfortness (e.g., bigger seat with more leg room) and superior services (e.g., free checked luggage, more food and drink choices, and early boarding and unboarding). In contrast, a regular product may only consist of a fraction of features and services included in the premium product.

Product differentiation is clearly an important decision when firms offer multiple products in the market. Normally the premium product should entail most, if not all possible services and features; however, the design of the regular product requires careful consideration. By adjusting the add-on services/features attached to the regular product, firms can influence the utility a consumer derives from the product. In the airline industry, companies adopt different baggage policies. Southwest Airlines allow domestic passengers to check two pieces of luggage for free, while many other airlines charge a fee for checked luggage[1]. Delta Airlines used to offer two free checked luggage for the economy class passengers on international flights between north America and Asia, then they began to charge a fee for the second checked luggage, but they changed back to the old policy recently[2]. In the hotel industry, some companies provide complimentary add-on services such as parking, wireless Internet, and access to fitness center (Shulman and Geng 2012). A particularly interesting add-on service is upgrading in the aforementioned industries. Sometimes firms upgrade a regular-class customer to a premium product without charging additional fees. In particular, firms can control the probability of offering product upgrade by imposing additional restrictions[3].

Increasing consumer diversification stimulates firms' awareness of consumer heterogeneity, which complicates firms' differentiation decisions. There are several contributing factors to such heterogeneity. First, there is discrepancy in consumers' acknowledgement of the complete characteristics of the product. For instance, consumers with limited time to research the product may be unaware of the add-on services (e.g., free parking and Internet) provided by a hotel when making their reservations (see Shulman and Geng 2012). Second, even with full knowledge of the product, consumers may have different preferences or purchase intentions over the same product. In the airline industry, passengers with or without checked bags may have different valuations of the same economy class ticket with free checked baggage

[1]Baggage Policies: `http://www.southwest.com/html/customer-service/baggage/`.

[2]Delta Airlines' previous bag fees: `http://www.delta.com/content/www/en_US/traveling-with-us/baggage/before-your-trip/checked/previous-bag-fees.html`. Current baggage fee policy: `http://www.delta.com/content/www/en_US/traveling-with-us/baggage/before-your-trip/checked.html`.

[3]Delta Airlines' Medallion Upgrades: `http://www.delta.com/content/www/en_US/traveling-with-us/check-in/requesting-medallion-upgrades.html`. American Airlines' 500-Mile Upgrades (Complimentary): `http://www.aa.com/i18n/urls/aadvantageupgrades.jsp?anchorLocation=DirectURL&title=aadvantageupgrades#500MileUpgradesComp`. United Airlines' Complimentary Premier Upgrades: `https://www.united.com/web/en-US/content/mileageplus/premier/upgrades.aspx`.

service. Lastly, in the upgrading case, consumers may have disparate capabilities in evaluating the value from potential upgrading. From these observations, a consumer needs to spend time and effort to accurately assess the value from the add-on services. In reality, it is natural that some consumers are naive (i.e., they simply ignore the value from add-on services), while some consumers are strategic (i.e., they are fully aware of the value from add-on services).

This chapter studies a firm's optimal strategies when selling two differentiated products to heterogeneous consumers. We focus on several research questions. First, how much add-on service the firm should attach to the regular product? As mentioned above, they may adjust the quality of the regular product by changing the add-on services included in the product. In the presence of strategic consumers, firms may also change the product quality by varying the frequency of offering complementary product upgrade (i.e., the firm can adjust the probability of providing a higher quality product to consumers who only pay the price of the low quality product). So how does increasing the add-on service to the regular product affect the firms' profitability?

Second, what is the impact of consumer behavior on the firm's optimal strategy? Consumer heterogeneity clearly affects the firm's quality and pricing decisions. In addition, the firm may use various tools to influence consumers' behavior. For example, advertisement is a useful tool to improve consumers' knowledge of the characteristics of the product. Moreover, the firm may intentionally train the consumers to be more strategic. There are many websites that provide advice to consumers on how to improve the chance of getting upgrades[4]. On one hand, these methods can help naive consumers better understand the regular product and therefore improve their valuations of the product - this allows the firm to charge a higher price on the regular product. On the other hand, a more attractive regular product may induce a strategic consumer to switch from the premium product to the regular one, hoping to receive a free upgrade - such opportunistic behavior may cannibalize the sales from

[4]How to get a free upgrade: `http://www.cnn.com/2012/03/19/travel/free-upgrade-strategies/`. Top 10 Ways to Get Upgraded on a Flight: `http://airtravel.about.com/od/travelindustrynews/a/upgrades.htm`.

the premium product. It is natural to explore the impact of consumer heterogeneity on firms' optimal strategy and profit. In particular, it is useful to know under what circumstances the firm can benefit from efforts influencing consumer behavior.

Third, how does capacity constraint affect the firm's optimal strategy? In most revenue management settings (e.g., airline, hotel, and car rental settings), the firm is often constrained by limited capacity, especially for the premium product. Thus it would be interesting to study the role the capacity constraint plays in the firms' optimal strategy. Specifically, we will investigate how capacity constraint may change the insights from the previous two questions.

We propose an analytical framework to address the above questions. In particular, we study a single-period model where a monopoly firm sells two products with distinct qualities, i.e., the regular product and the premium product. The premium product has a limited capacity, whereas the regular product has ample capacity. Consumers arrive to the market, each purchasing at most one unit of either product. All consumers agree with the valuation of the premium product but they differ on the valuation of the regular product: opportunistic consumers value the regular product more than the naive consumers because they are better informed and are more capable to evaluate the add-on services. Each consumer makes the purchase decision based on her own individual preference for quality, valuations of these products and the product prices. The firm's objective is to maximize its profit by setting the prices for both products.

We obtain three major results about firms' optimal strategies under such a model setting. First, the quality improvement of the regular product can change the firm's profit in both directions. The intuition is as follows. When increasing the quality of the regular product, the firm always decreases its price of the premium product while increasing the price of the regular product to limit the number of consumers who appreciate such quality improvement and purchase the regular product instead of the premium one. However, it is possible that the profit loss from the premium product dominates the profit gain from the regular product. We find that it happens when the quality difference between these two products are moderate, which is quite common in practice. Our results suggest that firms should conduct product improvement or upgrade with extra caution.

Second, improving consumers' knowledge about the regular product through advertisement may actually hurt the firm when the consumer awareness is low and the capacity of the premium product is large. The benefit of the advertisement is clear: since more consumers are aware the true quality of the product, there are more demands for the regular product and the firm may increase the corresponding price. But in this case, as more advertisement is used, firm has to decrease the price of the premium product to avoid undesirable capacity leftover. And the profit loss from the premium product always dominates the additional profit from the regular product.

Finally, the capacity constraint plays an important role in the firm's optimal quality improvement and advertising decisions. Specifically, quality improvement of the regular product and advertising such improvement will be more beneficial when there is a low capacity for the premium product. Recall that the main negative effect from quality improvement and advertising is cannibalization, i.e., the firm has to decrease the price of the premium product to prevent opportunistic consumers from switching to regular product. A more constrained capacity for the premium product, which is equal to the number of consumers purchasing the premium product in the optimal solutions, means the cannibalization effect will be weaker. Thus, the firm will have stronger incentives to offer more add-on services and publicize such services when the premium capacity is highly constrained.

The rest of this chapter is organized as follows. Section 3.2 reviews the related literature. Section 3.3 describes the model setting. Section 3.4 characterized the impact of the quality improvement and Section 3.5 investigates how the firm should affect the consumer heterogeneity. Section 3.6 presents the findings from numerical studies. The chapter concludes in Section 3.7. All proofs are given in the Appendices.

3.2 Literature Review

This chapter uniquely incorporates the consumer heterogeneity and the capacity constraint. There are two aspects of the consumer heterogeneity: First, consumers may have different valuations about the quality of the same product; Second, consumers may have different willingness to pay for the same quality.

The first type of the consumer heterogeneity can be explained by either the bounded rationality or the strategic behavior. As we mentioned earlier, consumers may be unable to account for all characteristics that are included in the product or service and thus make uninformative decisions. The concept of bounded rationality is introduced by Simon (1955). Conlisk (1996) review extensive evidences for incorporating bounded rationality in economic models and provided supports both theoretically and empirically. In our model, the consumer heterogeneity can also be the result of the strategic behavior. Specifically, when making a purchase decision, consumers choose from various alternatives, including the options of avoiding the purchase. For instance, whether or not an air passenger chooses an economy class ticket at a specific time depends on her expectation of the probability being upgraded to the business class, which has higher value to her. Such forward looking behavior has been widely studied in the consumer behavior literature. For instance, Jacobson and Obermiller (1990), Krishna et al. (1991), Ho et al. (1998) and Su and Zhang (2008). Although our model does not explicitly incorporate the strategic behavior, the essence is the same: there exists a difference in valuation of the same product or service among consumers. It is clear that such difference can be the result of whether or not the consumer is strategic. And we consider the impact of the difference on the firm's pricing decision and the optimal profit as the strategic consumer literature. However, in contrast to the above literatures, this chapter considers more than one product that introduces product cannibalization with the firm.

The second type of the consumer heterogeneity is the consumer differentiation. This topic has been extensively studied in the literature of product line design that includes both quality decisions of a quality-differentiated spectrum of goods and the corresponding pricing decisions. Mussa and Rosen (1978) consider a monopoly firm deciding the optimal set of price-quality schedules offered to consumers with heterogeneous tastes of the qualities, which is modeled by a continuous distribution. Similarly, Moorthy (1984) analyzes a monopolist serving discrete consumer segments of quality valuations. By extending the model into duopoly setting, Moorthy (1988) shows the equilibrium price-quality schedules for each firm facing the marginal production cost that is quadratic with respect to the quality. As extensions, Chambers et al. (2006) discuss the impact of variable production cost when duopoly firms decide quality and price sequentially, Lehmann-Grube (1997) studies the impact of the convex fixed cost of quality chosen by the firms in the first stage, Rhee (1996) considers the impact of the consumer's heterogeneity that is unobservable to the firms and is modeled as a random variable in the consumer's utility function, and Ronnen (1991) investigates the consequence of imposing a minimum quality standard in the duopoly case. When income disparities of the consumers have been privileged against taste differentiation, Gabszewicz and Thisse (1979) characterize the equilibrium prices of the duopoly firms selling two differentiated products whose qualities are predetermined. Gabszewicz and Thisse (1980) further extend the model to study the equilibrium prices in a competitive market with arbitrary number of firms. After incorporating the firms' decisions of entering the market, Shaked and Sutton (1982) and Gabszwicz and Thisse (1986) study the equilibrium when the entry decisions are made simultaneously, while Donnenfeld and Weber (1992) explore the equilibrium qualities and prices when firms sequentially make the entry decisions. Although these studies involve both the optimal quality and price decisions, consumers are assumed to have the save valuation of the qualities, and the capacity constraint is not considered.

This chapter also exhibits some similarities to the add-on pricing literature in the sense that we consider how profit is changed when the firm improves the quality of the regular product, and providing add-on service is clearly one of many possible quality improvements. Ellison (2005) identified how add-on pricing can actually lead to improved profit for the firm by creating an adverse selection problem in the competition. Gabaix and Laibson (2006) were among the first in the add-on literature to allow for boundedly rational consumers. Shulman and Geng (2012) examine the consequences of add-on pricing when firms are both horizontally and vertically differentiated. This chapter differs from the add-on literature because the firm in our model do not charge an additional fee for the improved or "add-on" quality, thus, the firm is unable to achieve price discrimination. In addition, our model innovatively considers the capacity constraint, which has great impact on the firms' optimal pricing, quality improvement and advertisement decisions.

3.3 Model

Consider a firm selling two differentiated products, product 1 (the premium product) and product 2 (the regular product). There is a continuum of consumers in the market with a deterministic total size 1. Each consumer needs only one unit of the product. Consumer utility is given by $U = \theta q - p$, where q is the product quality, p is the product price, and $\theta \geq 0$ is the parameter which measures the intensity of consumer's taste for quality. We assume θ is uniformly distributed on $[0, 1]$. The consumers unanimously agree on the quality of product 1, denoted by q_1 ($q_1 > 0$); however, they have different perceptions about the quality of product 2. Particularly, the naive consumers value product 2 at q_2 ($0 < q_2 < q_1$) while the opportunistic consumers value it at $q_2 + \delta$ ($0 < \delta < q_1 - q_2$). The parameter δ represent the additional value from add-on services (e.g., upgrading) discussed in the introduction. The naive consumers have a lower value for product 2 because they are either uninformed about the add-on services or they are unable to evaluate the value from the add-on services (e.g., it is hard to evaluate the probability of receiving an upgrade). We assume there is r ($0 \leq r \leq 1$) portion of opportunistic consumers in the market (and $1 - r$ of naive consumers). Thus r measures how strategic the consumer population is in the market. Although the firm understands there are two types of consumers, it cannot distinguish each individual consumer, and is therefore unable to charge different prices for product 2. The firm faces exogenously given capacity constraints (e.g., a flight has limited number of seats, a hotel has certain number of guest rooms). Let x_i ($x_2 > x_1 > 0$) be the capacity for product i. For analytical tractability, in the base model we assume x_2 is large enough such that it never constrains the sales. The fact that the premium capacity is more constrained than the regular capacity is consistent with most practical situations (e.g., airlines and hotels). Later we will demonstrate that the qualitative insights will remain after relaxing the assumption about x_2. The marginal usage costs of both products are normalized to zero. The firm's objective is to maximize its revenue $\pi(p_1, p_2)$ by choosing prices p_i ($i = 1, 2$) for the products. Since the capacities x_i are exogenously given, we will use profit and revenue functions exchangeably. For notational convenience, define $y \vee z = \max(y, z)$, $y \wedge z = \min(y, z)$, and $(z)^+ = \max(z, 0)$.

To analyze the firm's problem, we start with consumer demand functions. Define $d_i(p_1, p_2)$ ($i = 1, 2$) as the total demand for product i under prices (p_1, p_2). Consider an opportunistic consumer with a taste parameter θ, he chooses between products 1 and 2 as well as choosing whether to purchase at all. The utilities of buying product 1 and 2 are $\theta q_1 - p_1$ and $\theta(q_2 + \delta) - p_2$, respectively. The consumer purchases product 1 if $\theta q_1 - p_1 \geq \theta(q_2 + \delta) - p_2$ and $\theta q_1 - p_1 \geq 0$ or product 2 if $\theta(q_2 + \delta) - p_2 \geq \theta q_1 - p_1$ and $\theta(q_2 + \delta) - p_2 \geq 0$. Therefore, given the firm's prices (p_1, p_2), an opportunistic consumer purchases product 1 if his taste parameter $\theta \in [\frac{p_1 - p_2}{q_1 - q_2 - \delta} \vee \frac{p_1}{q_1}, 1]$ or product 2 if $\theta \in [\frac{p_2}{q_2 + \delta}, \frac{p_1 - p_2}{q_1 - q_2 - \delta} \vee \frac{p_1}{q_1}]$. The same argument can be applied to the naive consumers by replacing $q_2 + \delta$ with q_2. Thus, the demand for product 1 is

$$d_1(p_1, p_2) = r \left(1 - \frac{p_1 - p_2}{q_1 - q_2 - \delta} \vee \frac{p_1}{q_1}\right)^+ + (1 - r) \left(1 - \frac{p_1 - p_2}{q_1 - q_2} \vee \frac{p_1}{q_1}\right)^+. \tag{3.1}$$

Similarly, we can derive the demand for product 2 to be

$$\begin{aligned} d_2(p_1, p_2) = r & \left(1 \wedge \left(\frac{p_1 - p_2}{q_1 - q_2 - \delta} \vee \frac{p_1}{q_1}\right) - \frac{p_2}{q_2 + \delta}\right)^+ \\ & + (1 - r) \left(1 \wedge \left(\frac{p_1 - p_2}{q_1 - q_2} \vee \frac{p_1}{q_1}\right) - \frac{p_2}{q_2}\right)^+. \end{aligned} \tag{3.2}$$

Let $\mathbf{R} = \{(p_1, p_2): \; 0 \le p_1 \le q_1, \; 0 \le p_2 \le q_2 + \delta\}$ be the feasible region for the firm, then we only need to focus on $\mathbf{R}$ when searching for the firm's optimal prices. To see this, suppose $(p_1, p_2) \notin \mathbf{R}$ constitute the firm's optimal prices. Without loss of generality, we assume $p_1 > q_1$ and $0 \le p_2 \le q_2 + \delta$, then $\pi(q_1, p_2) = \pi(p_1, p_2)$ since $d_1(p_1, p_2) = d_1(q_1, p_2) = 0$ and $d_2(p_1, p_2) = d_2(q_1, p_2)$ by (3.1) and (3.2). The similar argument can be applied to $p_2 > q_2 + \delta$. Hence, we will restrict to the region $\mathbf{R}$ in subsequent analysis.

Now the firm's optimization problem can be written as follows:

$$\begin{aligned} \max \quad & \pi(p_1, p_2) = p_1 d_1(p_1, p_2) + p_2 d_2(p_1, p_2) \\ \text{s.t.} \quad & d_1(p_1, p_2) \le x_1, \;\; (p_1, p_2) \in \mathbf{R}. \end{aligned} \tag{3.3}$$

Note x_1, δ, and r are the key parameters in the firm's optimization problem. Later we will investigate how these parameters affect the firm's optimal revenue.

3.3.1 Analysis of Objective Function

As preparation, we first study the property of the firm's objective function in (3.3). Close scrutiny of the demand functions $d_i(p_1, p_2)$ $(i = 1, 2)$ reveals that the region $\mathbf{R}$ in (3.3) can be partitioned into sub-regions based on consumers' purchase decisions. In particular, $\mathbf{R}$ can be divided into 3 regions, $\mathbf{R}_i$ $(i = 1, 2, 3)$, where $\mathbf{R}_1$ is the region in which there is demand for product 1 from both types of consumers, $\mathbf{R}_2$ is the region in which the demand for product 1 is only from the naive consumers, and $\mathbf{R}_3$ is the region in which no consumer purchases product 1. Consumers' purchase decisions regarding product 2 further split the $\mathbf{R}_i$'s into sub-regions. For example, $\mathbf{R}_{11}$ is the sub-region in which there is demand for product 2 from both types of consumers. Table 3.1 summarizes the consumers' purchase decisions in these different regions, which are also illustrated by Figure 3.1. A full characterization of the sub-regions is lengthy and therefore given in the appendix.

From (3.3) and Table 3.1, it is clear that $\pi(p_1, p_2)$ has different expressions in different regions. The following lemma summarizes the properties of $\pi(p_1, p_2)$.

Consumer Type		Opportunistic		Naive	
Product		1	2	1	2
$\mathbf{R}_1$	$\mathbf{R}_{11}$	✓	✓	✓	✓
	$\mathbf{R}_{12}$	✓	✓	✓	
	$\mathbf{R}_{13}$	✓		✓	
$\mathbf{R}_2$	$\mathbf{R}_{21}$		✓	✓	✓
	$\mathbf{R}_{22}$		✓	✓	
$\mathbf{R}_3$	$\mathbf{R}_3$		✓		✓

Table 3.1.: Consumer Purchase Decisions (✓ means positive demand).

Lemma 3.3.1 *The objective function $\pi(p_1, p_2)$ is continuous in $\mathbf{R}$. Moreover, $\pi(p_1, p_2)$ is continuously differentiable and jointly concave in (p_1, p_2) in $\mathbf{R}_i$ $(i = 1, 2, 3)$, respectively.*

Now we consider the firm's optimal pricing decisions. From Lemma 3.3.1, the optimal solution of (3.3) is straightforward if the capacity constraint $d_1(p_1, p_2) \leq x_1$ does not exist. Let $d_i^* = d_i(p_1^*, p_2^*)$ $(i = 1, 2)$ be the demand of product i when the optimal prices are used, and $\pi^* = \pi(p_1^*, p_2^*)$ the firm's optimal profit. Actually, it can be readily shown that in the absence of the capacity constraint, the firm's optimal solution is given by

$$(p_1^*, p_2^*) = (\frac{q_1}{2}, \frac{q_2 + \delta}{2}), \quad d_1^* = \frac{1}{2}, \quad d_2^* = 0, \quad \text{and} \quad \pi = \frac{q_1}{4}.$$

That is, the firm only sells the premium product, i.e., there is $d_2^* = 0$ in the optimal solution. Furthermore, there must be $d_1^* = \frac{1}{2}$, i.e., only half of the consumers with $\theta \geq \frac{1}{2}$ will be served. This is consistent with the existing results in the product line design literature (see Mussa and Rosen 1978). In the literature, it has been assumed that all consumers have the same valuation for the regular product. So we can show such a result still holds even when the consumers have different valuations for product 2. The intuition is that the firm wants to eliminate product competition between its own products when there is no usage cost. Thus, the firm's prices make the premium product more attractive than the regular product (or essentially removing the regular product from the market). From this observation, we will focus on the case $x_1 \leq \frac{1}{2}$ in the rest of chapter.[5]

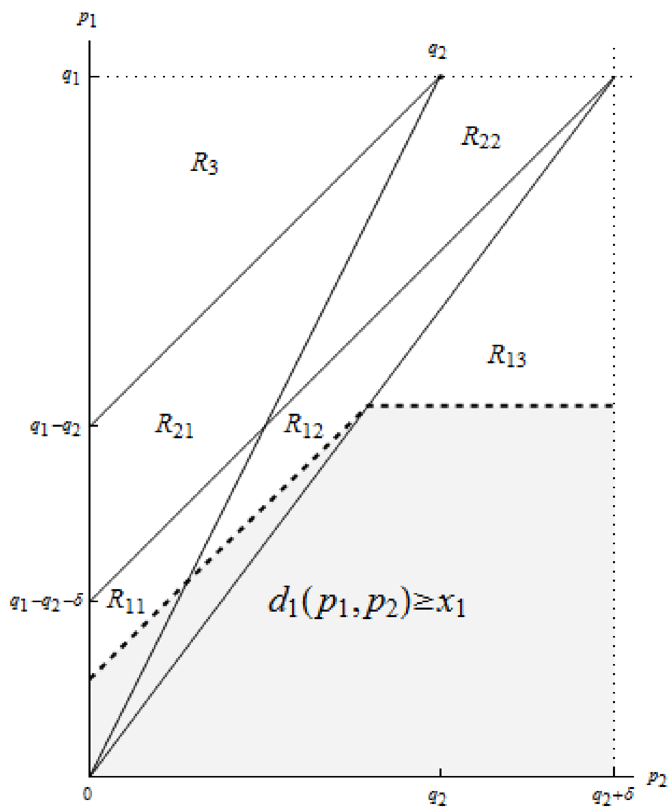


Figure 3.1.: Feasible Regions.

The analysis of the optimization problem in (3.3) becomes more involved when the capacity constraint is present, which can be highlighted in Figure 3.1. Note that the dashed lines in Figure 3.1 are the curves of prices (p_1, p_2) under which the capacity x_1 is fully utilized, i.e. $d_1(p_1, p_2) = x_1$. Intuitively, the premium product capacity is precious to the firm, which implies that the firm should always utilize such capacity to the fullest extent. Recall that only the premium product will be purchased when

[5]We have assumed the capacities are exogenously given. In reality, the capacity constraint would be determined by the cost of the capacity. For example, a capacity would be more constrained as its cost increases.

the capacity is sufficiently large (i.e., $x_1 \geq \frac{1}{2}$). As x_1 decreases from $\frac{1}{2}$, the firm needs to divert part of the demand of the premium product to the regular product due to the capacity limit. Since the opportunistic consumers have higher valuation of the regular product, the firm would charge a price to first divert some opportunistic consumers would change their decision and purchase the regular product. As x_1 further decreases, at certain point, even though all the opportunistic consumers have been diverted to the regular product, the limited premium capacity can not satisfy all demand from the naive consumers. Thus, the firm has to change the prices so that some of the naive consumers will also be diverted to the regular product. Define parameters k^j $(j = 1, ..., 6)$ that are independent of x_1 as follows:

$$
\begin{aligned}
k^1 &= \frac{(1-r)\delta}{q_1 - q_2}, \\
k^2 &= \frac{1}{2}\left(1 - \frac{\delta}{q_1 - q_2 - \delta}\sqrt{\frac{rq_1(q_1 - (q_2+\delta)(1-r))}{q_2(q_2 + (1-r)\delta)}}\right), \\
k^3 &= \frac{(1-r)(q_2 + (1-2r)\delta)}{2(q_2 + (1-r)\delta)}, \\
k^4 &= \frac{(q_2+\delta)(1-r)}{2q_1}, \\
k^5 &= \frac{(q_2-\delta)(1-r)}{2q_2}, \\
k^6 &= \frac{1}{4}(q_2+\delta)\left(\frac{q_2(1-r)}{q_1}\left(\frac{q_2+\delta}{q_1} - 2\right) + \frac{q_2}{q_2 + (1-r)\delta} - r\right).
\end{aligned}
\tag{3.4}
$$

The firm's optimal solution can be achieved in one of the four possible regions: $\mathbf{R}_{11}$, $\mathbf{R}_{12}$, $\mathbf{R}_{21}$, and $\mathbf{R}_{22}$. The corresponding prices, demands and profits are given as follows:

Region $\mathbf{R}_{11}$:

$$(p_1^*, p_2^*) = \left(\frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)} + \frac{(1-x_1)(q_1-q_2)(q_1-q_2-\delta)}{q_1-q_2-(1-r)\delta},\ \frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)}\right), \tag{3.5}$$

$$d_1^* = x_1,\ d_2^* = \frac{1}{2} - x_1,\ \pi^* = \frac{q_2(q_2+\delta)}{4(q_2+(1-r)\delta)} + \frac{x_1(1-x_1)(q_1-q_2)(q_1-q_2-\delta)}{q_1-q_2-(1-r)\delta};$$

Region $\mathbf{R}_{12}$:

$$(p_1^*, p_2^*) = \left(\frac{q_1((1-x_1)(q_1-q_2-\delta)+\frac{1}{2}r(q_2+\delta))}{q_1-(1-r)(q_2+\delta)},\ \frac{q_2+\delta}{2}\right), \tag{3.6}$$

$$d_1^* = x_1,\ d_2^* = \frac{rq_1(1-2x_1)}{2(q_1-(1-r)(q_2+\delta))},\ \pi^* = \frac{q_1(4x_1(1-x_1)(q_1-q_2-\delta)+r(q_2+\delta))}{4(q_1-(1-r)(q_2+\delta))};$$

Region $\mathbf{R}_{21}$:

$$(p_1^*, p_2^*) = \left(q_1 - \frac{(q_1-q_2)x_1}{1-r} - \frac{q_2(q_2+(1-2r)\delta)}{2(q_2+(1-r)\delta)},\ \frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)}\right), \tag{3.7}$$

$$d_1^* = x_1,\ d_2^* = \frac{1}{2} - x_1,\ \pi^* = (q_1-q_2)x_1\left(1-\frac{x_1}{1-r}\right) + \frac{q_2(q_2+\delta)}{4\left(q_2+(1-r)\delta\right)};$$

Region $\mathbf{R}_{22}$:

$$(p_1^*, p_2^*) = \left(q_1(1-\frac{x_1}{1-r}),\ \frac{q_2+\delta}{2}\right), \tag{3.8}$$

$$d_1^* = x_1,\ d_2^* = \frac{r}{2},\ \pi^* = \frac{r(q_2+\delta)}{4} + q_1x_1(1-\frac{x_1}{1-r}).$$

Proposition 3.3.1 characterizes the firm's optimal solutions.

Proposition 3.3.1 *Consider $x_1 \le \frac{1}{2}$. The firm's optimal solution is determined by x_1 and the thresholds k^j $(j = 1, \cdots, 6)$ defined in* (3.4). *Specifically,*

Case 1. If $k^1 < k^2$, then

$$(p_1^*, p_2^*) \in \begin{cases} \mathbf{R}_{21}, & \text{if } x_1 \in (0, k^1]; \\ \mathbf{R}_{11}, & \text{if } x_1 \in (k^1, k^2]; \\ \mathbf{R}_{12}, & \text{if } x_1 \in (k^2, \frac{1}{2}). \end{cases}$$

Case 2. If $k^1 \geq k^2$, $k^3 > k^4$ and $k^6 \geq 0$: there exists a threshold $\bar{k} \in [k^4, k^3]$, then

$$(p_1^*, p_2^*) \in \begin{cases} \mathbf{R}_{21}, & \textit{if } x_1 \in (0, \bar{k}]; \\ \mathbf{R}_{12}, & \textit{if } x_1 \in (\bar{k}, \frac{1}{2}). \end{cases}$$

Case 3. If $k^1 \geq k^2$, $k^3 > k^4$ and $k^6 < 0$: there exists a threshold $\bar{k} \in [0 \vee k^5, k^4]$, then

$$(p_1^*, p_2^*) \in \begin{cases} \mathbf{R}_{21}, & \textit{if } x_1 \in (0, \bar{k}]; \\ \mathbf{R}_{22}, & \textit{if } x_1 \in (\bar{k}, k^4]; \\ \mathbf{R}_{12}, & \textit{if } x_1 \in (k^4, \frac{1}{2}). \end{cases}$$

Case 4. If $k^1 \geq k^2$ and $k^3 \leq k^4$: there exists a threshold $\bar{k} \in [k^5, k^3]$, then

$$(p_1^*, p_2^*) \in \begin{cases} \mathbf{R}_{21}, & \textit{if } x_1 \in (0, \bar{k}]; \\ \mathbf{R}_{22}, & \textit{if } x_1 \in (\bar{k}, k^4]; \\ \mathbf{R}_{12}, & \textit{if } x_1 \in (k^4, \frac{1}{2}). \end{cases}$$

We may further explain the intuition behind Proposition 3.3.1. Recall that the opportunistic consumers have higher valuation of the regular product compared to the naive ones, which implies that there will be more opportunistic consumers buying the regular product than naive consumers under any price scheme. On the other hand, the premium product is less attractive to the opportunistic consumers, who are more sensitive to the price change of the premium product. When capacity is restrictive, the firm tends to increase the price of the premium product, which drives the opportunistic consumers to the regular product. Thus, the number of the opportunistic consumers buying the premium product decreases as the capacity becomes scarce. Table 3.2 illustrates the optimal solutions in each case described by Proposition 3.3.1. For instance, if $k^1 < k^2$, the optimal solutions are described in Case 1. As x_1 decreases from $\frac{1}{2}$ to 0, the optimal solution moves from region $\mathbf{R}_{12}$ to $\mathbf{R}_{11}$ to $\mathbf{R}_{21}$, which is shown in Table 3.2a.

Consumer Type	Opportunistic		Naive	
Product	1	2	1	2
$\mathbf{R}_{12}$	✓	✓	✓	
$\mathbf{R}_{11}$	✓	✓	✓	✓
$\mathbf{R}_{21}$		✓	✓	✓

(a) Case 1.

Consumer Type	Opportunistic		Naive	
Product	1	2	1	2
$\mathbf{R}_{12}$	✓	✓	✓	
$\mathbf{R}_{21}$		✓	✓	✓

(b) Case 2.

Consumer Type	Opportunistic		Naive	
Product	1	2	1	2
$\mathbf{R}_{12}$	✓	✓	✓	
$\mathbf{R}_{22}$		✓	✓	
$\mathbf{R}_{21}$		✓	✓	✓

(c) Case 3 and Case 4.

Table 3.2.: Evolution of the optimal solution (x_1 decreases from $\frac{1}{2}$ to 0).

3.4 Impact of δ

In this section, we study the impact of improving the quality of the regular product on the firm's profit. Specifically, we want to explore the how the firm's optimal profit changes when it adjusts the additional quality δ. The firm can influence δ by either providing more add-on services or increasing the probability of offering such services (e.g., free upgrades). Since quality improvement is one of the most important objectives for most firms, it is interesting to see whether it is beneficial. As a benchmark, we assume there is no cost in improving the quality of the regular product. Note such quality improvement only affects the opportunistic consumers' utilities but not the naive consumers' utilities.

The next proposition asserts that the optimal solution in region $\mathbf{R}_{11}$ will always occur when δ is sufficiently small (i.e., close to zero). That is, a small difference in the valuation of product 2 always leads to an optimal solution in which there are both types of consumers purchasing both products.

Proposition 3.4.1 *For any capacity level $x_1 \le \frac{1}{2}$, there exists a threshold $\bar{\delta}^1 \in (0, q_1 - q_2)$ such that the optimal solution of* (3.3) *belongs to $\mathbf{R}_{11}$ if and only if $\delta \in [0, \bar{\delta}^1]$.*

To see how the firm's optimal profit depends on δ, we first present the following lemma. Let π^*_{ij} denote the firm's optimal profit occured in region $\mathbf{R}_{ij}$.

Lemma 3.4.1 *The firm's optimal profit π^*_{11} in* (3.5) *is concave in δ, π^*_{12} in* (3.6) *is convex and increasing in δ, π^*_{21} in* (3.7) *is concave and increasing in δ, and π^*_{22} in* (3.8) *is linearly increasing in δ.*

The following proposition states that the optimal value π^* is increasing in δ if δ is sufficiently small; However, when x_1 is large enough, the optimal profit may also decrease in δ.

Proposition 3.4.2 *If $x_1 \in (0, \frac{1}{2})$, then*

1. $\frac{\partial}{\partial \delta}\pi^*(\delta) \mid_{\delta=0} > 0$;
2. *There exists $\bar{\delta}^2 \in (0, \bar{\delta}^1)$ and a threshold $k^* < \frac{1}{5}$ such that $\pi^*(\delta)$ is decreasing in $\delta \in [\bar{\delta}^2, \bar{\delta}^1]$ if and only if $x_1 > k^*$, and $\pi^*(\delta)$ is increasing in δ otherwise. Moreover, the threshold k^* does not depend on x_1 and r.*

Proposition 3.4.2 indicates there are two possible curves for $\pi^*(\delta)$. First, if $x_1 \le k^*$ ($k^* < \frac{1}{5}$), then $\pi^*(\delta)$ is monotonically increasing in δ (i.e., increasing the quality of add-on services or the probability of offering such services always improves the firm's profit). Second, if $x_1 > k^*$, then $\pi^*(\delta)$ is first increasing, then decreasing, and finally increasing again in δ (i.e., the firm's profit does not necessarily increases in δ). The decreasing range is $[\bar{\delta}^2, \bar{\delta}^1]$, which falls into region $\mathbf{R}_{11}$ (see the definition of $\bar{\delta}^1$ in Proposition 3.4.1). Why does a higher quality decrease the firm's profit? To understand this unexpected result, from (3.5), we can write the partial derivative of the optimal profit function π_{11} with respect to δ as follows:

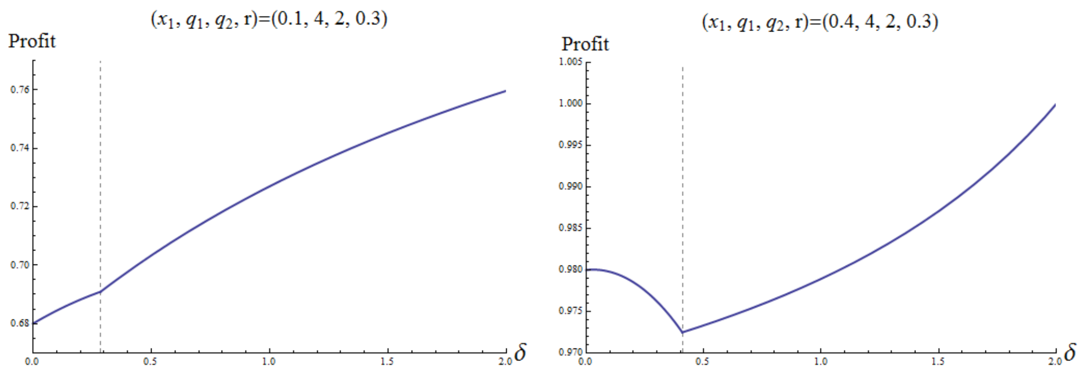
$$\frac{\partial \pi_{11}}{\partial \delta} = x_1 \frac{\partial p_1^*}{\partial \delta} + \left(\frac{1}{2} - x_1\right) \frac{\partial p_2^*}{\partial \delta}. \tag{3.9}$$

Since x_1 is a constant, the value of the derivative in (3.9) is determined by $\frac{\partial p_1^*}{\partial \delta}$ and $\frac{\partial p_2^*}{\partial \delta}$. Lemma 3.4.2 shows how the optimal prices (p_1^*, p_2^*) change with respect to δ.

Lemma 3.4.2 *Suppose (p_1^*, p_2^*) is in region $\mathbf{R}_{11}$ as (3.5), then p_1^* is concave and decreasing in δ, and p_2^* is concave and increasing in δ.*

With Lemma 3.4.2, we provide the following intuitive explanation for Proposition 3.4.2. When δ increases, product 2 becomes more valuable for the opportunistic consumers who have already chosen product 2. Thus, the firm will increase p_2 to capture the additional surplus from the increased consumer valuation; this is the direct, positive effect as measured by $(\frac{1}{2} - x_1)\frac{\partial p_2^*}{\partial \delta}$. Meanwhile, increasing δ also makes product 1 less attractive to the opportunistic consumers. So the firm has to lower price p_1 in order to prevent the opportunistic consumers from switching to product 2 (eventually the total demand of product 1 remains x_1). Since decreasing p_1 lowers the firm's profit from selling product 1, this is the indirect, negative effect as measured by $x_1 \frac{\partial p_1^*}{\partial \delta}$ (the negative effect is due to the firm's effort to prevent cannibalization). When δ is small, the competition or cannibalization between the two products is weak. Therefore, the direct, quality effect dominates the indirect, cannibalization effect, and the profit increases in δ, as shown in Part 1 of Proposition 3.4.2. As δ becomes larger, the cannibalization effect becomes stronger, and the profit decreases in δ if it dominates the positive direct effect. Note that a higher capacity level x_1

enhances the cannibalization effect but weakens the quality effect (a higher x_1 means more demand for product 1 while less demand for product 2). Therefore, when x_1 is relatively large, the latter effect can be easily dominated by the former one, which results in the non-monotone property of the optimal profit function as illustrated by Part 2 of Proposition 3.4.2. Figure 3.2 shows two cases of the monotonicity of $\pi^*(\delta)$. Note that Figure 3.2b represents the most interesting result in Proposition 3.4.2.



(a) $x_1 \leq k^* = 0.157$

(b) $x_1 > k^* = 0.157$

Figure 3.2.: The monotonicity of the firm's optimal profit $\pi^*(\delta)$ under different values for x_1.

Proposition 3.4.2 presents useful insights into how firms should manage quality for a product line. The optimal quality improvement decisions may depend on a variety of factors, such as the capacity of the premium product, the current quality, and so on. When the premium product is scarce, i.e., x_1 is relatively small, improving the quality of the regular produce will always benefit the firm. That is, offering more free add-on services will be the optimal strategy. On the other hand, if the capacity of the premium product is not restrictive, i.e., x_1 is relatively large, then the firm needs to be more cautious when making the quality improvement decisions. In this case, improving the quality of the regular product may significantly intensify the competition between the two products and thus hurt the firm's profit. For instance, choosing $\delta \in [0.1, 1.3]$ decreases the firm's profit in Figure 3.2b, where $q_1 = 4$, $q_2 = 2$ and $x_1 = 0.4$.

3.5 Impact of r

We proceed to examine the impact of increasing r on firm's optimal profit. In many practical situations, the firm can use advertisement to increase consumer awareness of the add-on services, or provide training to influence consumer behavior (e.g., there are websites that teach consumers how to obtain free upgrades in airlines and hotels). That is, a firm may change the parameter value r in our model setting. To help firms make this decision, it would be helpful to investigate how the firm's optimal profit depends on r. To focus on the impact of r on the firm's optimal profit, we assume the cost for changing r is zero.

The next lemma shows how the profit functions π_{11}^* and π_{12}^* depend r.

Lemma 3.5.1 *π_{11}^* in (3.5) is convex in r, and π_{12}^* in (3.6) is concave and strictly increasing in r.*

Similar to Proposition 3.4.2, the following proposition shows that the firm's optimal profit may either increase or decrease in r under different conditions.

Proposition 3.5.1 *Suppose $\frac{\delta}{q_1 - q_2} < \frac{1}{2}$. There exists $\bar{r} \in (0, 1)$ and a threshold $\tilde{k} \in [\frac{\delta}{q_1 - q_2}, \frac{1}{2})$ such that:*

1. *If $x_1 \in \left(\tilde{k}, \frac{1}{2}\right)$, then the firm's optimal profit is decreasing in $r \in [0, \bar{r}]$ and increasing in $r \in (\bar{r}, 1]$.*

2. *If $x_1 \in \left[\frac{\delta}{q_1 - q_2}, \tilde{k}\right]$, then the firm's optimal profit is increasing in $r \in [0, 1]$.*

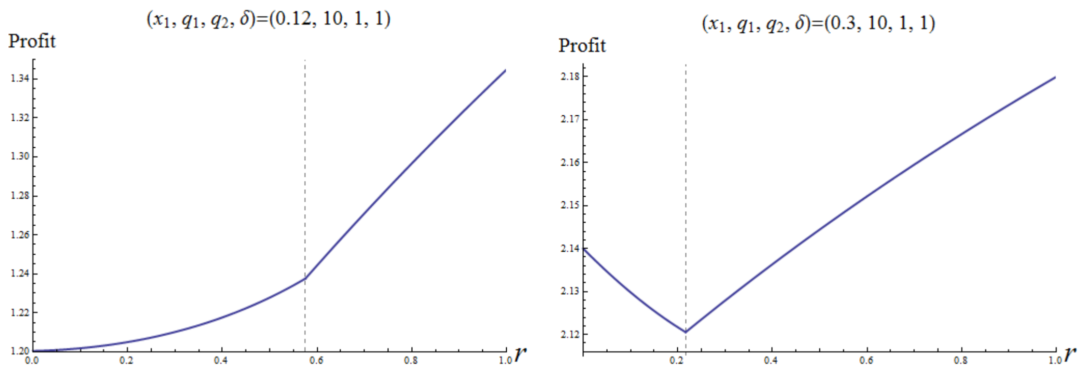
Proposition 3.5.1 indicates when capacity x_1 is large and the add-on quality δ is small, the firm's optimal profit is non-monotone in r. In addition, the decreasing part occurs in region $\mathbf{R}_{11}$ only. This result is interesting in two aspects. On one hand, it implies more consumers know the add-on services could be detrimental to the firm (the decreasing part); on the other hand, it means more opportunistic consumers may also benefit the firm (the increasing part). To help us understand this result, we can write the partial derivative of the optimal profit function π_{11} in (3.5) with respect to r as follows:

$$\frac{\partial \pi_{11}}{\partial r} = x_1 \frac{\partial p_1^*}{\partial r} + \left(\frac{1}{2} - x_1\right) \frac{\partial p_2^*}{\partial r}. \tag{3.10}$$

Lemma 3.5.2 states how the optimal prices (p_1^*, p_2^*) change with respect to r in region $\mathbf{R}_{11}$.

Lemma 3.5.2 *Suppose* (p_1^*, p_2^*) *is in region* $\mathbf{R}_{11}$ *as* (3.5)*, then* p_1^* *is non-monotone convex in* r*, and* p_2^* *is convex increasing in* r*.*

We provide the following intuitive explanation for the fact that optimal profit may decrease in r as demonstrated by Proposition 3.5.1. When r increases, there is a larger portion of consumers who are opportunistic and aware of the total value of product 2. Consequently, the firm should increase p_2 to capture the additional surplus due to the larger portion of high-valuation consumers; this is the direct, positive effect (which can be also called the advertisement effect), measured by $(\frac{1}{2} - x_1)\frac{\partial p_2^*}{\partial r}$. Meanwhile, increasing r enlarges the size of opportunistic consumers, which view product 1 as less attractive. So the firm has to lower price p_1 in order to prevent the opportunistic consumers from switching to product 2 (eventually the total demand of product 1 remains x_1). Thus, p_1^* decreases in r at the beginning, which lowers the firm's profit from selling product 1. This is the indirect, negative effect (which can be also called the cannibalization effect), measured by $x_1\frac{\partial p_1^*}{\partial r}$. Moreover, a larger capacity x_1 intensifies the "cannibalization effect" and dampens the "advertisement effect". When r is small, the cannibalization effect dominates the advertisement effect, and the profit will decrease in r, which corresponds to Part 1 of Proposition 3.5.1. On the other hand, when r is relatively large, the dominance relationship reverses and the optimal profit increases in r. Figure 3.3 illustrates these two cases using an numerical example. Note that Figure 3.3b represents the more interesting result in Proposition 3.5.1.



(a) $x_1 \leq \tilde{k} = 0.127$ (b) $x_1 > \tilde{k} = 0.127$

Figure 3.3.: The monotonicity of the firm's optimal profit $\pi^*(r)$ under different values for x_1.

3.6 Numerical Studies

So far we have characterized the firm's optimal pricing strategies and obtained two main results. § 3.4 shows that the firm's optimal profit may be lower when more free add-on services are provided; and § 3.5 gives the conditions under which advertisement may actually hurt the firm's profit. In this section, we relax some model assumptions to check the robustness of these results. First, we introduce a capacity constraint for regular products as well (§ 3.6.1). Second, we drop the zero usage cost assumption and assume that the firm incurs a positive cost for each unit of the premium product that is consumed (§ 3.6.2). Lastly, we consider a random market size (§ 3.6.3). An analytical investigation is challenging due to the complexity of the problem; thus we rely on numerical experiments in this section. We have explored a wide range of parameter settings; however, to save space, we will only present some representative examples.

3.6.1 Capacity Constraint for Product 2

The basic model assumes there is unlimited supply of product 2. Now we relax this assumption by introducing a capacity limit x_2 for product 2 (recall there is a capacity constraint x_1 for product 1). Then the firm's optimization problem in (3.3) becomes

$$\begin{aligned} \max_{(p_1,p_2)\in R} \quad & \pi = p_1 d_1(p_1,p_2) + p_2 d_2(p_1,p_2) \\ \text{s.t.} \quad & d_1(p_1,p_2) \le x_1,\ d_2(p_1,p_2) \le x_2, \\ & R = \{(p_1,p_2) | 0 \le p_1 \le q_1, 0 \le p_2 \le q_2 + \delta\}. \end{aligned} \tag{3.11}$$

Figure 3.4 illustrates how the firm's optimal profit varies with respect to δ and r when $q_1 = 4$, $q_2 = 2$ and $x_1 = 0.3$. The capacity for product 2 is $x_2 = 0.13$ on the left-hand panel and $x_2 = 0.18$ on the right-hand panel. We can see that the qualitative results remain unchanged with a capacity constraint for product 2: The firm's optimal profit may decrease in the quality level for the add-on services attached with product 2 and more advertisement may hurt the firm's profit.

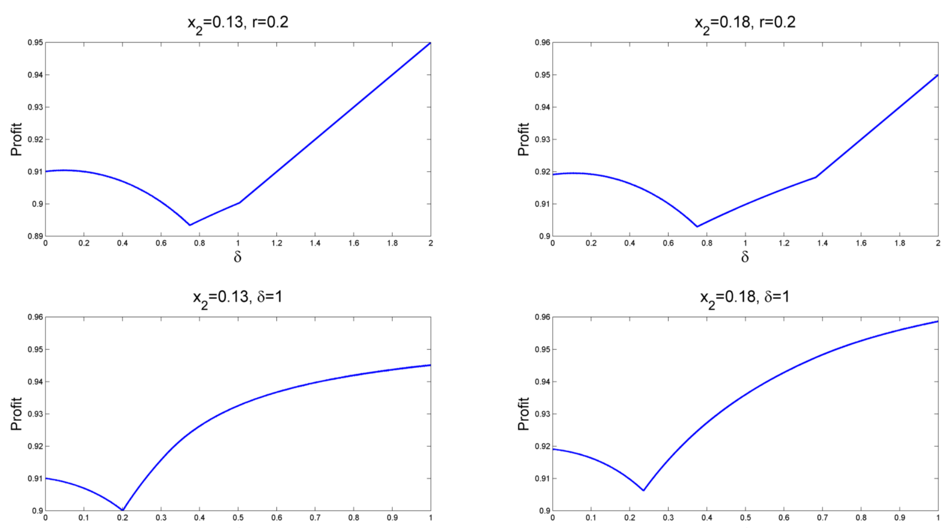


Figure 3.4.: Firm optimal profit as a function of δ and r (with capacity constraint for product 2).

3.6.2 Positive Usage Cost

The usage costs for both products are normalized to zero in the basic model. In service industries such as airlines and car rental companies, normally there is a usage cost associated with the product. To account for this fact, we assume there is a positive usage cost denoted by c for the premium product in this section. Since the usage cost for the regular product is usually lower than that for the premium product, we still assume there is a zero usage cost for the regular product. Given the usage cost $c > 0$ for product 1, the firm's optimization problem in (3.3) becomes

$$\begin{aligned} \max_{(p_1,p_2)\in R} \quad & \pi = (p_1 - c)d_1(p_1,p_2) + p_2 d_2(p_1,p_2) \\ \text{s.t.} \quad & d_1(p_1,p_2) \le x_1, \ \ R = \{(p_1,p_2) | 0 \le p_1 \le q_1, 0 \le p_2 \le q_2 + \delta\}. \end{aligned} \tag{3.12}$$

A representative numerical example is presented in Figure 3.5. Again, our main results still hold with different usage cost c's (e.g., $c = 0.1$ and $c = 0.3$) as shown in Figure 3.5 ($q_1 = 4$, $q_2 = 2$ and $x_1 = 0.4$ in this example).

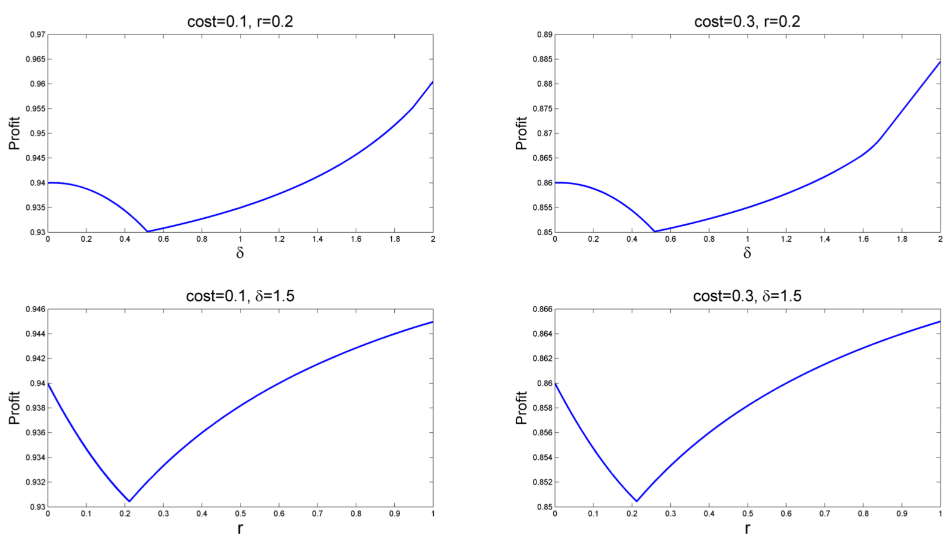


Figure 3.5.: Firm optimal profit as a function of δ and r (with positive usage cost for product 1).

3.6.3 Random Demand Size

In the last part of the numerical study, we incorporate random demand size into the basic model. Specifically, we assume the demand takes two possible values $1-z$ $(0 < z < 1)$ and $1+z$ with equal probability $\frac{1}{2}$. This two-point demand model is sufficient to capture the nature of random market size (the qualitative results will not change with a continuous random demand). The prices (p_1, p_2) are determined before the realization of the demand. Under this assumption, the demand of product i $(i = 1, 2)$ can be written as $d_i^-(p_1, p_2) = (1-z)d_i(p_1, p_2)$ for market realization $1-z$ and $d_i^+(p_1, p_2) = (1+z)d_i(p_1, p_2)$ for market realization $1+z$, where $d_i(p_1, p_2)$ is given in (3.1) and (3.2). Consequently, the new optimization problem becomes

$$\begin{aligned} \max_{(p_1,p_2)\in R} \qquad \pi =& \frac{1}{2}\left(p_1\left(d_1^-(p_1,p_2)\wedge x_1 + d_1^+(p_1,p_2)\wedge x_1\right) + p_2\left(d_2^-(p_1,p_2) + d_2^+(p_1,p_2)\right)\right) \\ \text{s.t.} \qquad & R = \{(p_1,p_2)|0 \le p_1 \le q_1, 0 \le p_2 \le q_2 + \delta\}. \end{aligned} \tag{3.13}$$

Figure 3.6 presents a numerical example with parameters $q_1 = 4$, $q_2 = 2$ and $x_1 = 0.4$. We can see that the firm's optimal profit can still be decreasing in both the quality of the free add-on services and the level of the advertisement. Thus our results are robust under random demand as well.

3.7 Conclusion

This chapter studies a monopoly firm's product quality and price decisions in the presence of the consumer heterogeneity and capacity constraint. The firm provides both a regular product and a premium product, the latter of which has a higher quality and limited capacity. Consumers are heterogeneous in two dimensions: First, they have different tastes for quality; second, they may or may not value the add-on services attached with the regular product. The firm needs to decide how much add-on service to offer with the regular product and then how to price the two products.

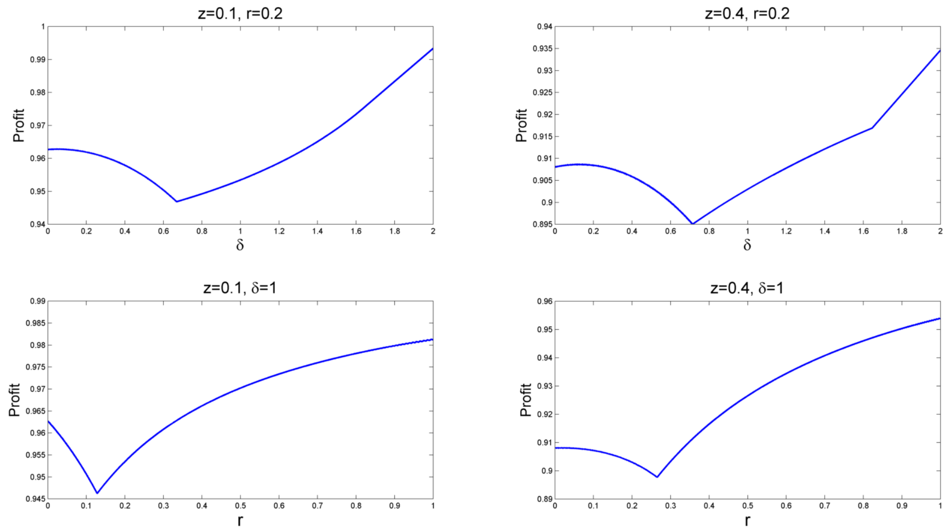


Figure 3.6.: Firm optimal profit as a function of δ and r (with random demand size).

We first show that the firm's optimal pricing strategy depends on the capacity constraint. With a sufficiently large supply of the premium product, the firm will just satisfy half of the consumers with only the premium product (i.e., the regular product will be taken out of the market). With a very limited supply of the premium product, the firm will always fully utilize the available capacity and prefers to sell it to consumers whose valuation of the regular product is lower. When the firm improves the quality of the regular product (by providing more add-on services), its profit may decrease due to the cannibalization effect (i.e., the quality improvement intensifies the competition between the firm's own products and drives the consumers away from purchasing the premium product). Furthermore, the profit decrease is more likely to happen when the capacity level of the premium product is higher since the firm has to lower the price of the premium product to mitigate the cannibalization effect. We also find that it is not always beneficial to the firm when it uses advertisement to increase the portion of consumers who value the add-on services associated with the regular product.

Chapter 4

Upgrading Under Opportunistic Consumers

4.1 Introduction

To cope with rapidly changing consumer trends, many firms have expanded their product lines and provided multiple products with different qualities. In addition to a regular product that includes basic functions desired by consumers, firms usually offer a premium product that incorporates additional services and features. On one hand, firms can achieve better market segmentation and extract higher profit by offering both products to consumers. On the other hand, firms face a more complicated problem in matching the supply with the demand because the demand is more predictable at the aggregated level than at the individual segment. Product upgrade as an operational strategy has been widely adopted in practice to mitigate the risk of mismatches between supply and demand. More details can be found in Chapter 2.

There are differences among consumers in terms of observable characteristics, such as opportunistic behavior and naive behavior, or in terms of unobservable characteristics, such as tastes for quality. When firms offer two products with the potential product upgrade, only the opportunistic (strategic) consumers take the upgrade probability into account, and the regular product becomes more attractive to an opportunistic consumer if she may receive a higher quality product with a positive probability while only paying the price of the regular product. Moreover, for each product offered by the firm, consumers may have different willingness to pay based on their tastes for quality. A detailed discussion is in Chapter 3.

A key assumption made in Chapter 3 is that consumers are heterogeneous in their perceptions of the quality of the regular product. Recall that δ is the valuation difference of the regular product between the opportunistic consumers and the naive ones. As the opportunistic consumers make their purchase decisions in anticipation of the potential product upgrade, δ can be viewed as the product of the quality difference between the premium product and the regular one and the probability of the product upgrade (i.e., obtaining the quality difference). However, as the qualities of both products are exogenously given, the probability of the product upgrade is also exogenously decided. Yet in reality, such probability depends on several factors including the capacity limits, the total number of consumers, and the fraction of consumers being opportunistic. For instance, a higher capacity limit of the premium product means there may be more potential leftover capacity that can be used as a product upgrade. A larger number of consumers has two-fold impacts: First, there are more consumers who may request the premium product, which results in less leftover capacity that can be used as upgrade; second, there are also more consumers who may purchase the regular product, which results in more consumers who are waiting for the product upgrade. A larger fraction of consumers being opportunistic implies that more consumers appreciate the potential product upgrade and may compete for the limited upgrading opportunities. Therefore, the probability of the product upgrade should be endogenously decided by both the firm and the consumers. And the opportunistic consumers should strategically learn such a probability by incorporating all contributing factors into their decision processes. The concept that consumers can rationally predict the future product availability has been studied in the operations management literature, for instance, Liu and van Ryzin (2008b), Su and Zhang (2008), Yin et al. (2009) and Cachon and Swinney (2009).

This chapter studies a monopoly firm's optimal pricing decisions when selling two differentiated products to heterogeneous consumers. There are two research questions that we would like to address: First, is consumers' opportunistic behavior beneficial or detrimental to the firm? As we discussed earlier, the opportunistic consumers choose between the premium product and the regular product based on their taste for quality as well as the probability of the future product upgrade. The firm may lose the sales of the premium product because some opportunistic consumers may instead purchase the regular product that yields a higher utility. However, as the utility of buying the regular product becomes higher, the firm may be able to increase the price of the regular product and extract more profit from the opportunistic consumers. It is unclear how the firm's profit may change when combining these two effects.

Second, how should the firm control the probability of the product upgrade? In service industries, firms may impose various restrictions on the eligibility of receiving the product upgrade. For example, airlines usually restrict the seat upgrade among consumers with certain booking codes and give a higher priority to consumers with elite status when assigning the seat upgrade. Similar examples can be found in the hotel industry as well as rental car companies. Furthermore, the firm can control the upgrade probability by adjusting the initial capacities. With a lower capacity level of the premium product, there may be less leftover capacities that can be used as an upgrade. Because firms clearly have the ability to adjust the frequency of the upgrade, we wish to understand how frequently the firm should offer the product upgrade.

The focus of this chapter is to extend the model in Chapter 3 and capture the opportunistic behavior. In particular, we study a specific upgrading problem where the probability of the product upgrade is endogenously determined by the consumer heterogeneity and the capacity limit. In a single-period model, a monopoly firm offers two products with differentiated qualities, the regular product and the premium product. The premium product has a limited capacity, whereas the regular product has ample capacity. Consumers with a random size arrive at the market, and each consumer prefers the premium product over the regular one at equal prices and chooses to buy one of them or nothing. After consumers make their product selections, the firm randomly distributes a fraction of the leftover capacity of the premium product to consumers who request the regular product as a free product upgrade. There are opportunistic consumers and naive consumers in the market, the former of which rationally predict the availability of the product upgrade and adjust their purchase decisions accordingly. The firm decides the prices for both products to achieve profit maximization.

Two major results about the firm's optimal strategies are obtained in this model. First, increasing the probability of the product upgrade has complicated impacts on the firm's profit. When increasing such a probability, the regular product becomes more attractive for the opportunistic consumers, and the firm may decrease the price of the premium product while increasing the price of the regular product, which restricts the number of opportunistic consumers changing their purchase decisions from the premium product to the regular one. Profit is more likely to increase if the quality difference between the two products is large. Our result suggests that the firm should pay extra attention when revising its upgrading policy.

Second, influencing consumers' opportunistic behavior through advertising can change the firm's profit in both directions. Consumers become more opportunistic when exposed to more advertisements or training about the product upgrade. Intuitively, more opportunistic consumers may increase the sales of the regular product due to their anticipation of the product upgrade. However, the profit from the premium product is reduced at the same time because the cannibalization effect between the two products becomes more severe. We recommend that firms use advertisements more wisely based on the model parameters.

The rest of this chapter is organized as follows. Section 4.2 reviews the strategic consumer behavior literature. The model setting is introduced in Section 4.3. Extensive numerical tests are conducted in Section 4.4 to derive managerial insights. The chapter concludes in Section 4.5.

4.2 Literature Review

This chapter extends the model in Chapter 3 by explicitly considering opportunistic (strategic) consumer behavior.

Consumers' purchase decisions depend on various factors including both future price and future product availability. Specifically, a group of papers study the scenario where consumers may strategically delay their purchases in anticipation of future price changes. For instance, Aviv and Pazgal (2008) study two classes of pricing strategies for a seasonal good with a limited quantity in the presence of forward-looking (strategic) customers. Elmaghraby et al. (2008) analyze the optimal markdown pricing mechanism in the presence of strategic buyers who request multiple units of the product. Su (2007) studies a dynamic pricing model with heterogeneous consumers and shows that the optimal price path could involve either markups or markdowns, depending on the composition of the customer pool. Another group of papers focus on the interaction between firms' inventory decisions and the strategic consumers' anticipation of future product availability. Liu and van Ryzin (2008b) investigate the firm's understocking quantity decisions, which may stimulate early purchase in a capacity-rationing model with strategic consumers. Liu and van Ryzin (2011) then extend the model into repeated seasons. Su and Zhang (2008) study a newsvendor seller facing strategic customers and find that either quantity or price commitment may improve the seller's profit. Su and Zhang (2009) further explore the benefit of product availability in attracting strategic consumer demand. Yin et al. (2009) consider how a fashion retailer can use two different inventory display formats (display all and display one) to mitigate the adverse impact of the strategic consumer behavior on the retailer's profit. Lai et al. (2010) consider a model where the seller has a posterior price matching policy and give the condition under which the price matching benefits the firm by eliminating strategic consumers' waiting incentive. Cachon and Swinney (2009) show that the value of the seller being able to procure additional inventory after obtaining updated demand information is generally much greater when strategic consumers are present.

In contrast, our model considers a different type of strategic behavior. We study the opportunistic consumers who anticipate the potential product upgrade while the above papers consider the anticipation of either the price change or the product availability. This chapter also extends the firm's product line by considering two quality differentiated products whereas the above models are limited to a single product. Lastly, our model incorporates the consumer differentiation, i.e., consumers have different tastes for quality, which is absent in the models above.

4.3 Model

We consider a model that has three groups of agents. In the supply side, there is a monopoly firm managing two types of products, product 1 with quality q_1 ($q_1 > 0$, the premium product) and product 2 with quality q_2 ($0 < q_2 < q_1$, the regular product). The firm has fixed capacities x_i ($i = 1, 2$) for product i. In the base model, the capacities are exogenously given, and we will use profit and revenue functions exchangeably. Similar to Chapter 3, we assume x_2 is large enough for analytical tractability such that it never constrains the sales. The firm decides the prices p_i for product i. After consumers make their decisions, the firm allocates both capacities to the corresponding demand. During the capacity allocation process, each consumer has equal probability to receive the product that she requested. Furthermore, the base model considers the complementary product upgrade, i.e., the firm may offer product 1 to consumers requesting product 2 at price p_2 as a courtesy. Specifically, the firm performs the product upgrade if there remain leftover capacities of product 1 after the capacity allocation process. However, the firm allocates at most κ ($0 \leq \kappa \leq 1$) fraction of the total leftover capacities of product 1 to consumers who originally purchase product 2. Thus, the firm uses κ to control the quantity or frequency of the upgrade. Let d_i ($i = 1, 2$) be the demand of product i before the capacity allocation, then the upgrade occurs if $d_1 < x_1$, and the upgrading probability is defined as $1 \wedge \frac{\kappa(x_1-d_1)^+}{d_2}$, where $(z)^+ = \max(z, 0)$ and $y \wedge z = \min(y, z)$. For notational convenience, we further define $y \vee z = \max(y, z)$.

In the demand side, there are two types of consumers, the opportunistic consumers and the naive consumers. The total mass of the consumers is a random variable Y that follows the distribution F and is independent of prices (p_1, p_2). A consumer chooses between the premium product and the regular one as well as choosing whether to purchase at all. Each of these consumers has a utility function $U = \theta q - p$ for a single unit of product, where q is the product quality, p is the product price, and $\theta \geq 0$ is the parameter which measures the intensity of the consumer's taste for quality. A consumer receives zero utility if she does not purchase. The utility function implies that all consumers prefer a higher quality product for a given price, but a consumer with a larger θ is more willing to pay to obtain the high quality product. We assume θ is uniformly distributed on the interval $[0, 1]$. The consumer chooses the option that yields the highest utility based on her taste for quality and consumer type. The opportunistic consumers are different from the naive ones in the sense that only they recognize the possibilities of product 1 being stock-out and the product being upgraded during the capacity allocation process. Specifically, an opportunistic consumer forms a private belief $\xi_o = (\xi_o^1, u_o)$ over probabilities of receiving product 1 if requesting product 1 and receiving product 1 if requesting product 2 (i.e., the product upgrade), respectively. For tractability, we assume that all opportunistic consumers share the same belief $\xi_o = (\xi_o^1, u_o)$. For the opportunistic consumers, the quality of the premium product is $\xi_o^1 q_1$, which incorporates the stock-out probability. And the quality of the regular product is $q_2 + u_o(q_1 - q_2)$, which considers the probability of receiving the product upgrade. In contrast, the quality of product i for the naive consumers remains q_i. r $(0 \leq r \leq 1)$ and $1 - r$ are the probabilities of consumers being opportunistic and naive, respectively. Note that r is observable to the firm.

4.3.1 Sequence of Events

We now summarize the timeline in our model. The firm forms the belief $\xi_f = (s_f^1, u_f)$, which is a belief over the same probabilities as the opportunistic consumers' belief ξ_o. Then, the firm optimally decides prices p_1 and p_2 to maximize the profit $\pi(p_1, p_2)$. Although the firm is informed about different types of consumers, it is unable to differentiate individual consumers and charge different prices. Both the opportunistic and naive consumers decide whether to buy and which product to buy. However, the opportunistic consumers form the belief $\xi_o = (s_o^1, u_o)$ and make their decisions accordingly. Next, the random consumer size Y is realized. Finally, sales occur at the prices (p_1, p_2) after the firm allocates the capacities to the corresponding demand and performs the product upgrade. Our model follows the definition of the rational expectation equilibrium proposed by Su and Zhang (2008) and requires that $\xi_f = \xi_o$.

4.3.2 Analysis

To analyze the firm's profit maximization problem, we start with consumer demand functions. Let $d_i(p_1, p_2, \xi_o)$ $(i = 1, 2)$ be the demand for product i under prices (p_1, p_2) and the common belief $\xi_o = (\xi_o^1, u_o)$ among the opportunistic consumers. The utilities of buying product 1 and 2 for an opportunistic consumer with the taste parameter θ are $\theta\xi_o^1 q_1 - p_1$ and $\theta\left(q_2 + u_o(q_1 - q_2)\right) - p_2$, respectively. The consumer purchases product 1 if $\theta\xi_o^1 q_1 - p_1 \geq \theta\left(q_2 + u_o(q_1 - q_2)\right) - p_2$ and $\theta\xi_o^1 q_1 - p_1 \geq 0$ or product 2 if $\theta\left(q_2 + u_o(q_1 - q_2)\right) - p_2 \geq \theta\xi_o^1 q_1 - p_1$ and $\theta\left(q_2 + u_o(q_1 - q_2)\right) - p_2 \geq 0$. Therefore, given the firm's prices (p_1, p_2), an opportunistic consumer purchases product 1 if her taste parameter

$$\theta \in \left[\frac{p_1 - p_2}{(\xi_o^1 - u_o)q_1 - (1 - u_o)q_2} \vee \frac{p_1}{\xi_o^1 q_1}, \quad 1\right],$$

or product 2 if

$$\theta \in \left[\frac{p_2}{q_2 + u_o(q_1 - q_2)}, \quad \frac{p_1 - p_2}{(\xi_o^1 - u_o)q_1 - (1 - u_o)q_2} \vee \frac{p_1}{\xi_o^1 q_1}\right].$$

The same argument can be applied to the naive consumers by using (q_1, q_2) instead of $(\xi_o^1 q_1, q_2 + u_o(q_1 - q_2))$. Therefore, the demand for product 1 is

$$d_1(p_1, p_2, \xi_o) = Y\left(r\left(1 - \frac{p_1 - p_2}{(\xi_o^1 - u_o)q_1 - (1 - u_o)q_2} \vee \frac{p_1}{\xi_o^1 q_1}\right)^+ + (1-r)\left(1 - \frac{p_1 - p_2}{q_1 - q_2} \vee \frac{p_1}{q_1}\right)^+\right). \tag{4.1}$$

And the demand for product 2 is

$$d_2(p_1, p_2, \xi_o) = Y\left(r\left(1 \wedge \left(\frac{p_1 - p_2}{(\xi_o^1 - u_o)q_1 - (1 - u_o)q_2} \vee \frac{p_1}{\xi_o^1 q_1}\right) - \frac{p_2}{q_2 + u_o(q_1 - q_2)}\right)^+ \right.$$
$$\left. +(1-r)\left(1 \wedge \left(\frac{p_1 - p_2}{q_1 - q_2} \vee \frac{p_1}{q_1}\right) - \frac{p_2}{q_2}\right)^+\right). \tag{4.2}$$

Define $\mathbf{R} = \{(p_1, p_2) : \ 0 \le p_1 \le q_1, \ 0 \le p_2 \le q_1\}$. We can show that it is sufficient to only examine $\mathbf{R}$ when searching for the firm's optimal prices. To prove this, suppose to the contrary that $(p_1, p_2) \notin \mathbf{R}$ constitute the firm's optimal prices. We assume without loss of generality that $p_1 > q_1$ and $0 \le p_2 \le q_1$. For any belief ξ_o, there is $\pi(q_1, p_2) = \pi(p_1, p_2)$ since $d_1(p_1, p_2, \xi_o) = d_1(q_1, p_2, \xi_o) = Y$ by (4.1) and $d_2(p_1, p_2, \xi_o) = d_2(q_1, p_2, \xi_o)$ by (4.2). Similarly, we can show that the optimal $p_2 \in [0, q_1]$. Hence, we will restrict to the region $\mathbf{R}$ in subsequent analysis.

Now the firm's optimization problem can be written as follows:

$$\max_{(p_1,p_2)\in\mathbf{R}} \quad \pi(p_1, p_2) = p_1\,\mathbb{E}\left[d_1(p_1, p_2, \xi_o) \wedge x_1\right] + p_2\,\mathbb{E}\left[d_2(p_1, p_2, \xi_o)\right] \tag{4.3}$$

$$\text{s.t.} \quad \xi_o^1 = \mathbb{E}\left[1 \wedge \frac{x_1}{d_1(p_1, p_2, \xi_o)}\right], \tag{4.4}$$

$$u_o = \mathbb{E}\left[1 \wedge \frac{\kappa\,(x_1 - d_1(p_1, p_2, \xi_o))^+}{d_2(p_1, p_2, \xi_o)}\right], \tag{4.5}$$

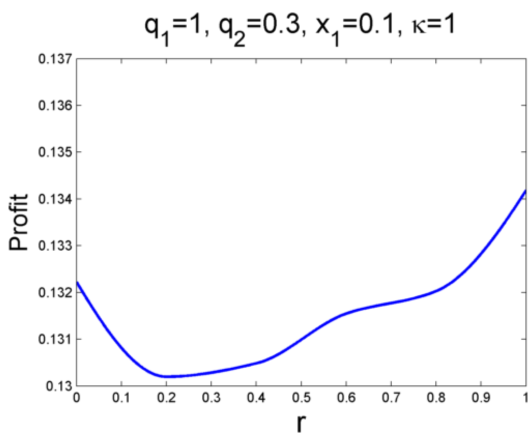
where $\frac{x_1}{d_1(p_1,p_2,\xi_o)}$ is defined to be 1 if $d_1(p_1, p_2, \xi_o) = 0$, and $\frac{\kappa(x_1 - d_1(p_1,p_2,\xi_o))^+}{d_2(p_1,p_2,\xi_o)}$ is defined to be 0 if $d_2(p_1, p_2, \xi_o) = 0$.

4.4 Numerical Studies

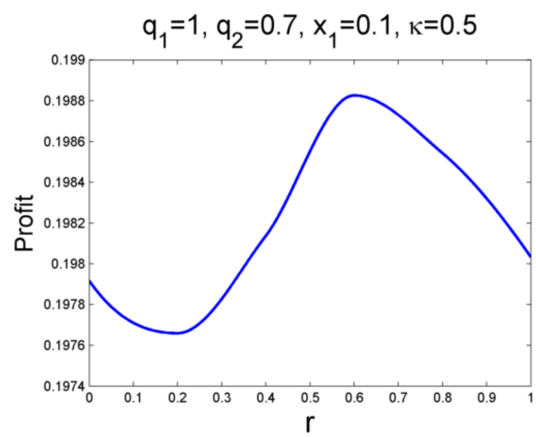
Similar to the model in Chapter 3, the profit function (4.3) is not unimodal. Furthermore, the objective is maximized subject to the rational expectation equilibrium constraints (4.4) and (4.5), which means that we have to solve a non-linear equations system for each pair of prices (p_1, p_2). Due to the complexity of this problem, the results of this section are derived from extensive numerical tests. This numerical study includes two parts: the impacts of the firm's advertising decision and upgrading decision.

First, we examine how the firm's advertising decision affects its profit. Recall that advertisement is a useful instrument to improve the fraction of consumers being opportunistic. On one hand, advertising the potential product upgrade can educate naive consumers and make them purchase strategically, which increases the number of consumers who have a higher valuation of the regular product and allows the firm to raise the price of the regular product. On the other hand, advertising increases the number of opportunistic consumers who may switch from the premium product to the regular product, hoping to get a potential free upgrade, which may cannibalize the sales from the premium product. In terms of the probability of the product upgrade, having more consumers purchase the regular product should decrease their individual probability to receive the product upgrade, however, having less consumers purchase the premium product should increase such probability. Therefore, it is unknown how the firm's profit changes with respect to r, which represents the fraction of the opportunistic consumers. From Figure 4.1, it is clear that the firm's optimal profit is non-monotone in r. Specifically, Figure 4.1a shows that the highest profit is reached when $r = 1$ if the quality difference between the two products is large. In this case, the firm benefits from a higher price for the regular product, which is the result of the opportunistic behavior. And such a profit gain is amplified by the large $\kappa = 1$, which encourages the opportunistic consumers by offering all leftover capacities of the premium product as product upgrade. However, Figure 4.1b illustrates a different scenario where the highest profit is achieved when $0 < r < 1$. Note that the quality difference between the two products is smaller in this case. We offer the following explanation: The cannibalization effect between the two products is stronger among the opportunistic consumers, whose valuation of the regular product is greater than

q_2, than among the naive consumers. As r increases from 0 to 1, the firm benefits from the opportunistic behavior initially. But when r is close to 1, most consumers are opportunistic. With a small quality difference $q_1 - q_2$, the cannibalization effect becomes dominant and lowers the firm's profit.



(a) $q_2 = 0.3$ and $\kappa = 1$.

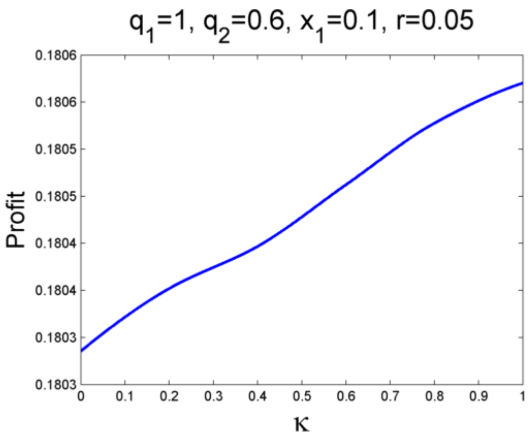


(b) $q_2 = 0.7$ and $\kappa = 0.5$.

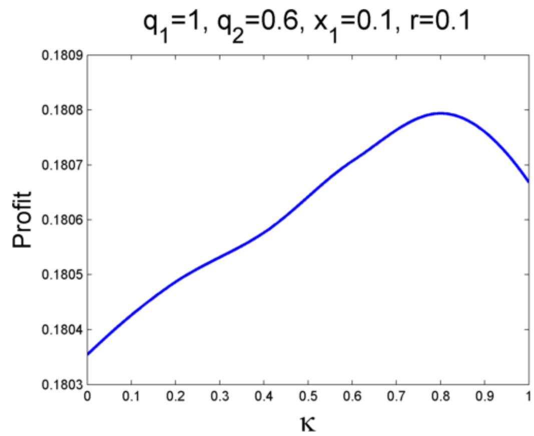
Figure 4.1.: Firm optimal profit as a function of r.

The second part of the numerical study considers the firm's decision about the upgrading probability. The firm only uses κ fraction of the leftover capacity of the premium product to upgrade consumers requesting the regular product. When the firm chooses a larger κ, there is a greater probability of the potential product upgrade, and the firm may be able to charge a higher price for the regular product because the opportunistic consumers appreciate the product upgrade. However, with a smaller κ, there is a bigger quality difference between the regular product and the premium one for the opportunistic consumers, and the cannibalization effect may be reduced. Figure 4.2 shows that the firm's profit can either decrease or increase with respect to κ and the highest profit may occur in the interior of $[0, 1]$. Figure 4.2b and 4.2d represent the most interesting cases. In particular, the highest profit is achieved when $0 < \kappa < 1$ in Figure 4.2b. To understand it, note that $r = 0.1$ in Figure 4.2b, which is larger than $r = 0.05$ in Figure 4.2a. If q_2 is small as in Figure 4.2a and 4.2b, increasing κ benefits the firm in general because the profit increase from the higher price of the regular product is larger than the profit loss due to the cannibalization effect between the two products. However, as r becomes larger, i.e., there is a larger fraction of consumers who are opportunistic and can be directly affected by the firm's decision κ, the firm's profit may eventually decrease in κ because the number of opportunistic consumers, who may choose the regular product instead of the premium one in anticipation of the potential product upgrade, becomes much larger. However, if q_2 is large as in Figure 4.2c and 4.2d, the aforementioned cannibalization effect is already strong even without the product upgrade (i.e., $\kappa = 0$). As r increases from Figure 4.2c to 4.2d, the firm benefits from the additional profit from selling the regular product to the opportunistic consumers at a higher price, which can be further increased by offering a more generous product upgrade. Hence, the firm's profit can increase in κ when κ is large.

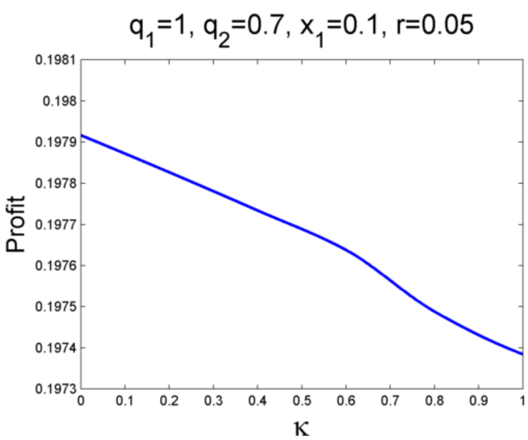
Last, we examine the impact of changing the quality difference $q_1 - q_2$ and the capacity constraint x_1 for the premium product. As q_1 is fixed, improving the quality q_2 of the regular product is equivalent to decreasing the quality difference $q_1 - q_2$. From Figures 4.3 and 4.4, we can see that the firm's optimal profit is always increasing in the quality of q_2 as well as the capacity x_1. The intuition is straightforward: without the cost of increasing the quality q_2, the firm has a better product to sell, which implies that the firm is in a better position to generate a higher profit. Even though the cannibalization effect becomes stronger, the additional benefit from the higher quality q_2 dominates the negative effect. Similar argument can be applied to changing x_1. Acquiring more capacity x_1 increases the probability of product upgrade, which implicitly increases the quality of the regular product. Note that the profit curves will correspondingly change if we incorporate the cost of the quality improvement or the capacity acquisition.



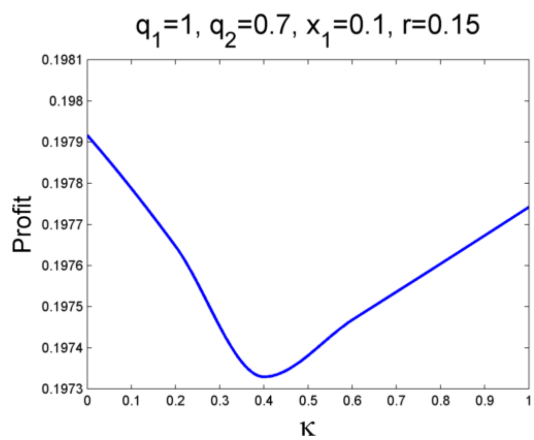
(a) $q_2 = 0.6$ and $r = 0.05$.



(b) $q_2 = 0.6$ and $r = 0.1$.



(c) $q_2 = 0.7$ and $r = 0.05$.

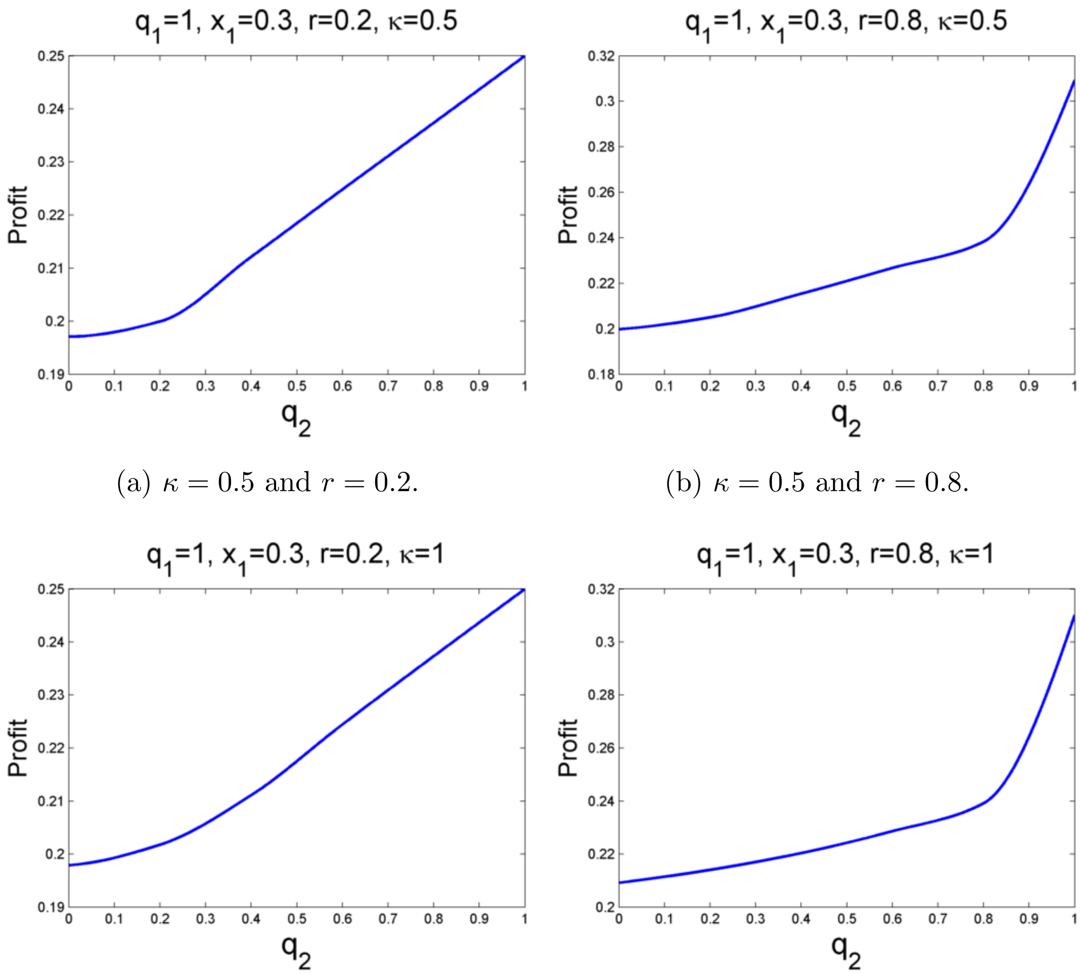


(d) $q_2 = 0.7$ and $r = 0.15$.

Figure 4.2.: Firm optimal profit as a function of κ.

4.5 Conclusion

This chapter studies a monopoly firm's price and upgrading decisions in the presence of the consumer heterogeneity and capacity constraint. The firm offers both the regular product and the premium product with capacity limits and uses a fraction of the leftover capacity of the premium product to upgrade consumers purchasing the regular product. There are two types of consumers whose total size is a random variable. The opportunistic consumers rationally predict the probability of the potential product upgrade and make their purchase decisions accordingly, whereas the naive



(a) $\kappa = 0.5$ and $r = 0.2$.

(b) $\kappa = 0.5$ and $r = 0.8$.

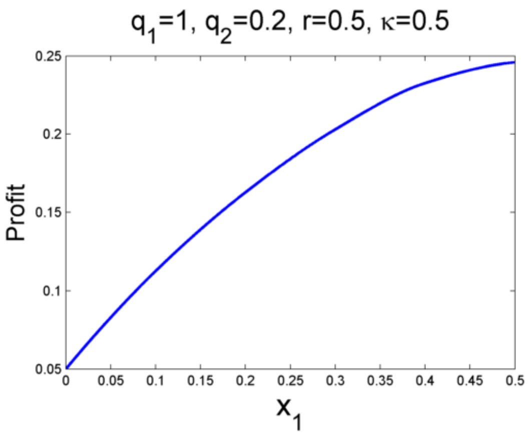
(c) $\kappa = 1$ and $r = 0.2$.

(d) $\kappa = 1$ and $r = 0.8$.

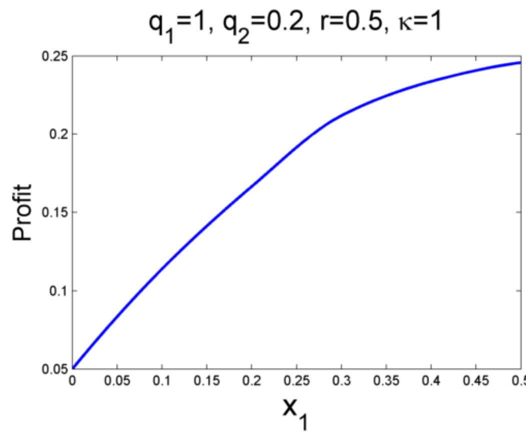
Figure 4.3.: Firm optimal profit as a function of q_2.

consumers ignore such a probability. The firm decides the price for each product to maximize its profit.

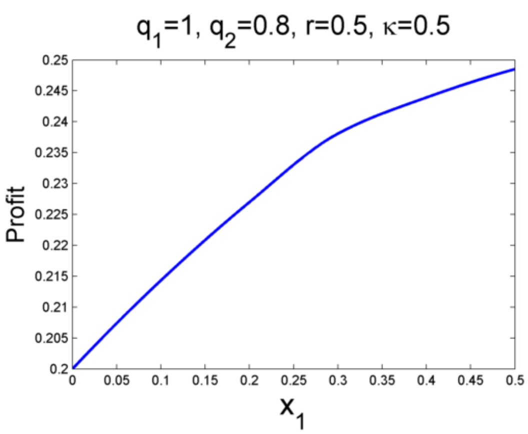
We confirm that the analytical results discussed in Chapter 3 still hold in this specific upgrading problem. Particularly, the firm's advertising and upgrading decisions can change the profit in either direction. The firm needs to pay extra attention to these decisions whose impacts depend on the quality difference between the two products and other model specifications.



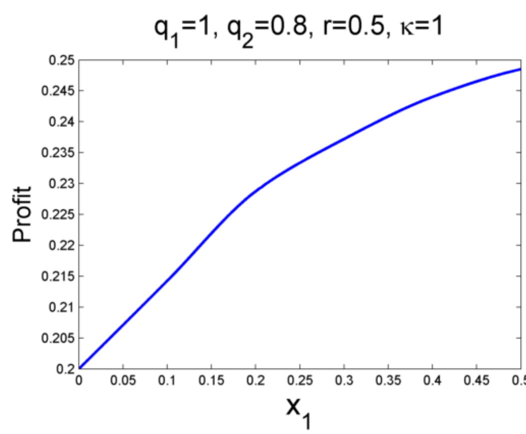
(a) $q_2 = 0.2$ and $\kappa = 0.5$.



(b) $q_2 = 0.2$ and $\kappa = 1$.



(c) $q_2 = 0.8$ and $\kappa = 0.5$.



(d) $q_2 = 0.8$ and $\kappa = 1$.

Figure 4.4.: Firm optimal profit as a function of x_1.

Chapter 5

Conclusion and Future Research

This dissertation studies the impact of upgrading on firms' operational strategies. Three problem settings have been considered. The main results can be summarized as follows.

Chapter 2 studies a firm's capacity investment and allocation decisions in a dynamic setting with stochastic demand. There are N classes of products, each of which corresponds to a demand class that arrives in each period. The firm has to decide the initial capacity for each class of product before the beginning of the selling season, then allocate capacities to incoming consumers in each period before future demand is realized. The model considers a general upgrading rule that covers most of the practical applications. We show that a Parallel and Sequential Rationing (PSR) policy is the optimal allocation rule in each period for any given initial capacity. The complexity of the allocation problem can be greatly reduced by the PSR policy, where the firm first satisfies demand by the same-class capacity as much as possible and then sequentially upgrades leftover demand. Chapter 2 also proposes an efficient heuristic, Refined Certainty Equivalence Control (RCEC), that exploits the structural properties of the PSR policy and yields close-to-optimum solutions for the firm. With the help of the RCEC heuristic, extensive numerical studies show that the multi-step upgrading is highly valuable when the capacities are not balanced. Moreover, it is illustrated that the allocation decision is much more important than the initial capacity decision for the firm.

Chapter 3 studies a monopolist firm's optimal strategies in the presence of the consumer heterogeneity and capacity constraint. The firm offers two differentiated products, the regular product and the premium product, to two types of consumers, the opportunistic consumers and the naive ones. The difference between the two consumer types is that only the opportunistic consumers appreciate the add-on services included in the regular product and thus have a higher valuation of the regular product. The consumer's product selection depends on his type and taste for quality. The firm needs to decide how much add-on service to offer and then how to price the two products. We show that the firm's optimal pricing strategy depends on the capacity constraint. With a limited capacity of the premium product, the firm always fully utilizes the available capacity and prefers to sell it to naive consumers whose valuation of the regular product is lower. We find that improving the quality of the add-on services can change the firm's profit in both directions. Similar result has been found about influencing the fraction of opportunistic consumers, i.e., a larger fraction of opportunistic consumers may increase or decrease the firm's profit. Numerical results confirm the robustness of these results in more general model settings. Chapter 4 extends Chapter 3 by endogenizing the upgrading probability. A specific upgrading problem is introduced, where a consumer may be upgraded if there is leftover capacity of the premium product after satisfying an uncertain demand. An opportunistic consumer rationally predicts the probability of the potential product upgrade and adjusts his product selection accordingly, while a naive consumer disregards the information about the product upgrade. The upgrading probability is determined by the capacity level of the premium product as well as the consumers' strategic behavior. The main results in Chapter 3 have been verified in this complex but more realistic setting, which also provides insight into how firms should control the upgrading frequency.

Future research may be done in the following directions. In Chapter 2, we assume that firms' cost parameters are constant over time and unmet demand is backlogged. There are several interesting extensions of this research. First, it is worthwhile exploring models with general non-stationary model parameters. The PSR policy remains optimal if the profit margin is monotonically decreasing over time. However, with general non-stationary model parameters, the optimal policy is still unknown. Second, it is a challenge to analyze models with lost sales. The backorder assumption used in this chapter is critical for the optimal PSR allocation policy. It is not clear how the optimal policy looks under the lost-sales assumption. Third, it would be interesting to take pricing decisions into account, i.e., the firm may adjust prices over time depending on the evolution of demand and remaining capacity levels.

In Chapter 3, we assume that the initial qualities of both products are exogenously given and there is only one firm in the market. There are two potential extensions of this research. First, it would be interesting to incorporate the firm's decisions about the initial qualities of the products (i.e., the firm can determine the optimal qualities of both the premium product and the regular one before deciding the qualities of the add-on services). Second, though challenging, we may extend our model to a competitive setting. It would be interesting to see how competition will change the results derived from our monopoly model.

LIST OF REFERENCES

LIST OF REFERENCES

Acimovic, J. and S. C. Graves (2013). Making better fulfillment decisions on the fly in an online retail environment. Working paper.

Archibald, T. W., S. A. E. Sassen, and L. C. Thomas (1997). An optimal policy for a two depot inventory problem with stock transfer. *Management Science 43*(2), 173–183.

Aviv, Y. and A. Pazgal (2008, Summer). Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management 10*(3), 339–359.

Axsäter, S. (2003). A new decision rule for lateral transshipments in inventory systems. *Management Science 49*(9), 1168–1179.

Bassok, Y., R. Anupindi, and R. Akella (1999, July/August). Single-period multi-product inventory models with substitution. *Operations Research 47*(4), 632–642.

Bertsekas, D. P. (2005a). *Dynamic Programming and Optimal Control*, Volume 1. Athena Scientific.

Bertsekas, D. P. (2005b). Dynamic programming and suboptimal control: A survey from adp to mpc. *European Journal of Control 11*(4-5), 310–334.

Bish, E. K. and Q. Wang (2004, November/December). Optimal investment strategies for flexible resources, considering pricing and correlated demands. *Operations Research 52*(6), 954–964.

Brenner, U. (2008). A faster polynomial algorithm for the unbalanced hitchcock transportation problem. *Operations Research Letters 36*(4), 408–413.

Cachon, G. P. and R. Swinney (2009). Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science 55*(3), 497–511.

Chambers, C., P. Kouvelis, and J. Semple (2006). Quality-based competition, profitability, and variable costs. *Management Science 52*(12), 1884–1895.

Chen, X., M. Sim, P. Sun, and J. Zhang (2008, Mar). A linear decision-based approximation approach to stochastic programming. *Operations Research 56*(2), 344–357.

Chod, J. and N. Rudi (2005, May/June). Resource flexibility with responsive pricing. *Operations Research 53*(3), 532–548.

Chou, M. C., G. A. Chua, C.-P. Teo, and H. Zheng (2010, January/February). Design for process flexibility: Efficiency of the long chain and sparse structure. *Operations Research 58*(1), 43–58.

Chou, M. C., G. A. Chua, C.-P. Teo, and H. Zheng (2011, September/October). Process flexibility revisited: The graph expander and its applications. *Operations Research 59*(5), 1090–1105.

Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature 34*(2), 669–700.

de Véricourt, F., F. Karaesmen, and Y. Dallery (2001, Spring). Assessing the benefits of different stock-allocation policies for a make-to-stock production system. *M&SOM 3*(2), 105–121.

de Véricourt, F., F. Karaesmen, and Y. Dallery (2002). Optimal stock allocation for a capacitated supply system. *Management Science 48*(11), 1486–1501.

Deshpande, V., M. A. Cohen, and K. Donohue (2003). A threshold inventory rationing policy for service-differentiated demand classes. *Management Science 49*(6), 683–703.

Ding, Q., P. Kouvelis, and J. M. Milner (2006). Dynamic pricing through discounts for optimizing multiple-class demand fulfillment. *Operations Research 54*(1), 169–183.

Donnenfeld, S. and S. Weber (1992). Vertical product differentiation with entry. *International Journal of Industrial Organization 10*(3), 449–472.

Ellison, G. (2005). A model of add-on pricing. *The Quarterly Journal of Economics 120*(2), 585–637.

Elmaghraby, W., A. Gülcü, and P. Keskinocak (2008). Designing optimal preannounced markdowns in the presence of rational customers with multiunit demands. *Manufacturing & Service Operations Management 10*(1), 126–148.

Fine, C. H. and R. M. Freund (1990). Optimal investment in product-flexible manufacturing capacity. *Management Science 36*(4), 449–466.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Gabszewicz, J. J. and J.-F. Thisse (1979). Price competition, quality and income disparities. *Journal of Economic Theory 20*(3), 340–359.

Gabszewicz, J. J. and J.-F. Thisse (1980). Entry (and exit) in a differentiated industry. *Journal of Economic Theory 22*(2), 327–338.

Gabszwicz, J. J. and J.-F. Thisse (1986). On the nature of competition with differentiated products. *The Economic Journal 96*(381), 160–172.

Gallego, G. and C. Stefanescu (2009). Upgrades, upsells and pricing in revenue management. Available at SSRN: http://ssrn.com/abstract=1334341.

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

Ha, A. Y. (1997). Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. *Naval Research Logistics 44*(5), 457–472.

Ha, A. Y. (2000). Stock rationing in an $M/E_k/1$ make-to-stock queue. *Management Science 46*(1), 77–87.

Ho, T.-H., C. S. Tang, and D. R. Bell (1998). Rational shopping behavior and the option value of variable pricing. *Management Science 44*(12), 145–160.

Hsu, A. and Y. Bassok (1999, March/April). Random yield and random demand in a production system with downward substitution. *Operations Research 47*(2), 277–290.

Jacobson, R. and C. Obermiller (1990). The formation of expected future price: A reference price for forward- looking consumers. *Journal of Consumer Research 16*(4), 420–432.

Jordan, W. C. and S. C. Graves (1995, April). Principles on the benefits of manufacturing process flexibility. *Management Science 41*(4), 577–594.

Krishna, A., I. S. Currim, and R. W. Shoemaker (1991). Consumer perceptions of promotional activity. *Journal of Marketing 55*(2), 4–16.

Lai, G., L. G. Debo, and K. Sycara (2010). Buy now and match later: Impact of posterior price matching on profit with strategic consumers. *Manufacturing & Service Operations Management 12*(1), 33–55.

Leachman, R. (1987). Preliminary design and development of a corporate-level production system for the semiconductor industry. Working Paper ORC 86-11. Operations Research Center, University of California, Berkeley, CA.

Lehmann-Grube, U. (1997). Strategic choice of quality when quality is costly: The persistence of the high-quality advantage. *The RAND Journal of Economics 28*(2), 372–384.

Liu, Q. and G. van Ryzin (2008a, Spring). On the choice-based linear programming model for network revenue management. *M&SOM 10*(2), 288–310.

Liu, Q. and G. van Ryzin (2011). Strategic capacity rationing when customers learn. *Manufacturing & Service Operations Management 13*(1), 89–107.

Liu, Q. and G. J. van Ryzin (2008b). Strategic capacity rationing to induce early purchases. *Management Science 54*(6), 1115–1131.

Moorthy, K. S. (1984). Market segmentation, self-selection, and product line design. *Marketing Science 3*(4), 288–307.

Moorthy, K. S. (1988, Spring). Product and price competition in a duopoly. *Marketing Science 7*(2), 141–168.

Murty, K. G. (1983). *Linear Programming.* New York: John Wiley & Sons.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Netessine, S., G. Dobson, and R. A. Shumsky (2002, March/April). Flexible service capacity: Optimal investment and the impact of demand correlation. *Operations Research 50*(2), 375–388.

Parlar, M. and S. Goyal (1984). Optimal ordering decisions for two substitutable product with stochastic demands. *OPSEARCH 21*, 1–15.

Pasternack, B. A. and Z. Drezner (1991, April). Optimal inventory policies for substitutable commodities with stochastic demand. *Naval Research Logistics 38*(2), 221–240.

Rhee, B.-D. (1996). Consumer heterogeneity and strategic quality decisions. *Management Science 42*(2), 157–172.

Robinson, L. W. (1990, March/April). Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Operations Research 38*(2), 278–295.

Rockafellar, R. T. (1996). *Convex Analysis*. Princeton University Press.

Ronnen, U. (1991). Minimum quality standards, fixed costs, and competition. *The RAND Journal of Economics 22*(4), 490–504.

Savin, S. V., M. A. Cohen, N. Gans, and Z. Katalan (2005, July/August). Capacity management in rental businesses with two customer bases. *Operations Research 53*(4), 617–631.

Shaked, A. and J. Sutton (1982). Relaxing price competition through product differentiation. *The Review of Economic Studies 49*(1), 3–13.

Shulman, J. D. and X. Geng (2012). Add-on pricing by asymmetric firms. *Management Science 59*(4), 899–917.

Shumsky, R. A. and F. Zhang (2009, May/June). Dynamic capacity management with substitution. *Operations Research 57*(3), 671–684.

Simchi-Levi, D., X. Chen, and J. Bramel (2014). *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management* (3 ed.). Springer.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics 69*(1), 99–118.

Steinhardt, C. and J. Gönsch (2012). Integrated revenue management approaches for capacity control with planned upgrades. *European Journal of Operational Research 223*(2), 380–391.

Su, X. (2007). Intertemporal pricing with strategic customer behavior. *Management Science 53*(5), 726–741.

Su, X. and F. Zhang (2008, October). Strategic customer behavior, commitment, and supply chain performance. *Management Science 54*(10), 1759–1773.

Su, X. and F. Zhang (2009, May). On the value of commitment and availability guarantees when selling to strategic consumers. *Management Science 55*(5), 713–726.

Talluri, K. T. and G. J. van Ryzin (2004a). Revenue management under a general discrete choice model of consumer behavior. *Management Science 50*(1), 15–33.

Talluri, K. T. and G. J. van Ryzin (2004b). *The Theory and Practice of Revenue Management*. New York: Springer-Verlag.

Topkis, D. M. (1968). Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Science 15*(3), 160–176.

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.

Van Mieghem, J. A. (1998). Investment strategies for flexible resources. *Management Science 44*(8), 1071–1078.

Van Mieghem, J. A. (2003, Fall). Capacity management, investment, and hedging: Review and recent developments. *M&SOM 5*(4), 269–302.

Van Mieghem, J. A. and N. Rudi (2002, Fall). Newsvendor networks: Inventory management and capacity investment with discretionary activities. *M&SOM 4*(4), 313–335.

Wall, M. (2003). Manufacturing flexibility. *Automotive Industries 183*(10), 44–45.

Xu, H., D. D. Yao, and S. Zheng (2011). Optimal control of replenishment and substitution in an inventory system with nonstationary batch demand. *Production and Operations Management 20*(5), 727–736.

Yao, D. D. and S. Zheng (2003). Substitutable inventory: Single- and multi-period models. In J. G. Shanthikumar, D. D. Yao, and Z. Henk (Eds.), *Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains*, Volume 63, pp. 177–201. Boston, MA: Kluwer Academic Publishers.

Yin, R., Y. Aviv, A. Pazgal, and C. S. Tang (2009). Optimal markdown pricing: Implications of inventory display formats in the presence of strategic customers. *Management Science 55*(8), 1391–1408.

Zhang, D. (2011, Winter). An improved dynamic programming decomposition approach for network revenue management. *M&SOM 13*(1), 35–52.

Zhang, D. and W. L. Cooper (2005, May/June). Revenue management for parallel flights with customer-choice behavior. *Operations Research 53*(3), 415–431.

Zorich, V. A. (2004). *Mathematical Analysis*, Volume 2. Springer.

APPENDICES

Appendix A

Appendices: Dynamic Capacity Management with Upgrading

Appendix A.1: Preliminary

A.1.1 Notations

The following notations are used in this appendix to simplify our exposition. Consider a vector $\mathbf{Z} = (z_1, \cdots, z_N) \in \Re^N$, for $1 \leq i < j \leq N$, we define

$$
\begin{aligned}
(\mathbf{Z})_i &= z_i \\
(\mathbf{Z})_{i,\cdots,j} &= (z_i, z_{i+1}, \cdots, z_j) \\
\mathbf{Z}_{ij} &= (z_1, \cdots, z_{i-1}, z_i + 1, z_{i+1}, \cdots, z_{j-1}, z_j - 1, z_{j+1}, \cdots, z_N).
\end{aligned}
$$

Notice that the above notations are still valid for $\mathbf{Z} = (z_r, \cdots, z_k)$ $(1 < r \leq i \leq j \leq k < N)$ if we artificially set $\mathbf{Z} = (0, \cdots, 0, z_r, \cdots, z_k, 0, \cdots, 0) \in \Re^N$.

For state vector $\mathbf{N}^t$, recall the effective state $\hat{\mathbf{N}}_r^t$ of classes $(1, \cdots, r)$ defined in Definition 2.5.1. If $r = N$, we use $\hat{\mathbf{N}}^t$ instead of $\hat{\mathbf{N}}_N^t$ to simplify our notation.

Moreover, for class i $(1 \leq i \leq N)$ in period t $(1 \leq t \leq T)$, we define

$$
\partial_i^- \Theta^t(\mathbf{Z}) = \frac{\partial}{\partial z_i^-} \Theta^t(\mathbf{Z}), \qquad \partial_i^+ \Theta^t(\mathbf{Z}) = \frac{\partial}{\partial z_i^+} \Theta^t(\mathbf{Z}).
$$

Recall Δ_{ij}^{-+} and Δ_{ij}^{+-} $(1 \leq i < j \leq N)$, we have

$$
\Delta_{ij}^{-+} \Theta^t(\mathbf{Z}) = \frac{\partial}{\partial z_i^-} \Theta^t(\mathbf{Z}) - \frac{\partial}{\partial z_j^+} \Theta^t(\mathbf{Z}), \qquad \Delta_{ij}^{+-} \Theta^t(\mathbf{Z}) = \frac{\partial}{\partial z_i^+} \Theta^t(\mathbf{Z}) - \frac{\partial}{\partial z_j^-} \Theta^t(\mathbf{Z}).
$$

Using the notations above, the protection level $p_{ij} = p$ in period t if and only if $\Delta_{ij}^{+-} \Theta^{t+1}(\mathbf{N}) \leq \alpha_{ij} \leq \Delta_{ij}^{-+} \Theta^t(\mathbf{N})$ from (2.8), where

$$
\mathbf{N} = (n_1^t, \cdots, n_{i-1}^t, p, 0, \cdots, 0, -p, n_{j+1}^t, \cdots, n_N^t).
$$

In the essence of Δ_{ij}^{-+} and Δ_{ij}^{+-}, we define the marginal perturbation of class i and j (referred to as MP_{ij} hereafter) as $\Theta^t(\mathbf{Z}+\epsilon(\mathbf{e}_i-\mathbf{e}_j))-\Theta^t(\mathbf{Z})$, where $\epsilon \in \Re$ is a small number and $\mathbf{e}_s$ $(s=i,j)$ is the unit vector with 1 in position s.

A.1.2 Independence Property

Consider a state vector $\mathbf{N}^t = (n_1^t, \cdots, n_N^t)$ and its effective state $\hat{\mathbf{N}}_{i-1}^t = (\hat{n}_1^t, \cdots, \hat{n}_{i-1}^t, n_i^t, \cdots, n_N^t)$ in period t. In Lemma 2.5.2 and A.2.1, we will show Θ^t has the following independence property if $(n_{i+1}^t, \cdots, n_{j-1}^t) \leq 0$ and $n_j^t < 0$:

1. In period t $(1 \leq t \leq T-1)$,

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t).$$

2. $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t)$ and $\Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t)$ are independent of the values of $(n_j^t, \cdots, n_N^t)$.

Given the independence property of Θ^{t+1}, the protection levels in period t have a similar property. Specifically, consider two different state vectors $\mathbf{N} = (n_1, \cdots, n_N)$ and $\mathbf{N}' = (n_1', \cdots, n_N')$ with the same effective state for the first $i-1$ classes. If $(n_{i+1}, \cdots, n_{j-1}) = (n_{i+1}', \cdots, n_{j-1}') \leq 0$ and $n_j = n_j' < 0$, then the protection level p_{ij} under state $\mathbf{N}$ is the same as that under $\mathbf{N}'$. Furthermore, the protection level p_{ij} under state $\mathbf{N}$ is independent of the values of $(n_j, \cdots, n_N)$. Hereafter, when speaking of the independence property, we do not distinguish between Θ^{t+1} and the protection levels in period t, since the proper interpretation is usually clear from the context.

Remark A.1.1 *Note that the independence property holds under the conditions $(n_{i+1}^t, \cdots, n_{j-1}^t) \leq 0$ and $n_j^t < 0$. However, in the proofs of Lemma 2.5.2 and A.2.1, we only need $n_j^t \leq 0$ to prove the results of $\Delta_{ij}^{+-}\Theta^{t+1}$.*

A.1.3 Foundation Results

Lemma 2.4.1 gives the condition of splitting the N-class general upgrading problem into subproblems, which reduces the complexity of the analysis.

Lemma 2.4.1 *Consider an N-class general upgrading problem with state $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)^\intercal$ in period t. If $\sum_{s=k}^{i} n_s^t \leq 0$ for all class $k \leq i$, then the problem can be separated into two independent subproblems: an upper part consisting of classes $(1, \cdots, i)$, and a lower part consisting of classes $(i+1, \cdots, N)$.*

Proof. This result holds if none of the optimal policies would upgrade demand j by capacity k when there remains unmet demand i $(k < i < j)$ in the same period. For simplicity, we only prove the latter claim in the integer case. For any demand sample path $(\mathbf{D}^t, \cdots, \mathbf{D}^T)$, let $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ be the optimal decisions. We assume without loss of generality that $y_{i-1,j}^t = (\mathbf{Y}^t)_{i-1,j} \geq 1$ $(i < j)$ while there remains unmet demand i after period t.

We construct decisions $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, T)$ that yield higher profit than the optimal decisions, which will be a contradiction. Let $\bar{\mathbf{Y}}^t$ be the same as $\mathbf{Y}^t$ except that $\bar{y}_{i-1,i}^t = y_{i-1,i}^t + 1$ and $\bar{y}_{i-1,j}^t = y_{i-1,j}^t - 1$. In the remaining periods s $(s = t+1, \cdots, T)$, we apply allocation decision $\bar{\mathbf{Y}}^s = \mathbf{Y}^s$ whenever $\mathbf{Y}^s$ is feasible. If the optimal decisions are feasible in periods $t+1$ to T, the profit increase by using $\bar{\mathbf{Y}}^s$ $(t \leq s \leq T)$ instead of the optimal decisions is $\alpha_{i-1,i} - \alpha_{i-1,j} + (T-t+1)(g_i - g_j) > 0$, which is a contradiction.

Otherwise, let l $(t+1 \leq l \leq T)$ be the first period that $\mathbf{Y}^l$ is not feasible. From our construction, it is clear that there exists $y_{ki}^l \geq 1$ $(k < i)$ in $\mathbf{Y}^l$ that is not feasible after applying $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, l-1)$. Let $\bar{\mathbf{Y}}^l$ be the same as $\mathbf{Y}^l$ except that $\bar{y}_{ki}^l = y_{ki}^l - 1$ and $\bar{y}_{kj}^l = y_{kj}^l + 1$. Since the states after applying $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, l)$ are the same as that for $\mathbf{Y}^s$ $(s = t, \cdots, l)$, $\bar{\mathbf{Y}}^s = \mathbf{Y}^s$ $(s = l+1, \cdots, T)$ are feasible in the remaining periods. Thus, the profit increase by using $\bar{\mathbf{Y}}^s$ $(t \leq s \leq T)$ instead of the optimal ones is $(l-t)(g_i - g_j) > 0$, which contradicts the optimality assumption.

This concludes our proof. $\square$

Lemmas A.1.1 and A.1.2 illustrate the bounds of the profit differences under different states.

Lemma A.1.1 *Consider a state vector* $\mathbf{N} = (n_1, \cdots, n_N)$ *with* $n_i \geq 0$ *and* $n_j \geq 0$ $(1 \leq i < j \leq N)$. *Then,*

$$\partial_i^+ \Theta^t(\mathbf{N}) - \partial_j^+ \Theta^t(\mathbf{N}) \geq u_j - u_i \tag{A.1}$$

and

$$\partial_i^- \Theta^t(\mathbf{N}) - \partial_j^- \Theta^t(\mathbf{N}) \geq u_j - u_i \qquad \text{if } n_i > 0 \text{ and } n_j > 0. \tag{A.2}$$

Proof. We use the sample path argument to prove (A.1). For each demand sample path, it is sufficient to prove

$$\Theta^t(\mathbf{N} + \epsilon \mathbf{e}_i) - \Theta^t(\mathbf{N} + \epsilon \mathbf{e}_j) \geq \epsilon(u_j - u_i), \tag{A.3}$$

where $\epsilon > 0$, $\mathbf{e}_s$ $(s = i, j)$ is the unit vector with 1 in position s. The same argument can be applied to (A.2).

Given a demand sample path $(\mathbf{D}^t, \cdots, \mathbf{D}^T)$, let $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ be the corresponding optimal solutions in period t to T under initial state $\mathbf{N} + \epsilon \mathbf{e}_j$ in period t. For initial state $\mathbf{N}+\epsilon\mathbf{e}_i$, we sequentially construct solutions $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$ based on $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ from period t to T. Specifically, $\bar{\mathbf{Y}}^l = \mathbf{Y}^l$ in period l $(t \leq l \leq T)$ if $\mathbf{Y}^l$ is feasible, and we write $\epsilon_l = 0$. Otherwise, if $\mathbf{Y}^l$ is not feasible, from the assumption of the initial states, the total demands which are satisfied by capacity j in $\mathbf{Y}^l$ is greater than the existing capacity j with initial state $\mathbf{N} + \epsilon \mathbf{e}_i$, and we denote the difference as ϵ_l $(0 < \epsilon_l \leq \epsilon_1)$. To construct a feasible solution $\bar{\mathbf{Y}}^l$, we use capacity i to satisfy demands which cannot be fulfilled by capacity j. By applying such $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$, the unmet demands in periods t to T are the same for both initial states, and $\sum_{l=t}^T \epsilon_l \leq \epsilon$.

Note that $\alpha_{si} - \alpha_{sj} = u_j - u_i < 0$ for any class s $(s \geq j)$, and unmet demand vectors in period t to T are the same for both initial states. Since $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$ are feasible solutions to the general upgrading problem with initial state $\mathbf{N} + \epsilon \mathbf{e}_i$, we have

$$\Theta^t(\mathbf{N} + \epsilon \mathbf{e}_i) - \Theta^t(\mathbf{N} + \epsilon \mathbf{e}_j) \geq (u_j - u_i) \sum_{l=t}^{T} \epsilon_l \geq \epsilon(u_j - u_i),$$

which completes the proof. $\square$

Lemma A.1.2 *Consider a state vector* $\mathbf{N} = (n_1, \cdots, n_N)$ *with* $n_i \leq 0$ *and* $n_j \leq 0$ $(1 \leq i < j \leq N)$. *Then,*

$$\partial_i^+ \Theta^t(\mathbf{N}) - \partial_j^+ \Theta^t(\mathbf{N}) \geq r_j - r_i \qquad \text{if } n_i < 0 \text{ and } n_j < 0$$

and

$$\partial_i^- \Theta^t(\mathbf{N}) - \partial_j^- \Theta^t(\mathbf{N}) \geq r_j - r_i.$$

Proof. It is similar to the proof of Lemma A.1.1. □

Appendix A.2: Proofs of the Main Results

This section presents the proofs of the main results in the chapter. The proofs of some intermediate results are lengthy and therefore presented in the Electronic Companion (§A.5), including Lemmas A.5.1 to A.5.5 and Propositions A.5.1 to A.5.3.

In §A.2.1, we prove the desired properties in period T. §A.2.2 considers a general period t by following the similar logic for period T. §A.2.3 completes the optimality proof. §A.2.4 proves two properties of the protection levels.

A.2.1 Final Period T

LEMMA 2.5.1 *The PSR algorithm solves the general upgrading problem* (2.2) *in period* T *with all protection levels being* 0.

Proof. Note that $\Theta^{T+1} \equiv 0$ and the solution $\mathbf{Y}^T$ generated by the PSR is a Monge sequence which solves the general upgrading problem in period T (see Bassok et al. 1999). □

We follow the notations in Chapter 2. Recall the state vector $\mathbf{N}^t = (n_1^t, \cdots, , n_N^t)$ in period t, and $\hat{\mathbf{N}}_{i-1}^t = (\hat{n}_1^t, \cdots, \hat{n}_{i-1}^t, n_i^t, \cdots, n_N^t)$, where $(\hat{n}_1^t, \cdots, \hat{n}_{i-1}^t)$ is the effective state of $(n_1^t, \cdots, n_{i-1}^t)$. Then, Lemma 2.5.2 shows the independence property of Θ^T.

LEMMA 2.5.2 *Consider an N-class general upgrading problem in period $T-1$ with state vector $\mathbf{N}^{T-1}$, where $(n_{i+1}^{T-1}, \cdots, n_{j-1}^{T-1}) \leq 0$ and $n_j^{T-1} < 0$. Then,*

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{+-}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}), \qquad \Delta_{ij}^{-+}\Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{-+}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}). \tag{A.4}$$

In addition, they are independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$.

Proof. For any $t = 1, \cdots, T$, given $\mathbf{D}^t = (d_1, \cdots, d_N)$ as realized demand in period t, we have

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}^{t-1}) = \Delta_{ij}^{+-}\,\mathbb{E}\left\{\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} = \mathbb{E}\left\{\Delta_{ij}^{+-}\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} \tag{A.5}$$

and

$$\Delta_{ij}^{-+}\Theta^t(\mathbf{N}^{t-1}) = \Delta_{ij}^{-+}\,\mathbb{E}\left\{\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} = \mathbb{E}\left\{\Delta_{ij}^{-+}\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\}. \tag{A.6}$$

Both the continuity of $\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)$ and the existence of its left and right derivatives (see Rockafellar 1996) assure the last equality in (A.5-A.6) (see Zorich 2004, P.409).

We focus on Δ_{ij}^{+-} in (A.4) since the same method applies to Δ_{ij}^{-+}. For any demand realization $\mathbf{D}^T = (d_1, \cdots, d_N)$ in period T, we next show

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}|\mathbf{D}^T) = \Delta_{ij}^{+-}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}|\mathbf{D}^T), \tag{A.7}$$

and it is independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$.

For any $\mathbf{D}^T$, without loss of generality, we assume classes $(1, \cdots, N)$ can not be separated based on $\mathbf{N}^{T-1} - \mathbf{D}^T$. Otherwise, from Lemma 2.4.1, we can consider independent subproblems instead. With this assumption, classes $(1, \cdots, N)$ are also not separable based on $\hat{\mathbf{N}}_{i-1}^{T-1} - \mathbf{D}^T$ by Proposition A.5.1 given in the Supplementary Appendix (§A.5).

To solve the N-class general upgrading problem in period T, we first solve subproblems $(1, \cdots, i-1)$ with initial state $(\mathbf{N}^{T-1})_{1,\cdots,i-1}$ and $(\hat{\mathbf{N}}_{i-1}^{T-1})_{1,\cdots,i-1}$ by the PSR. Then, we use the PSR to solve the subproblem $(1, \cdots, N)$, where the initial states of classes $(1, \cdots, i-1)$ are the states after solving the subproblem $(1, \cdots, i-1)$ by the PSR.

Since the upgrading problem in period T is a transportation problem, given the special cost structure, the optimal allocation decisions in subproblem $(1, \cdots, i-1)$ are independent from classes $(i, \cdots, N)$. Particularly, the optimal decisions within classes $(1, \cdots, i-1)$ remain unchanged with respect to MP_{ij}. Moreover, from Proposition A.5.2, the result of applying the PSR to subproblem $(1, \cdots, i-1)$ with initial state $(\mathbf{N}^{T-1})_{1,\cdots,i-1}$ is the same as that with initial state $(\hat{\mathbf{N}}_{i-1}^{T-1})_{1,\cdots,i-1}$. In other words, the initial states in subproblem $(1, \cdots, N)$ are the same for both initial states $(\mathbf{N}^{T-1}, \mathbf{D}^T)$ and $(\hat{\mathbf{N}}_{i-1}^{T-1}, \mathbf{D}^T)$. Thus, (A.7) is true. In addition, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}|\mathbf{D}^T)$ is independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$ from Lemma A.5.2. This completes the proof. $\square$

A.2.2 Earlier Periods

Lemma A.2.1 proves the independence property of Θ^{t+1} by backward induction.

Lemma A.2.1 *Consider an N-class general upgrading problem in period t with state vector $\mathbf{N}^t$, where $(n_{i+1}^t, \cdots, n_{j-1}^t) \leq 0$ and $n_j^t < 0$. If the PSR algorithm solves the general upgrading problem in period $t+1$ and the independence property holds for Θ^{t+2}, then,*

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t). \tag{A.8}$$

In addition, $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t)$ and $\Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t)$ are independent of the values of $(n_j^t, \cdots, n_N^t)$.

Proof. As discussed in the proof of Lemma 2.5.2, Δ_{ij}^{+-} and Δ_{ij}^{-+} in (A.8) are well-defined. We prove the equality regarding Δ_{ij}^{+-} in (A.8) and the corresponding independence property for any demand realization $\mathbf{D}^{t+1} = (d_1, \cdots, d_N)$ in period $t+1$. From Lemma 2.4.1, we can assume classes $(1, \cdots, N)$ are not separable under $\mathbf{N}^t - \mathbf{D}^{t+1}$, which is also true under $\hat{\mathbf{N}}_{i-1}^t - \mathbf{D}^{t+1}$ by Proposition A.5.1.

Splitting the N-class general upgrading problem into subproblems: $(1, \cdots, i-1)$, $(1, \cdots, j)$ and $(1, \cdots, N)$, we start with the subproblem $(1, \cdots, i-1)$.

1. Because the protection levels within classes $(1, \cdots, i-1)$ in period $t+1$ are defined by Θ^{t+2}, which satisfies the independence property by assumption, the allocation decisions within classes $(1, \cdots, i-1)$ in period $t+1$ remain unchanged with respect to MP_{ij}. Let $\mathbf{N}'_{i-1}$ be the outcome of applying the PSR algorithm to subproblem $(1, \cdots, i-1)$ with states $((\mathbf{N}^t)_{1,\cdots,i-1}, (\mathbf{D}^{t+1})_{1,\cdots,i-1})$. Denote k $(1 \le k \le i-1)$ as the highest class such that $(\mathbf{N}'_{i-1})_{k,\cdots,i-1} \ge 0$ and $(\mathbf{N}'_{i-1})_k > 0$. Since the PSR is optimal in period $t+1$ by assumption, we only need to consider upgrading decisions among classes $(k, \cdots, N)$ in the rest of the subproblems. Similarly, we can define $\hat{\mathbf{N}}'_{i-1}$ and $\hat{k}$ for subproblem $(1, \cdots, i-1)$ with states $((\hat{\mathbf{N}}^t_{i-1})_{1,\cdots,i-1}, (\mathbf{D}^{t+1})_{1,\cdots,i-1})$. From Proposition A.5.3, we know that $\hat{k} = k$ and $(\hat{\mathbf{N}}'_{i-1})_{k,\cdots,i-1} = (\mathbf{N}'_{i-1})_{k,\cdots,i-1}$. In other words, after solving subproblem $(1, \cdots, i-1)$, the initial state of classes $(k, \cdots, N)$ are the same for both $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t_{i-1}$. Notice that we assume both k and $\hat{k}$ exist; otherwise, both k and $\hat{k}$ do not exist from Proposition A.5.3, which means that considering upgrading decisions in classes $(i, \cdots, N)$ is sufficient, which is a simpler case.

2. From the definition of the protection levels, although there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$, the states of classes $(1, \cdots, k-1)$ can still affect the protection levels within classes $(k, \cdots, N)$ in period $t+1$. Fortunately, the effective state of $(\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1}$ is the same as that of $(\mathbf{N}'_{i-1})_{1,\cdots,k-1}$ by Proposition A.5.2. From the independence property assumption of Θ^{t+2}, the protection levels within classes $(k, \cdots, N)$ are the same for both initial states.

To summarize, for initial states $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t_{i-1}$, the capacities of classes $(k, \cdots, i-1)$ after solving subproblem $(1, \cdots, i-1)$, which can upgrade the demands in classes $(i, \cdots, N)$, are the same. Moreover, the protection levels within classes $(k, \cdots, N)$ are also the same. Therefore, we only analyze the allocation decisions within classes $(k, \cdots, N)$ under initial state $\mathbf{N}^t$, which can again be split into subproblems $(k, \cdots, j)$ and $(k, \cdots, N)$.

Apply the PSR to subproblem $(k, \cdots, j)$ with state $(\mathbf{N}_j, (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}))$, where $\mathbf{N}_j = ((\mathbf{N}'_{i-1})_{k,\cdots,i-1}, (\mathbf{N}^t)_{i,\cdots,j})$, and let $\mathbf{N}'_j$ be the resulting states of classes

$(k,\cdots,j)$ after applying $\mathbf{Y}_j$, which are the optimal allocation decisions within classes $(k,\cdots,j)$. Since $(\mathbf{N}^t)_{i+1,\cdots,j} \le 0$, the protection levels used in subproblem $(k,\cdots,j)$, which determine the upgrades from classes $(k,\cdots,i)$, only depend on $(\mathbf{N}'_{i-1})_{k,\cdots,i-1}$ by the independence property assumption of Θ^{t+2}. We consider two cases based on whether there is unmet demand j in $\mathbf{N}'_j$:

1. $(\mathbf{N}'_j)_j = 0$: Define h $(k \le h \le i)$ as the class which satisfies the last unit of demand j when the PSR solves subproblem $(k,\cdots,j)$. In fact,

$$h = \begin{cases} r, & \text{if } r < i \text{ and } \sum_{s=r+1}^{i-1}(\mathbf{N}'_{i-1})_s \le -\sum_{s=i}^{j}((\mathbf{N}^t)_s - d_s) < \sum_{s=r}^{i-1}(\mathbf{N}'_{i-1})_s \\ i, & \text{if } \sum_{s=i}^{j}((\mathbf{N}^t)_s - d_s) > 0. \end{cases}$$

   In this case, $\mathbf{N}'_j$ is the same as the result of applying the greedy upgrading to subproblem $(k,\cdots,j)$, i.e., $\mathbf{N}'_j = \hat{\mathbf{N}}_j$, where $\hat{\mathbf{N}}_j$ is the effective state of $\mathbf{N}_j$. Specifically,

$$(\hat{\mathbf{N}}_j)_l = \begin{cases} (\mathbf{N}'_{i-1})_l, & \text{if } k \le l < h \\ \sum_{s=h}^{i-1}(\mathbf{N}'_{i-1})_s + \sum_{s=i}^{j}((\mathbf{N}^t)_s - d_s), & \text{if } l = h < i \\ \sum_{s=i}^{j}((\mathbf{N}^t)_s - d_s), & \text{if } l = h = i \\ 0, & \text{otherwise,} \end{cases} \tag{A.9}$$

   for class l $(k \le l \le j)$. Note that class h $(k \le h \le i)$ must exist since classes $(1,\cdots,N)$ are not separable, and h and $\hat{\mathbf{N}}_j$ remain the same with respect to MP_{ij}. Furthermore, from the discussion of $\mathbf{N}'_j$, we can see that $\mathbf{Y}_j$ is the same as optimal allocation decisions given initial state $(\mathbf{N}_j, (0,\cdots,0,(\mathbf{D}^{t+1})_{i,\cdots,j}))$ in period T where the protection levels are zero. Hence,

$$\begin{aligned} &\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1}) \\ =&\Theta^T\left((\mathbf{N}^t)_{1,\cdots,i-1} - \mathbf{N}'_{i-1}|(\mathbf{D}^{t+1})_{1,\cdots,i-1}\right) + \Theta^T\left(\mathbf{N}_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{i,\cdots,j})\right) \\ &+\Theta^{t+1}\left(((\mathbf{N}'_{i-1})_{1,\cdots,k-1}, \hat{\mathbf{N}}_j, (\mathbf{N}^t)_{j+1,\cdots,N})|(0,\cdots,0,(\mathbf{D}^{t+1})_{j+1,\cdots,N})\right), \end{aligned} \tag{A.10}$$

where the first two terms are the corresponding revenues of subproblems $(1, \cdots, i-1)$ and $(k, \cdots, j)$, and the last term is the sum of the current revenue of subproblem $(k, \cdots, j)$ and the expected value in the remaining periods. Thus,

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1}) = \Delta_{ij}^{+-}\Theta^T\left(\mathbf{N}_j|(0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j})\right), \tag{A.11}$$

which is clearly independent of $(n_{j+1}^t, \cdots, n_N^t)$. Also, (A.11) is independent of n_j^t by Lemma A.5.2. Note that the first term in (A.10) has been omitted from (A.11) since the allocation decisions in subproblem $(1, \cdots, i-1)$ remain unchanged with respect to MP_{ij}. Moreover, the last term in (A.10) has also been dropped from (A.11) because its initial states remain the same with respect to MP_{ij}.

Similarly, for initial state $\hat{\mathbf{N}}_{i-1}^t$, we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t|\mathbf{D}^{t+1}) = \Delta_{ij}^{+-}\Theta^T\left(\mathbf{N}_j|(0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j})\right)$$

since the allocation decisions in subproblem $(k, \cdots, j)$ are the same for both initial state $\mathbf{N}^t$ and $\hat{\mathbf{N}}_{i-1}^t$. Therefore, we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1}) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t|\mathbf{D}^{t+1}),$$

which is independent of the values of $(n_j^t, \cdots, n_N^t)$ by (A.11);

2. $(\mathbf{N}_j')_j < 0$: Since the PSR is optimal in period $t+1$, there is no upgrade between classes $(k, \cdots, j)$ and $(j+1, \cdots, N)$. By the definition of the effective state, $\hat{\mathbf{N}}_j$ in (A.9), which remains unchanged with respect to MP_{ij}, is the effective state of $\mathbf{N}_j'$. Thus, the allocation decisions within classes $(j+1, \cdots, N)$ stay the same with respect to MP_{ij} by the independence property assumption of Θ^{t+2},

and we denote $\mathbf{N}'_{j+}$ as the result of applying the PSR to classes $(j+1,\cdots,N)$. Therefore, we have

$$\begin{aligned}
&\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1})\\
=&\Theta^T\left((\mathbf{N}^t)_{1,\cdots,i-1}-\mathbf{N}'_{i-1}|(\mathbf{D}^{t+1})_{1,\cdots,i-1}\right)+\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)\\
&+\Theta^T\left((\mathbf{N}^t)_{j+1,\cdots,N}-\mathbf{N}'_{j+}|(0,\cdots,0,(\mathbf{D}^{t+1})_{j+1,\cdots,N})\right)\\
&+\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{A.12}$$

where the first three terms are the corresponding revenues of subproblems $(1,\cdots,i-1)$, $(k,\cdots,j)$, and $(j,\cdots,N)$, and the last term is the expected revenue-to-go function. As we discussed earlier, we have

$$\begin{aligned}
&\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1})\\
=&\Delta_{ij}^{+-}\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)+\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
&-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{A.13}$$

where the first term is independent of $(n_{j+1}^t,\cdots,n_N^t)$ by construction. Moreover, recall that the protection levels used in subproblem $(k,\cdots,j)$ only depend on $(\mathbf{N}'_{i-1})_{k,\cdots,i-1}$, and demand j is not fully satisfied in this case, thus the allocation decisions $\mathbf{Y}_j$ as well as $\mathbf{N}_j-\mathbf{N}'_j$, which is the capacity used in subproblem $(k,\cdots,j)$, do not depend on n_j^t. Hence, the first term in (A.13) is also independent of n_j^t. Similarly, for initial state $\hat{\mathbf{N}}_{i-1}^t$, we have

$$\begin{aligned}
&\Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t|\mathbf{D}^{t+1})\\
=&\Delta_{ij}^{+-}\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)+\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
&-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right).
\end{aligned} \tag{A.14}$$

To complete the proof, from (A.13) and (A.14), we use the induction assumption of Θ^{t+2} to show

$$\begin{aligned}
&\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
=&\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{A.15}$$

which is independent of $(n_j^t,\cdots,n_N^t)$. First of all, since there is no upgrade between classes $(1,\cdots,k-1)$ and $(k,\cdots,N)$ in period $t+1$, and the PSR sequentially satisfies demands in each class, the marginal change of n_i^t only affects the state of a single class in $\mathbf{N}'_j$, which is the same for both initial states $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t_{i-1}$. Denote such a class as r, then $k \leq r \leq j$. Given $(\mathbf{N}'_{i-1})_{1,\cdots,k-1}$ and $(\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1}$ have the same effective state from the previous argument, to apply the induction assumption, we only need to show $(\mathbf{N}'_j)_{r+1,\cdots,j} \leq 0$ where $(\mathbf{N}'_j)_j < 0$ by assumption. Suppose to the contrary that $(\mathbf{N}'_j)_l > 0$ for class l $(r < l < j)$. Note that initial states $(\mathbf{N}^t)_{i+1,\cdots,j} \leq 0$, thus class $l \leq i$. Since the demands in classes $(i,\cdots,j)$ should be satisfied by class l prior to class r by the PSR, given $(\mathbf{N}'_j)_l > 0$, there is no upgrade between classes $(k,\cdots,l-1)$ and $(l,\cdots,j)$, i.e., the marginal change of n_i^t should not affect the state of class r, which is a contradiction. Hence, by applying the induction assumption to (A.15), we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1}) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}^t_{i-1}|\mathbf{D}^{t+1}),$$

which is independent of the values of $(n_j^t,\cdots,n_N^t)$. This concludes the proof.

□

A.2.3 Optimality

Proposition 2.5.1 *1. The PSR algorithm solves the general upgrading problem in period t;*

2. For a state vector $\mathbf{N}^t$ with $(n_{i+1}^t, \cdots, n_{j-1}^t) \le 0$ and $n_j^t < 0$, we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t).$$

In addition, they are independent of the values of $(n_j^t, \cdots, n_N^t)$.

Proof. In the proof, we show the two properties in Proposition 2.5.1 can be preserved under backward induction. The proof of period T is given in the end of this proof.

Suppose they are true for Θ^{t+1}, we verify the two properties for Θ^t.

1. Optimality of the PSR algorithm

Consider initial state $\mathbf{X}^t = (x_1, \cdots, x_N)$ and $\tilde{\mathbf{D}}^t = (\tilde{d}_1, \cdots, \tilde{d}_N)$, for any demand realization $\mathbf{D}^t = (d_1, \cdots, d_N)$ in period t, we next verify $\mathbf{Y}^t = (y_{ij})_{N\times N}$ derived by the PSR are optimal in period t. First, from the discussion in Chapter 2, $y_{kk} = \min(d_k + \tilde{d}_k, x_k)$ $(1 \le k \le N)$ in the PSR is optimal.

For upgrading decisions y_{ij} $(i > j)$ in $\mathbf{Y}^t$, we consider an equivalent representation of the general upgrading problem in (2.2). Let $\mathbf{Z} = (z_1, \cdots, z_N)^\top = \mathbf{Y}^t\mathbf{1}$, the optimal solution $\mathbf{W} = (w_{ij})_{N\times N}$ in the following linear program is the same as $\mathbf{Y}^t = (y_{ij})_{N\times N}$ in period t:

$$\begin{aligned} \max_{\mathbf{W}\ge 0} \quad & \sum_{1\le i\le j\le N} \alpha_{ij} w_{ij} \\ \text{s.t.} \quad & \sum_j w_{ij} \le z_i, \quad i = 1, 2, \cdots, N, \\ & \sum_i w_{ij} \le d_j + \tilde{d}_j, \quad j = 1, 2, \cdots, N. \end{aligned} \tag{A.16}$$

Since the parallel allocation is optimal, $z_i = x_i$ $(1 \le i \le N)$ in $\mathbf{Z}$ is optimal if $x_i \le d_i + \tilde{d}_i$. Furthermore, we need to show the optimality of z_i for all classes i's with $x_i > d_i + \tilde{d}_i$, i.e., the classes with surplus capacities after the parallel allocation. Since the general upgrading problem is concave, we only need to examine $\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N})$ and $\frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N})$, where

$$\mathbf{N} = \mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t - \mathbf{Y}^t\mathbf{1} + (\mathbf{Y}^t)^\intercal\mathbf{1} = \mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t - \mathbf{Z} + (\mathbf{W})^\intercal\mathbf{1}$$

is the state at the beginning of period $t+1$.

Without loss of generality, we assume class 1 is the highest class with $x_1 > d_1 + \tilde{d}_1$ and analyze the optimality of z_1 by cases.

1. $z_1 = d_1 + \tilde{d}_1$: We only need to prove that increasing z_1 is suboptimal since $z_1 \geq y_{11} = d_1 + \tilde{d}_1$. Let k ($k > 1$) be the highest class with $(\mathbf{N})_k < 0$. Note that z_1 is clearly optimal if class k does not exist, i.e., there is no backlogged demand in classes $(1, \cdots, N)$ in $\mathbf{N}$.

   (a) $(\mathbf{N})_{2,\cdots,k-1} = 0$: When solving the protection level p_{1k} and the allocation decision y_{1k} by the PSR, $(\mathbf{N})_{1,\cdots,k}$ are the states of classes $(1, \cdots, k)$. Meanwhile, the upgrading decisions within classes $(k+1, \cdots, N)$ have not been considered, whose states are the states after the parallel allocation, i.e., $(\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{k+1,\cdots,N}$. Thus,

$$\begin{aligned} 0 \geq & \alpha_{1k} - \Delta_{1k}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,k}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{k+1,\cdots,N}\right) \\ = & \alpha_{1k} - \Delta_{1k}^{-+}\Theta^{t+1}(\mathbf{N}) = \alpha_{1k} + \frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}), \end{aligned} \tag{A.17}$$

   where the first equality is from the independence property assumption of Θ^{t+1}, and the second equality follows from the fact that $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_1 + \mathbf{e}_k)$ when z_1 marginally changes to $z_1 + \epsilon$, where $\epsilon > 0$. Hence, increasing z_1 is suboptimal.

   (b) There exists class i ($1 < i < k$) with $(\mathbf{N})_i > 0$: Without loss of generality, we assume that i is the lowest class in $(2, \cdots, k-1)$ with $(\mathbf{N})_i > 0$. In this case, the PSR considers protection level p_{ik} and ignores the potential upgrade from class 1 to k, and we will show it is indeed optimal to do so. Since $(\mathbf{N})_{1,\cdots,k}$ are the states of classes $(1, \cdots, k)$ when considering the protection level p_{ik} by the PSR, and $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_i + \mathbf{e}_k)$ when z_i marginally changes to $z_i + \epsilon$, where $\epsilon > 0$. We have

$$0 \geq \alpha_{ik} - \Delta_{ik}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,k}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{k+1,\cdots,N}\right) = \alpha_{ik} + \frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N}).$$

Moreover, because $(\mathbf{N})_1 > 0$ and $(\mathbf{N})_i > 0$,

$$\partial_1^- \Theta^{t+1}(\mathbf{N}) - \partial_i^- \Theta^{t+1}(\mathbf{N}) \geq u_i - u_1$$

from Lemma A.1.1.

Notice that $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_1 + \mathbf{e}_k)$ when z_1 marginally changes to $z_1 + \epsilon$, then

$$\frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}) - \frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N}) = \partial_i^- \Theta^{t+1}(\mathbf{N}) - \partial_1^- \Theta^{t+1}(\mathbf{N}). \tag{A.18}$$

Thus, from $\alpha_{ik} - \alpha_{1k} = u_1 - u_i$, we have

$$\frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}) + \alpha_{1k} \leq \frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N}) + \alpha_{1k} + u_1 - u_i = \frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N}) + \alpha_{ik} \leq 0, \tag{A.19}$$

which means increasing z_1 is not optimal.

2. $z_1 > d_1 + \tilde{d}_1$: Let j $(j > 1)$ be the lowest class with $y_{1j} > 0$ in $\mathbf{Y}^t$. Similar to the previous case, from the PSR, $(\mathbf{N})_{1,\cdots,j}$ are the states after performing the last unit of upgrade y_{1j}. In this case, $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(\mathbf{e}_1 - \mathbf{e}_j)$ when z_1 marginally changes to $z_1 - \epsilon$, where $\epsilon > 0$, then

$$0 \leq \alpha_{1j} - \Delta_{1j}^{+-}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,j}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{j+1,\cdots,N}\right) = \alpha_{1j} + \frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}). \tag{A.20}$$

Thus, decreasing current z_1 is costly.

Furthermore, for all class i $(1 < i < j)$ with $x_i > d_i + \tilde{d}_i$, $z_i = x_i$ by the PSR algorithm. First, we only need to show decreasing these z_i's is not optimal.

When z_i marginally changes to $z_i - \epsilon$, there is a *chain reaction*. From (A.16), decreasing z_i by ϵ is equivalent to reducing the upgrade y_{ik_i} by ϵ, where k_i is the lowest class upgraded by capacity i. Then, unmet demand k_i increases by ϵ unit, and demand k_i will be upgraded by capacity s $(1 \leq s < i)$, which is the lowest class with $x_s > \tilde{d}_s + d_s$. Meanwhile, k_s, the lowest class of demands upgraded by capacity s prior to changing z_i, has an additional ϵ unit unmet demand, which can be similarly analyzed as class k_i. The *chain reaction* continues, and $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(\mathbf{e}_i - \mathbf{e}_j)$, i.e., the unmet demand j is increased by ϵ unit.

When z_i marginally changes to $z_i - \epsilon$, only unmet demand j and capacity i changed in the aforementioned *chain reaction*, then the objective function in (A.16) decreases by $\epsilon\alpha_{ij}$. Meanwhile, given $(\mathbf{N})_1 \geq 0$ and $(\mathbf{N})_i = 0$, similar to (A.18), we have

$$\frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}) - \frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N}) \leq u_1 - u_i$$

by Lemma A.1.1. Thus, from $\alpha_{ij} - \alpha_{1j} = u_1 - u_i$,

$$\frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N}) + \alpha_{ij} \geq \frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}) + u_i - u_1 + \alpha_{ij} = \frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}) + \alpha_{1j} \geq 0.$$

Hence, $z_i = x_i$ is optimal for all class i $(1 < i < j)$ with $x_i > d_i + \tilde{d}_i$.

Next, we have to prove that increasing z_1 itself is also suboptimal.

(a) $(\mathbf{N})_j < 0$: In this case, the protection level p_{1j} is binding in the PSR, i.e., the upgrade between classes 1 and j stops when the quantity of capacity 1 reaches p_{1j}. From the definition of p_{1j}, and the fact that $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_1 + \mathbf{e}_j)$ when z_1 marginally changes to $z_1 + \epsilon$, we have

$$\begin{aligned} 0 \geq &\alpha_{1j} - \Delta_{1j}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,j}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{j+1,\cdots,N}\right) \\ = &\alpha_{1j} + \frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}). \end{aligned} \tag{A.21}$$

From (A.20) and (A.21), we know the optimality of z_1.

(b) $(\mathbf{N})_j = 0$: The upgrading decision y_{1j} is bounded because there is no unmet demand j remaining, and we do not have (A.21) directly from solving p_{1j}. However, similar to the case when $z_1 = d_1 + \tilde{d}_1$, increasing z_1 is still suboptimal. Particularly, if there exists k $(k > j)$ as the highest class with $(\mathbf{N})_k < 0$, and $(\mathbf{N})_s = 0$ for all class s $(j < s < k)$, then we have (A.17) that affirms the optimality of z_1. On the other hand, if there exists class i $(j < i < k)$ with $(\mathbf{N})_i > 0$, then (A.19) is valid, which also proves the optimality of z_1.

To summarize, we have proved that z_1 is optimal. In addition, if $z_1 > d_1 + \tilde{d}_1$ and class j is the lowest class with $y_{1j} > 0$ in $\mathbf{Y}^t$, we have also shown the optimality of

z_i $(1 < i < j)$ with $x_i > d_i + \tilde{d}_i$. The same argument can be sequentially applied to the rest of z_s's since $(\mathbf{N})_{1,\cdots,s-1}$ are the states of classes $(1, \cdots, s-1)$ when solving the protection levels within classes $(s, \cdots, N)$ in the PSR algorithm.

Therefore, the PSR algorithm solves the general upgrading problem in period t.

2. Independence property of Θ^t

As the PSR solves the general upgrading problem in period t, and the independence property of Θ^{t+1} holds by Property 2 of the induction assumption, all requirements of Lemma A.2.1 are satisfied, thus the independence property of Θ^t also holds.

To conclude the proof, we now consider period T. The PSR solves the general upgrading problem in period T by Lemma 2.5.1. And Lemma 2.5.2 asserts the independence property of Θ^T. Therefore, we can use the backward induction and complete the proof. $\square$

A.2.4 Properties of the Protection Levels

Proposition 2.5.2 *If initial capacity $\mathbf{X}^1$ and demand $\mathbf{D}^1, \cdots, \mathbf{D}^T$ are integer valued, there exists an integer valued optimal policy $\mathbf{Y}^1, \cdots, \mathbf{Y}^T$ derived by the PSR algorithm.*

Proof. The proof is similar to the proof of Proposition 3 in Shumsky and Zhang (2009). $\square$

Proposition 2.5.3 *For the same $(n_1^t, \cdots, n_{i-1}^t)$ in period t $(1 \le t \le T)$, $p_{ij} \le p_{i,j+1}$ when $i < j$.*

Proof. Suppose to the contrary that $p_{ij} > p_{i,j+1}$ in period t. Let $\bar{p} = \frac{p_{ij}+p_{i,j+1}}{2}$, and denote $\mathbf{N} = (n_1^t, \cdots, n_{i-1}^t, \bar{p}, 0, \cdots, 0, n_{j+2}^t, \cdots, n_N^t)$. From the concavity in Proposition 2.4.1 and the independence property in Proposition 2.5.1, we have $\Delta_{i,j+1}^{+-}\Theta^{t+1}(\mathbf{N}) \le \alpha_{i,j+1}$ given $\bar{p} > p_{i,j+1}$. Similarly, we have $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}) \ge \alpha_{ij}$ since $\bar{p} < p_{ij}$.

However, since $\alpha_{ij} - \alpha_{i,j+1} = r_j + g_j - r_{j+1} - g_{j+1}$ and $g_j > g_{j+1}$,

$$\begin{aligned}
&\Delta^{+-}_{i,j+1}\Theta^{t+1}(\mathbf{N}) = \partial^+_i\Theta^{t+1}(\mathbf{N}) - \partial^-_{j+1}\Theta^{t+1}(\mathbf{N}) \\
\geq &\partial^+_i\Theta^{t+1}(\mathbf{N}) - \partial^-_j\Theta^{t+1}(\mathbf{N}) + r_{j+1} - r_j = \Delta^{+-}_{ij}\Theta^{t+1}(\mathbf{N}) + r_{j+1} - r_j \\
\geq &\alpha_{ij} + r_{j+1} - r_j > \alpha_{i,j+1},
\end{aligned}$$

where the first inequality follows from Lemma A.1.2. This is a contradiction. Hence, $p_{ij} \leq p_{i,j+1}$ when $i < j$. □

Appendix A.3: Multi-Horizon Model with Replenishment

Proposition 2.6.1 *Suppose the firm starts with an initial capacity $\mathbf{X} \leq \mathbf{X}^*$. The firm's optimal replenishment policy in the multi-horizon model is a base stock policy with the optimal base stock level $\mathbf{X}^*$ in* (2.11). *Furthermore, the PSR algorithm solves the optimal allocation decisions within each horizon.*

Proof. We prove this proposition by induction. Let $V_k(\mathbf{X}, \tilde{\mathbf{D}})$ $(1 \leq k \leq K)$ denote the expected revenue-to-go function at the beginning of the k-th horizon with capacity $\mathbf{X}$ and backlogged demand $\tilde{\mathbf{D}}$. Where possible, the index of periods in each horizon is denoted by superscripts while subscripts denote the index of horizons. We inductively prove the following three properties.

1. The PSR algorithm optimally solves the capacity allocation decisions in the k-th horizon;
2. The optimal replenishment policy in the k-th horizon is a base stock policy with the optimal base stock level $\mathbf{X}^*$;
3. If $\mathbf{X} \leq \mathbf{X}^*$, $V_k(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\alpha - \mathbf{C}$.

Suppose all properties hold in the $(k+1)$-th horizon. It suffices to consider capacity $\mathbf{X} \le \mathbf{X}^*$. Since $V_{k+1}(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $(\mathbf{X}, \tilde{\mathbf{D}})$, in horizon k we have

$$
\begin{aligned}
&\Theta_k^T(\mathbf{X}, \tilde{\mathbf{D}}) \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) - \mathbf{h}\mathbf{X}_k^{T+1} + \gamma V_{k+1}(\mathbf{X}_k^{T+1}, \tilde{\mathbf{D}}_k^{T+1}) \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) - \mathbf{h}\mathbf{X}_k^{T+1} + \gamma V_{k+1}(\mathbf{X}^*, 0) + \gamma \mathbf{C}(\mathbf{X}_k^{T+1} - \mathbf{X}^*) + \gamma(\alpha - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) + (\gamma \mathbf{C} - \mathbf{h})\mathbf{X}_k^{T+1} + \gamma(\alpha - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} + \gamma \left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^* \right) \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) + (\gamma \mathbf{C} - \mathbf{h})\mathbf{X}_k^{T+1} + \gamma(\alpha - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} \right] \right\} + \gamma \left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^* \right),
\end{aligned}
\tag{A.22}
$$

where $\gamma \left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^* \right)$ is a constant. From the proof of Proposition 2.5.1, the PSR algorithm optimally solves the T-th period allocation decisions in the k-th horizon, where the protection levels are based on $\Theta^{T+1}(\mathbf{X}^{T+1}, \tilde{\mathbf{D}}^{T+1})$ in (2.10). Moreover, it is clear that $\Theta_k^T(\mathbf{X}, \tilde{\mathbf{D}})$ is also concave in $(\mathbf{X}, \tilde{\mathbf{D}})$ from the proof of Proposition 2.4.1. Inductively, for $t = T-1, \cdots, 1$, we know

$$
\Theta_k^t(\mathbf{X}, \tilde{\mathbf{D}}) = \mathop{\mathbb{E}}_{\mathbf{D}_k^t} \left\{ \max_{\mathbf{Y}_k^t} \left[H(\mathbf{Y}_k^t | \tilde{\mathbf{D}}; \mathbf{D}_k^t) + \Theta_k^{t+1}(\mathbf{X}_k^{t+1}, \tilde{\mathbf{D}}_k^{t+1}) \right] \right\}
$$

is concave in $(\mathbf{X}, \tilde{\mathbf{D}})$, and we can show that the PSR algorithm solves the capacity allocation decisions for horizon k.

From the Bellman equation, we have

$$
V_k(\mathbf{X}, \tilde{\mathbf{D}}) = \max_{\mathbf{Z} \ge \mathbf{X}} \; G(\mathbf{Z}),
$$

where $G(\mathbf{Z}) = \Theta_k^1(\mathbf{Z}, \mathbf{0}) + \alpha \tilde{\mathbf{D}} - \mathbf{C}(\mathbf{Z} - \mathbf{X} + \tilde{\mathbf{D}})$. From (A.22), we have

$$
G(\mathbf{Z}) = \Pi \left(\mathbf{Z}; \gamma \mathbf{C} - \mathbf{h}, \gamma(\alpha - \mathbf{C}) \right) + (\alpha - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}(\mathbf{X} - \gamma \mathbf{X}^*) + \gamma V_{k+1}(\mathbf{X}^*, 0).
$$

By the definition of $\mathbf{X}^*$, the optimal replenishment policy in the k-th horizon is a base stock policy with optimal base stock level $\mathbf{X}^*$. Furthermore, for $\mathbf{X} \le \mathbf{X}^*$,

$$
V_k(\mathbf{X}, \tilde{\mathbf{D}}) = \Pi \left(\mathbf{X}^*; \gamma \mathbf{C} - \mathbf{h}, \gamma(\alpha - \mathbf{C}) \right) + (\alpha - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}(\mathbf{X} - \gamma \mathbf{X}^*) + \gamma V_{k+1}(\mathbf{X}^*, 0)
$$

is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\alpha - \mathbf{C}$.

To conclude the proof, we consider the last horizon profit, $V_K(\mathbf{X}, \tilde{\mathbf{D}})$. Since

$$\Theta_K^T(\mathbf{X}, \tilde{\mathbf{D}}) = \mathop{\mathbb{E}}_{\mathbf{D}_K^T} \left\{ \max_{\mathbf{Y}_K^T} \left[H(\mathbf{Y}_K^T | \tilde{\mathbf{D}}; \mathbf{D}_K^T) - \mathbf{h}\mathbf{X}_K^{T+1} + \gamma \mathbf{C}\mathbf{X}_K^{T+1} + \gamma(\alpha - \mathbf{C})\tilde{\mathbf{D}}_K^{T+1} \right] \right\}$$

by definition, the optimality of the PSR algorithm can be similarly proved. Meanwhile, if $\mathbf{X} \le \mathbf{X}^*$,

$$V_K(\mathbf{X}, \tilde{\mathbf{D}}) = \max_{\mathbf{Z} \ge \mathbf{X}} \left[\Pi\left(\mathbf{Z}; \gamma\mathbf{C} - \mathbf{h}, \gamma(\alpha - \mathbf{C})\right) + (\alpha - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}\mathbf{X} \right],$$

so the base stock policy is optimal and $V_K(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\alpha - \mathbf{C}$. Therefore, all properties hold for the K-th horizon, which completes the proof. □

Appendix A.4: Additional Numerical Studies

A.4.1 Numerical study with $N = 4$ and $T = 3$

In Table 2.1, we consider problems with size $N = 4$ and $T = 3$. For such a problem size, we can use backward induction to calculate the firm's optimal revenue, which serves as a benchmark to evaluate the performance of the RCEC heuristics. Below we describe the design of the numerical study in detail. The description consists of three parts: demand patterns, economic parameters, and initial capacity.

Demand patterns

To cover a wide range of demand scenarios, we consider 13 evolution patterns for product demand means in Table A.1. For each evolution pattern, we define vectors $\mu^t = (\mu_1^t, \cdots, \mu_N^t)^\intercal$ $(t = 1, \cdots, T)$, where μ_i^t is the demand mean of product i in period t. The demand mean patterns in Table A.1 cover some typical realistic scenarios. For instance, in pattern 4, the expected demand for high-quality products are higher than that for low-quality products when the period is close to the end of horizon, which corresponds to revenue management situations.

Pattern	Description	Example ($T = 3$)
Linear	1. Product 1 demand increases, product 2 demand is flat, product 3 and 4 demands decrease with the same rate.	$\begin{pmatrix} 2 & 4 & 6 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	2. Product 1 demand increases, product 2 demand is flat, product 3 demand decreases, product 4 demand decreases in half of the rate of product 3.	$\begin{pmatrix} 2 & 4 & 6 & 4 \\ 4 & 4 & 4 & 3 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	3. Product 1, 2, 3 and 4 demands are flat.	$\begin{pmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{pmatrix}$
	4. Product 1 and 2 demands increase with the same rate, product 3 and 4 demands decrease with the same rate.	$\begin{pmatrix} 2 & 2 & 6 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 2 & 2 \end{pmatrix}$
	5. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand decreases, product 4 demand decreases in half of the rate of product 3.	$\begin{pmatrix} 2 & 2 & 6 & 4 \\ 4 & 3 & 4 & 3 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	6. Product 1 and 2 demands increase with the same rate, product 3 demand is flat, product 4 demand decreases.	$\begin{pmatrix} 2 & 2 & 4 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 4 & 2 \end{pmatrix}$
	7. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand is flat, product 4 demand decreases.	$\begin{pmatrix} 2 & 2 & 4 & 6 \\ 4 & 3 & 4 & 4 \\ 6 & 4 & 4 & 2 \end{pmatrix}$
	8. Product 1 and 2 demands increase with the same rate, product 3 and 4 demands are flat.	$\begin{pmatrix} 2 & 2 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 4 & 4 \end{pmatrix}$
	9. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 and 4 demands are flat.	$\begin{pmatrix} 2 & 2 & 4 & 4 \\ 4 & 3 & 4 & 4 \\ 6 & 4 & 4 & 4 \end{pmatrix}$
Alternating	10. Products 1 and 3 start with positive demand, while products 2 and 4 start with zero demand.	$\begin{pmatrix} 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}$
	11. Products 1 and 3 start with positive demand, where demand 3 is smaller than demand 1 in each period, and products 2 and 4 start with zero demand, where demand 4 is smaller than demand 2 in each period.	$\begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 6 & 0 & 2 \\ 6 & 0 & 2 & 0 \end{pmatrix}$
	12. Products 2 and 4 start with positive demand, while products 1 and 3 start with zero demand.	$\begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \end{pmatrix}$
	13. Products 2 and 4 start with positive demand, where demand 4 is smaller than demand 2 in each period, and products 1 and 3 start with zero demand, where demand 3 is smaller than demand 1 in each period.	$\begin{pmatrix} 0 & 6 & 0 & 2 \\ 6 & 0 & 2 & 0 \\ 0 & 6 & 0 & 2 \end{pmatrix}$

Table A.1.: Demand patterns with 4 products.

Given an evolution pattern μ^t $(t = 1, \cdots, T)$ for the demand means, we generate a sample of random demands for each product in each period. Specifically, given the demand mean μ^t in period t, we generate demand $\mathbf{D}^t$ by using either Poisson distribution or multivariate normal distribution with covariance matrix

$$\begin{pmatrix} 0.5 & 0.15 & 0.075 & 0.0375 \\ 0.15 & 0.5 & 0.15 & 0.075 \\ 0.075 & 0.15 & 0.5 & 0.15 \\ 0.0375 & 0.075 & 0.15 & 0.5 \end{pmatrix} * \mu^t,$$

and

$$\begin{pmatrix} 0.5 & -0.15 & -0.075 & -0.0375 \\ -0.15 & 0.5 & -0.15 & -0.075 \\ -0.075 & -0.15 & 0.5 & -0.15 \\ -0.0375 & -0.075 & -0.15 & 0.5 \end{pmatrix} * \mu^t.$$

The first covariance matrix represents positive correlation between the products, while the second represents negative correlation between the products. For normal distribution, we truncate the demand realizations at zero and round them to the nearest integer values. By the above construction, there are total $39 = 3 * 13$ demand scenarios.

Economic parameters

We also consider a wide variety of values for the economic parameters while using the same backorder cost $(g_1, g_2, g_3, g_4) = (1.0, 0.9, 0.8, 0.7)$. Recall the upgrading revenue is given by $\alpha_{ij} = r_j + g_j - u_i$; instead of specifying r_j and u_i, we choose to specify α_{ij}, which is sufficient for the numerical study. Four different matrices of $\alpha = (\alpha_{ij})_{4\times 4}$ have been considered in the numerical study. The capacity costs (c_1, c_2, c_3, c_4) are decided by $c_i = 0.3\alpha_{ii}$ $(i = 1, \cdots, 4)$ for each matrix.

1. Matrix 1: The parallel revenue decreases in product class, and the upgrading revenues are close to the parallel revenue.

$$\begin{pmatrix} 16 & 14 & 12 & 10 \\ 0 & 15 & 13 & 11 \\ 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$

2. Matrix 2: The parallel revenue decreases in product class, and the upgrading revenues are decreasing in the number of levels of upgrading (e.g., 1-step upgrading revenue is 11 and 2-step upgrading is either 7 or 8).

$$\begin{pmatrix} 16 & 11 & 7 & 4 \\ 0 & 15 & 11 & 8 \\ 0 & 0 & 14 & 11 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$

3. Matrix 3: The parallel revenue decreases in product class, and α_{12} and α_{34} are higher than α_{23}.

$$\begin{pmatrix} 16 & 14 & 5 & 3 \\ 0 & 15 & 6 & 4 \\ 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$

4. Matrix 4: The parallel revenue is constant across products, the 2-step upgrading revenue is constant, and α_{23} is higher than the other 1-step upgrading revenue.

$$\begin{pmatrix} 16 & 10 & 9 & 3 \\ 0 & 16 & 15 & 9 \\ 0 & 0 & 16 & 10 \\ 0 & 0 & 0 & 16 \end{pmatrix}$$

Initial capacity

When choosing the initial capacity, we start with optimal capacity level $\mathbf{X}_{\text{RCEC}}$ using the RCEC heuristic. To ensure the robustness of the results, we also consider a number of variants of $\mathbf{X}_{\text{RCEC}}$, among which some are extreme capacity scenarios. In particular, we use $\mathbf{X}_{\text{RCEC}} = (x_1^{\text{RCEC}}, x_2^{\text{RCEC}}, x_3^{\text{RCEC}}, x_4^{\text{RCEC}})$ to construct the following five patterns of initial capacity:

1. $\mathbf{X} = \lambda \mathbf{X}_{\text{RCEC}}$
2. For each $i \in \{1, 2, 3\}$: $(\mathbf{X})_i = \lambda(x_i^{\text{RCEC}} + x_{i+1}^{\text{RCEC}}),\ (\mathbf{X})_{i+1} = 0,\ (\mathbf{X})_s = \lambda x_s^{\text{RCEC}}$, $\forall\ s \in \{1, 2, 3, 4\} \setminus \{i, i+1\}$
3. For each $i \in \{2, 3, 4\}$: $(\mathbf{X})_i = 0,\ (\mathbf{X})_s = \lambda x_s^{\text{RCEC}},\ \forall\ s \in \{1, 2, 3, 4\} \setminus \{i\}$
4. $\mathbf{X} = \lambda(x_1^{\text{RCEC}} + x_2^{\text{RCEC}}, 0, x_3^{\text{RCEC}} + x_4^{\text{RCEC}}, 0)$
5. $\mathbf{X} = \lambda(x_1^{\text{RCEC}} + x_3^{\text{RCEC}}, x_2^{\text{RCEC}} + x_4^{\text{RCEC}}, 0, 0)$,

where $\lambda = \{0.75,\ 1,\ 1.25\}$. Each pattern corresponds to a realistic or extreme scenario. For instance, in Pattern 2, a certain product has extremely low inventory level while the adjacent high-quality product is abundant; in Pattern 5, the last two products have extremely low investment while there are plenty of higher level products. Note that in some of the patterns (e.g., Patterns 2-5), upgrading would be frequently performed. The parameter λ is used to adjust the capacity-demand ratio (e.g., $\lambda = 0.75$ implies that the aggregate capacity level is relatively low). For each λ, there are 9 initial capacity scenarios; so there are totally 27 capacity scenarios.

To summarize, we test $39 * 4 * 27 = 4212$ problem instances by the above construction. They cover a wide range of possible situations that may arise in practice.

A.4.2 Numerical study with $N = 5$ and $T \in \{3, 15, 30\}$

This is the major numerical study in Chapter 2; it serves several purposes. First, we test the performance of the RCEC heuristic for problems with larger sizes in Tables 2.2 and 2.3; second, we examine the value of multi-step upgrading in Tables

2.4 and 2.5; third, we investigate the importance of the allocation mechanism and the capacity sizing decision in Table 2.6. To make the results comparable across different T values, we make a couple of assumptions: (1) For each product i, the expected total demand throughout the sales horizon is the same for different T values; that is, the sum $\sum_{t=1}^{T} \mu_i^t$ in each demand evolution pattern $\mu^t = (\mu_1^t, \cdots, \mu_N^t)$ $(t = 1, \cdots, N)$ is the same for different T values, which is set to be 60 for each i. (2) For each parameter combination, the capacity cost is the same for different T's. Below we describe the design of the numerical study in detail. Again the description consists of three parts: demand patterns, economic parameters, and initial capacity.

Demand patterns

Similar to the first numerical study in Section A.4.1, we consider 13 demand evolution patterns in Table A.2.

Again, given an evolution pattern μ^t $(t = 1, \cdots, T)$ for the demand means, we generate a sample of random demands for each product in each period. Specifically, given the demand mean μ^t in period t, we generate demand $\mathbf{D}^t$ by using either Poisson distribution or multivariate normal distribution with a positive covariance matrix

$$\begin{pmatrix} 0.5 & 0.15 & 0.12 & 0.09 & 0.06 \\ 0.15 & 0.5 & 0.15 & 0.12 & 0.09 \\ 0.12 & 0.15 & 0.05 & 0.15 & 0.12 \\ 0.09 & 0.12 & 0.15 & 0.5 & 0.15 \\ 0.06 & 0.09 & 0.12 & 0.15 & 0.5 \end{pmatrix} * \mu^t,$$

and a negative covariance matrix

$$\begin{pmatrix} 0.5 & -0.15 & -0.12 & -0.09 & -0.06 \\ -0.15 & 0.5 & -0.15 & -0.12 & -0.09 \\ -0.12 & -0.15 & 0.05 & -0.15 & -0.12 \\ -0.09 & -0.12 & -0.15 & 0.5 & -0.15 \\ -0.06 & -0.09 & -0.12 & -0.15 & 0.5 \end{pmatrix} * \mu^t.$$

Pattern	Description	Example ($T = 3$)
Linear	1. Product 1 demand increases, product 2 and 3 demands are flat, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 8 & 8 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	2. Product 1 demand increases, product 2 and 3 demands are flat, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 8 & 8 & 12 & 8 \\ 8 & 8 & 8 & 8 & 6 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	3. Product 1, 2, 3, 4 and 5 demands are flat.	$\begin{pmatrix} 8 & 8 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \end{pmatrix}$
	4. Product 1, 2 and 3 demands increase with the same rate, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 4 & 4 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 12 & 4 & 4 \end{pmatrix}$
	5. Product 1 demand increases, product i ($i = 2, 3$) demand increases in half of the rate of product $i - 1$, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 4 & 4 & 12 & 8 \\ 8 & 6 & 5 & 8 & 6 \\ 12 & 8 & 6 & 4 & 4 \end{pmatrix}$
	6. Product 1 and 2 demands increase with the same rate, product 3 demand is flat, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 4 & 8 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 8 & 4 & 4 \end{pmatrix}$
	7. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand is flat, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 4 & 8 & 12 & 8 \\ 8 & 6 & 8 & 8 & 6 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	8. Product 1 and 2 demands increase with the same rate, product 3, 4 and 5 demands are flat.	$\begin{pmatrix} 4 & 4 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 8 & 8 & 8 \end{pmatrix}$
	9. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3, 4 and 5 demands are flat.	$\begin{pmatrix} 4 & 4 & 8 & 8 & 8 \\ 8 & 6 & 8 & 8 & 8 \\ 12 & 8 & 8 & 8 & 8 \end{pmatrix}$
Alternating	10. Products 1, 3 and 5 start with positive demand, while products 2 and 4 start with zero demand.	$\begin{pmatrix} 8 & 0 & 8 & 0 & 8 \\ 0 & 8 & 0 & 8 & 0 \\ 8 & 0 & 8 & 0 & 8 \end{pmatrix}$
	11. Products 1, 3 and 5 start with positive demand, where demand i ($i = 3, 5$) is smaller than demand $i - 2$ in each period, and products 2 and 4 start with zero demand, where demand 4 is smaller than demand 2 in each period.	$\begin{pmatrix} 16 & 0 & 8 & 0 & 4 \\ 0 & 16 & 0 & 8 & 0 \\ 16 & 0 & 8 & 0 & 4 \end{pmatrix}$
	12. Products 2 and 4 start with positive demand, while products 1, 3 and 5 start with zero demand.	$\begin{pmatrix} 0 & 8 & 0 & 8 & 0 \\ 8 & 0 & 8 & 0 & 8 \\ 0 & 8 & 0 & 8 & 0 \end{pmatrix}$
	13. Products 2 and 4 start with positive demand, where demand 4 is smaller than demand 2 in each period, and products 1, 3 and 5 start with zero demand, where demand i ($i = 3, 5$) is smaller than demand $i - 2$ in each period.	$\begin{pmatrix} 0 & 16 & 0 & 8 & 0 \\ 16 & 0 & 8 & 0 & 4 \\ 0 & 16 & 0 & 8 & 0 \end{pmatrix}$

Table A.2.: Demand patterns with 5 products.

The rest of the details are the same as in the first numerical study and therefore omitted. There are totally $3 * 13 = 39$ demand scenarios.

Economic parameters

For all problem instances, we use the same backorder cost vector $(g_1, \cdots, g_5) = (6.0, 5.7, 5.4, 5.1, 4.8)$. Four different matrices of $\alpha = (\alpha_{ij})_{5\times 5}$ have been considered. The capacity cost $(c_1, \cdots, c_5)$ are decided by $c_i = 0.3 * \alpha_{ii}$ $(i = 1, \cdots, 5)$ for each matrix.

1. Matrix 1: Upgrading revenue is close to the parallel revenue.

$$\begin{pmatrix} 17 & 15 & 13 & 11 & 9 \\ 0 & 16 & 14 & 12 & 10 \\ 0 & 0 & 15 & 13 & 11 \\ 0 & 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 0 & 13 \end{pmatrix}$$

2. Matrix 2: Revenues of 1-step upgrading are identical for different classes.

$$\begin{pmatrix} 17 & 12 & 8 & 5 & 3 \\ 0 & 16 & 12 & 9 & 7 \\ 0 & 0 & 15 & 12 & 10 \\ 0 & 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 0 & 13 \end{pmatrix}$$

3. Matrix 3: α_{12} is much smaller than parallel revenue α_{11}. However, α_{23}, α_{34} and α_{45} are close to α_{22}, α_{33} and α_{44}, respectively.

$$\begin{pmatrix} 17 & 11 & 9 & 7 & 4 \\ 0 & 16 & 14 & 12 & 9 \\ 0 & 0 & 15 & 13 & 10 \\ 0 & 0 & 0 & 14 & 11 \\ 0 & 0 & 0 & 0 & 13 \end{pmatrix}$$

4. Matrix 4: α_{12} and α_{45} are close to parallel revenues α_{11} and α_{44}, respectively. However, α_{23} and α_{34} are much smaller than α_{22} and α_{33}.

$$\begin{pmatrix} 17 & 15 & 10 & 6 & 4 \\ 0 & 16 & 11 & 7 & 5 \\ 0 & 0 & 15 & 11 & 9 \\ 0 & 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 0 & 13 \end{pmatrix}$$

Initial capacity

Similar to the first numerical study, we use $\mathbf{X}_{\mathrm{RCEC}}$ to construct the following five patterns of initial capacity.

1. $\mathbf{X} = \lambda \mathbf{X}_{\mathrm{RCEC}}$
2. For each $i, j \in \{1, 2, 3, 4, 5\}$ with $i < j$: $(\mathbf{X})_i = \lambda \left((\mathbf{X}_{\mathrm{RCEC}})_i + (\mathbf{X}_{\mathrm{RCEC}})_j \right)$, $(\mathbf{X})_j = 0$, $(\mathbf{X})_s = \lambda (\mathbf{X}_{\mathrm{RCEC}})_s$, $\forall\ s \in \{1, 2, 3, 4, 5\} \setminus \{i, j\}$
3. For each $i \in \{2, 3, 4, 5\}$: $(\mathbf{X})_i = 0$, $(\mathbf{X})_s = \lambda (\mathbf{X}_{\mathrm{RCEC}})_s$, $\forall\ s \in \{1, 2, 3, 4, 5\} \setminus \{i\}$
4. $\mathbf{X} = \lambda \left((\mathbf{X}_{\mathrm{RCEC}})_1 + (\mathbf{X}_{\mathrm{RCEC}})_2, 0, (\mathbf{X}_{\mathrm{RCEC}})_3 + (\mathbf{X}_{\mathrm{RCEC}})_4, 0, (\mathbf{X}_{\mathrm{RCEC}})_5 \right)$
5. $\mathbf{X} = \lambda \left((\mathbf{X}_{\mathrm{RCEC}})_1, (\mathbf{X}_{\mathrm{RCEC}})_2 + (\mathbf{X}_{\mathrm{RCEC}})_4, 0, (\mathbf{X}_{\mathrm{RCEC}})_3 + (\mathbf{X}_{\mathrm{RCEC}})_5, 0 \right)$,

where $\lambda = \{0.7, 0.9, 1.0, 1.1, 1.3\}$. Again the parameter λ is used to adjust the capacity-demand ratio (e.g., $\lambda = 0.7$ implies that the aggregate capacity level is relatively low). For each λ, there are 17 initial capacity scenarios; so there are totally 85 capacity scenarios.

To summarize, we test $39 * 4 * 85 = 13260$ problem instances by the above construction. They cover a wide range of possible situations that may arise in practice.

A.4.3 Impact of Allocation Mechanism: Suboptimal k-Step Upgrading

We examine the profit loss of adopting suboptimal, k-step upgrading ($k = 0, \cdots, N-2$). Given $\mathbf{X}_{\text{RCEC}}$ as the optimal initial capacity under full-step upgrading, define the profit loss of using only k-step upgrading as

$$\Delta_{k-step} = \left| \frac{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}) - \Pi^k_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})}{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})} \right| * 100\%, \qquad k = 0, 1, 2, 3,$$

where $\Pi^k_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})$ is the revenue from the k-step upgrading. The statistics are presented in Table A.3. We can see that the magnitudes of profit losses are still generally much larger than those for $\Delta_{\mathbf{X}_{\text{CB}}}$ and $\Delta_{\mathbf{X}_{\text{NV}}}$ (given in Table 2.6).

	Mean	Std.	Median	90%-percentile	Max.
Δ_{0-step}	4.28559	4.28222	2.84912	9.36076	31.80136
Δ_{1-step}	1.08141	1.38881	0.45090	3.36765	7.71787
Δ_{2-step}	0.33484	0.56076	0.09118	1.03939	3.53501
Δ_{3-step}	0.10276	0.25266	0.00874	0.28810	1.97719

Table A.3.: Profit loss of suboptimal allocation with k-step upgrading.

Appendix A.5: Other

The following lemma shows the relation between $\mathbf{N}$ and its effective state $\hat{\mathbf{N}}$.

Lemma A.5.1 *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$, *then* $\sum_{s=i}^{j} \hat{n}_s \leq \sum_{s=i}^{j} n_s$ *if* $\hat{n}_i > 0$, *and* $\sum_{s=i}^{j} \hat{n}_s \geq \sum_{s=i}^{j} n_s$ *if* $\hat{n}_{j+1} > 0$. *Especially,* $\sum_{s=i}^{N} \hat{n}_s \geq \sum_{s=i}^{N} n_s$.

Proof. The proof follows from the definition of the effective state. For any class k ($1 \leq k \leq N$), when applying the greedy upgrading to $\mathbf{N}$, there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ if $\hat{n}_k > 0$, and such an upgrade may exist if $\hat{n}_k \leq 0$. Hence, $\sum_{s=i}^{j} \hat{n}_s \leq \sum_{s=i}^{j} n_s$ if $\hat{n}_i > 0$, and $\sum_{s=i}^{j} \hat{n}_s \geq \sum_{s=i}^{j} n_s$ if $\hat{n}_{j+1} > 0$. The same argument shows that $\sum_{s=i}^{N} \hat{n}_s \geq \sum_{s=i}^{N} n_s$. $\square$

The next proposition shows that separation can be preserved under the effective state operation.

Proposition A.5.1 *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$. *For any demand realization* $\mathbf{D}$, *class* i $(i < N)$ *is the lowest class which is separable in* $\mathbf{N} - \mathbf{D}$ *if and only if* i *is the lowest class which is separable in* $\hat{\mathbf{N}} - \mathbf{D}$.

Proof. Suppose class i is the lowest separable class in $\mathbf{N} - \mathbf{D}$ but is not separable in $\hat{\mathbf{N}} - \mathbf{D}$. Then, there exists class a $k \leq i$ such that $\sum_{s=k}^{i} (\hat{n}_s - d_s) > 0$. First, we must have $k < i$; otherwise, $n_i \geq \hat{n}_i > d_i \geq 0$, which means class i is not separable in $\mathbf{N} - \mathbf{D}$ and is a contradiction. Given $k < i$, without loss of generality, we assume k is the lowest class with $\sum_{s=k}^{i} (\hat{n}_s - d_s) > 0$, which also implies $\hat{n}_k > d_k \geq 0$ since $\sum_{s=k+1}^{i} (\hat{n}_s - d_s) \leq 0$. Thus, $\sum_{s=k}^{i} n_s \geq \sum_{s=k}^{i} \hat{n}_s$ by Lemma A.5.1, and $\sum_{s=k}^{i} (n_s - d_s) \geq \sum_{s=k}^{i} (\hat{n}_s - d_s) > 0$ which contradicts the assumption of class i being separable in $\mathbf{N} - \mathbf{D}$. Hence, class i must be separable in $\hat{\mathbf{N}} - \mathbf{D}$ as well.

Next, we prove that i is the lowest separable class in $\hat{\mathbf{N}} - \mathbf{D}$. Suppose to the contrary that class $i' > i$ is the lowest separable class in $\hat{\mathbf{N}} - \mathbf{D}$, i.e., $\sum_{s=k}^{i'} (\hat{n}_s - d_s) \leq 0$ for all classes k $(1 \leq k \leq i')$. Then, $\hat{n}_{i'+1} - d_{i'+1} > 0$; otherwise, $i' + 1$ will be the lowest separable class. Because class i is the lowest separable class in $\mathbf{N} - \mathbf{D}$ and $i' > i$, there exists class $r \leq i'$ such that $\sum_{s=r}^{i'} (n_s - d_s) > 0$. Given $\hat{n}_{i'+1} > d_{i'+1} \geq 0$ and Lemma A.5.1, there is $\sum_{s=r}^{i'} \hat{n}_s \geq \sum_{s=r}^{i'} n_s$ and $\sum_{s=r}^{i'} (\hat{n}_s - d_s) \geq \sum_{s=r}^{i'} (n_s - d_s) > 0$, which is a contradiction since $\sum_{s=r}^{i'} (\hat{n}_s - d_s) \leq 0$. Therefore, class i is the lowest separable class in $\hat{\mathbf{N}} - \mathbf{D}$.

The necessary condition can be similarly proved. This completes the proof. $\square$

For any demand realization $\mathbf{D}$ in period t $(1 \leq t \leq T)$, let $\hat{\mathbf{N}}$ be the effective state of $\mathbf{N}$, Proposition A.5.2 gives the relation between the outcomes of applying the PSR algorithm to initial states $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$.

Proposition A.5.2 *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$, *and the PSR algorithm solves the general upgrading problem in period* t $(1 \le t \le T)$*. For any demand realization* $\mathbf{D}$ *in period* t*, let* $\mathbf{N}' = (n'_1, \cdots, n'_N)$ *and* $\hat{\mathbf{N}}' = (\hat{n}'_1, \cdots, \hat{n}'_N)$ *be the effective states of the outcomes of applying the PSR algorithm to* $(\mathbf{N}, \mathbf{D})$ *and* $(\hat{\mathbf{N}}, \mathbf{D})$*, respectively. Then,* $\mathbf{N}' = \hat{\mathbf{N}}'$ *if classes* $(1, \cdots, N)$ *are not separable under* $\mathbf{N} - \mathbf{D}$*. Especially,* $\mathbf{N}'$ *and* $\hat{\mathbf{N}}'$ *are the outcomes of applying the PSR algorithm to* $(\mathbf{N}, \mathbf{D})$ *and* $(\hat{\mathbf{N}}, \mathbf{D})$ *in period* T*.*

Proof. From Proposition A.5.1, classes $(1, \cdots, N)$ are also not separable under $\hat{\mathbf{N}} - \mathbf{D}$.

First, we must have $\mathbf{N}' \ge 0$. Suppose to the contrary that class k is the highest class with $n'_k < 0$. Since $\mathbf{N}'$ is the effective state of the outcome of applying the PSR to $(\mathbf{N}, \mathbf{D})$, there is $n'_1 = \cdots = n'_{k-1} = 0$. Note that any allocation decision is a transfer between two classes, which is true in both the PSR and the greedy upgrading. Thus, we have $0 > \sum_{s=r}^{k} n'_s \ge \sum_{s=r}^{k}(n_s - d_s)$ for any class $r < k$, where the second inequality is strict if there is any upgrade between classes $(1, \cdots, r-1)$ and $(r, \cdots, k)$ when applying the PSR or generating the effective state. Hence, class k is separable, which contradicts the assumption. Similarly, we know $\hat{\mathbf{N}}' \ge 0$.

Next, we show $\mathbf{N}' = \hat{\mathbf{N}}'$. Let class k be the lowest class such that $\hat{n}'_k \ne n'_k$. From the above argument, we have $\sum_{s=k}^{N} n'_s = \sum_{s=k}^{N}(n_s - d_s)$ if there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ in either solving $(\mathbf{N}, \mathbf{D})$ by the PSR or generating the effective state $\mathbf{N}'$. Furthermore, such an upgrade exists only if $n'_k = 0$ by the optimality of the PSR and the definition of the greedy upgrading, in which case $\sum_{s=k}^{N} n'_s > \sum_{s=k}^{N}(n_s - d_s)$. The same argument can be applied to $(\hat{\mathbf{N}}, \mathbf{D})$. With these observations, we derive contradictions for all possible cases.

1. $\sum_{s=k}^{N} \hat{n}_s = \sum_{s=k}^{N} n_s$.
   (a) $\hat{n}'_k > 0, n'_k > 0$: For both initial states $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$, there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ in either applying

the PSR or generating the effective state. Moreover, $\hat{n}_k \geq \hat{n}'_k > 0$ implies that $\sum_{s=k}^{N} \hat{n}_s = \sum_{s=k}^{N} n_s$ by Lemma A.5.1. Thus,

$$\sum_{s=k}^{N} (\hat{n}_s - d_s) = \sum_{s=k}^{N} \hat{n}'_s \neq \sum_{s=k}^{N} n'_s = \sum_{s=k}^{N} (n_s - d_s),$$

which is a contradiction;

(b) $\hat{n}'_k > n'_k = 0$: Similar to the previous case, since $\hat{n}_k \geq \hat{n}'_k > 0$, then

$$\sum_{s=k}^{N} \hat{n}'_s = \sum_{s=k}^{N} (\hat{n}_s - d_s) = \sum_{s=k}^{N} (n_s - d_s) \leq \sum_{s=k}^{N} n'_s,$$

which violates the assumption of class k;

(c) $n'_k > \hat{n}'_k = 0$: From Lemma A.5.1, we similarly have

$$\sum_{s=k}^{N} \hat{n}'_s \geq \sum_{s=k}^{N} (\hat{n}_s - d_s) \geq \sum_{s=k}^{N} (n_s - d_s) = \sum_{s=k}^{N} n'_s,$$

which is also a contradiction;

2. $\sum_{s=k}^{N} \hat{n}_s > \sum_{s=k}^{N} n_s$. In this case, $\hat{n}_k = 0$ by the definition of the effective state $\hat{\mathbf{N}}$. Since $\hat{n}'_k \geq 0$ from the previous discussion, there is $\hat{n}_k = \hat{n}'_k = 0$. Meanwhile, $n'_k \neq \hat{n}'_k$ by the assumption of k. From $n'_k \geq 0$, we must have $n'_k > \hat{n}'_k = 0$, and

$$\sum_{s=k}^{N} \hat{n}'_s \geq \sum_{s=k}^{N} (\hat{n}_s - d_s) > \sum_{s=k}^{N} (n_s - d_s) = \sum_{s=k}^{N} n'_s,$$

where the first inequality is from the fact that there might be upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ while solving $(\hat{\mathbf{N}}, \mathbf{D})$ by the PSR and generating the effective state $\hat{\mathbf{N}}'$. This is a contradiction since $n'_k > \hat{n}'_k$ and $\sum_{s=k+1}^{N} n'_s = \sum_{s=k+1}^{N} \hat{n}'_s$ by assumption of k.

Therefore, $\mathbf{N} = \hat{\mathbf{N}}'$. Note that the PSR optimally solves the general upgrading problem with protection levels being 0 in period T by Lemma 2.5.1. Since the greedy upgrading is equivalent to the PSR with protection levels being 0, we know $\mathbf{N}'$ and $\hat{\mathbf{N}}'$ are the outcomes of applying the PSR algorithm to $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$ in period T, which completes the proof. $\square$

Lemma A.5.2 considers a general upgrading problem with special states in period T, which can be used to simplify the proof of Lemma 2.5.1.

Lemma A.5.2 *Consider an N-class general upgrading problem with states $\mathbf{N} = (n_1, \cdots, n_N)$ and demand realization $\mathbf{D}$ in period T. Suppose classes $(1, \cdots, N)$ are not separable based on $\mathbf{N}-\mathbf{D}$, $(n_{i+1}, \cdots, n_{j-1}) \leq 0$ and $n_j < 0$. Then, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ and $\Delta_{ij}^{-+}\Theta^T(\mathbf{N}|\mathbf{D})$ are independent of the values of $(n_j, \cdots, n_N)$.*

Proof. Since $\Theta^T(\mathbf{N}|\mathbf{D})$ is piecewise linear and concave (see Murty 1983), both $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ and $\Delta_{ij}^{-+}\Theta^T(\mathbf{N}|\mathbf{D})$ exist.

We focus on the proof of Δ_{ij}^{+-}, and the same argument applies to Δ_{ij}^{-+}. We consider the dual form of the general upgrading problem with initial state $(\mathbf{N}, \mathbf{D})$, and let the dual variables be $(\lambda_1, \cdots, \lambda_N)$, where λ_i corresponds to the constraint of class i. The dual problem is

$$\begin{aligned} \min_{(\lambda_1,\cdots,\lambda_N)\geq 0} \; & \sum_{s=1}^{N} |n_s - d_s|\lambda_s && \text{(A.23)} \\ \text{s.t.} \quad & \lambda_s + \lambda_r \geq \alpha_{sr}, \\ & s, r \in \{s, r | (\mathbf{N}-\mathbf{D})_s \geq 0, (\mathbf{N}-\mathbf{D})_r < 0, \; 1 \leq s < r \leq N\}. \end{aligned}$$

1. $n_i \geq 0$: By Linear Programming theory, there is

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \begin{cases} \lambda_i^* + \lambda_j^* - g_j, & \text{if } n_i \geq d_i \\ -\lambda_i^* + \lambda_j^* + \alpha_{ii} - g_j, & \text{if } n_i < d_i, \end{cases}$$

   where $\lambda^* = (\lambda_1^*, \cdots, \lambda_N^*)$ is optimal in the dual problem (A.23).

   (a) $n_i > d_i$: Given classes $(1, \cdots, N)$ are not separable, we have $y_{kj}^* > 0$ for some class k $(1 \leq k \leq i)$. By the complementary slackness in the linear program, $\lambda_k^* + \lambda_j^* = \alpha_{kj}$. Assume without loss of generality that $i+1$ is the highest class s $(i+1 \leq s \leq j-1)$ with $n_s - d_s < 0$. Since it is optimal to first use class i's remaining capacity $n_i - d_i$ to satisfy demands from $(i+1, \cdots, j)$, there is $y_{i,i+1}^* > 0$, which implies $\lambda_i^* + \lambda_{i+1}^* = \alpha_{i,i+1}$.

By examining constraints $\lambda_i^* + \lambda_j^* \geq \alpha_{ij}$, $\lambda_k^* + \lambda_{i+1}^* \geq \alpha_{k,i+1}$ in the dual problem (A.23), as well as the assumption $\alpha_{kj} + \alpha_{i,i+1} = \alpha_{ij} + \alpha_{k,i+1}$, we have $\lambda_i^* + \lambda_j^* = \alpha_{ij}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \lambda_i^* + \lambda_j^* - g_j = \alpha_{ij} - g_j$. Note that $y_{ij} > 0$ if $n_s - d_s = 0$ for all classes s $(i+1 \leq s \leq j-1)$, which implies $\lambda_i^* + \lambda_j^* - g_j = \alpha_{ij} - g_j$.

(b) $n_i < d_i$: The non-separable assumption implies that there exist classes r $(r < i)$ and k $(k < i)$ such that $y_{ri}^* > 0$ and $y_{kj}^* > 0$. Thus, $\lambda_k^* + \lambda_j^* = \alpha_{kj}$ and $\lambda_r^* + \lambda_i^* = \alpha_{ri}$. We similarly have $\lambda_r^* + \lambda_j^* = \alpha_{rj}$ and $\lambda_k^* + \lambda_i^* = \alpha_{ki}$ by using the constraints in (A.23) and the assumption $\alpha_{kj} + \alpha_{ri} = \alpha_{rj} + \alpha_{ki}$. Thus, $-\lambda_i^* + \lambda_j^* = \alpha_{rj} - \alpha_{ri}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = -\lambda_i^* + \lambda_j^* + \alpha_{ii} - g_j = \alpha_{ij} - g_j$.

Since $\Theta^T(\mathbf{N}|\mathbf{D})$ is piecewise linear in n_i and n_j, then $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \alpha_{ij} - g_j$ when $n_i \geq 0$.

2. $n_i < 0$: In this case,

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = -\lambda_i^* + \lambda_j^* + g_i - g_j.$$

Note that this is similar to the case when $0 \leq n_i < d_i$. Hence, $-\lambda_i^* + \lambda_j^* = \alpha_{rj} - \alpha_{ri}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = r_j - r_i$.

Hence, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ is independent of the values of $(n_j, \cdots, n_N)$, which concludes the proof. $\square$

Suppose the PSR is optimal in period t. Then similar to Proposition A.5.2, the following proposition shows the relation between the outcomes of $\mathbf{N}$ and its effective state $\hat{\mathbf{N}}$ after applying the PSR given any demand realization $\mathbf{D}$.

Proposition A.5.3 *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$. *If the PSR algorithm solves the general upgrading problem in period* t, *and the protection levels in period* t *have the independence property. For any demand realization* $\mathbf{D}$, *let* $\mathbf{N}' = (n_1', \cdots, n_N')$ *and* $\hat{\mathbf{N}}' = (\hat{n}_1', \cdots, \hat{n}_N')$ *be the outcomes of applying the PSR algorithm to* $(\mathbf{N}, \mathbf{D})$ *and* $(\hat{\mathbf{N}}, \mathbf{D})$, *respectively.*

Let k be the highest class in $\mathbf{N}'$ such that $(n'_k, \cdots, n'_N) \geq 0$ and $n'_k > 0$, where $k = N+1$ if such a class does not exist in $\mathbf{N}'$. $\hat{k}$ is similarly defined in $\hat{\mathbf{N}}'$. Then, $k = \hat{k}$ and $(\mathbf{N}')_{k,\cdots,N} = (\hat{\mathbf{N}}')_{k,\cdots,N}$ if classes $(1, \cdots, N)$ are not separable under $\mathbf{N}-\mathbf{D}$.

Proof. We first show $(\hat{\mathbf{N}}')_{k,\cdots,N} = (\mathbf{N}')_{k,\cdots,N}$. Let i $(k \leq i \leq N)$ be the lowest class such that $\hat{n}'_i \neq n'_i \geq 0$. There are three cases.

1. $\hat{n}'_i > n'_i \geq 0$: Since capacity i may be used when applying the PSR, we have $\hat{n}_i \geq \hat{n}'_i > 0$, which implies $\sum_{s=i}^{N} \hat{n}_s = \sum_{s=i}^{N} n_s$ by Lemma A.5.1. Because $\hat{n}'_i > 0$, there is no upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\hat{\mathbf{N}}, \mathbf{D})$, and $\sum_{s=i}^{N}(\hat{n}_s - d_s) = \sum_{s=i}^{N} \hat{n}'_s$. On the other hand, $n'_i \geq 0$ implies that there could be upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when solving $(\mathbf{N}, \mathbf{D})$, thus $\sum_{s=i}^{N} n'_s \geq \sum_{s=i}^{N}(n_s - d_s)$. From the above, we have

$$\sum_{s=i}^{N} n'_s \geq \sum_{s=i}^{N}(n_s - d_s) = \sum_{s=i}^{N}(\hat{n}_s - d_s) = \sum_{s=i}^{N} \hat{n}'_s.$$

   This is a contradiction given the assumption of class i.

2. $n'_i > \hat{n}'_i$ and $n'_i > 0$: In this case, $n'_i > 0$ implies that there is no upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\mathbf{N}, \mathbf{D})$. Thus, $\sum_{s=i}^{N} n'_s = \sum_{s=i}^{N}(n_s - d_s)$. However, there could be upgrade between classes $(1, \cdots, i-1)$ and $(i, \cdots, N)$ when generating $\hat{\mathbf{N}}$ as well as applying the PSR to $(\hat{\mathbf{N}}, \mathbf{D})$, thus

$$\sum_{s=i}^{N} n'_i = \sum_{s=i}^{N}(n_s - d_s) \leq \sum_{s=i}^{N}(\hat{n}_s - d_s) \leq \sum_{s=i}^{N} \hat{n}'_s, \tag{A.24}$$

   which is a contradiction.

3. $n'_i = 0 > \hat{n}'_i$ and $i > k$: From (A.24), we only need to consider the case when $\sum_{s=i}^{N} n'_i > \sum_{s=i}^{N}(n_s - d_s)$, i.e., there is upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\mathbf{N}, \mathbf{D})$. Without loss of generality, we assume that $i-1$ is the highest class that upgrades the demands in classes $(i, \cdots, N)$ under initial state $(\mathbf{N}, \mathbf{D})$, and l $(l \geq i)$ is the lowest class being

upgraded by capacity $i-1$. Since there is no upgrade from classes $(1, \cdots, i-2)$ to $(i-1, \cdots, N)$ when solving $(\mathbf{N}, \mathbf{D})$, similar to (A.24), there is

$$\sum_{s=i-1}^{N} n'_i = \sum_{s=i-1}^{N} (n_s - d_s) \leq \sum_{s=i-1}^{N} (\hat{n}_s - d_s) \leq \sum_{s=i-1}^{N} \hat{n}'_s. \tag{A.25}$$

Since $n'_i = 0 > \hat{n}'_i$ and $(\hat{\mathbf{N}}')_{i+1,\cdots,N} = (\mathbf{N}')_{i+1,\cdots,N} \geq 0$ by assumption of class i, (A.25) implies $\hat{n}'_{i-1} > n'_{i-1} \geq 0$. Moreover, $\hat{n}_{i-1} > 0$ if $\hat{n}'_{i-1} > 0$.

Next, we show that the profit can be increased by upgrading demand i by capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$, which violates the optimality assumption of the PSR. Since $\hat{n}'_{i-1} > 0$ and the assumption of class $i-1$, there is no upgrade between classes $(1, \cdots, i-2)$ and $(i-1, \cdots, N)$ when generating the effective state $\hat{\mathbf{N}}$ as well as applying the PSR to both $(\hat{\mathbf{N}}, \mathbf{D})$ and $(\mathbf{N}, \mathbf{D})$. From Proposition A.5.2, given classes $(1, \cdots, N)$ are not separable under $\mathbf{N} - \mathbf{D}$, the effective states of $(\hat{\mathbf{N}}')_{1,\cdots,i-2}$ are the same as those of $(\mathbf{N}')_{1,\cdots,i-2}$.

If $l = i$, from the independence property of the protection levels, $p_{i-1,i}$ is the same for both $(\hat{\mathbf{N}}, \mathbf{D})$ and $(\mathbf{N}, \mathbf{D})$. Because $\hat{n}'_{i-1} > n'_{i-1}$ and capacity $i-1$ upgrades demand i under $(\mathbf{N}, \mathbf{D})$, it is also optimal to upgrade demand i by capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$.

If $l > i$, we have $(\hat{\mathbf{N}}')_{i+1,\cdots,l} = (\mathbf{N}')_{i+1,\cdots,l}$ by the assumption of class i. Moreover, $(\mathbf{N}')_{i+1,\cdots,l} = 0$ since capacity $i-1$ upgrades demand l under initial state $(\mathbf{N}, \mathbf{D})$, and n'_{i-1} is the remaining capacity after such upgrading. From the PSR, there is

$$\begin{aligned}
\alpha_{i-1,l} \geq & \Delta^{+-}_{i-1,l}\Theta^{t+1}(\mathbf{N}') \\
= & \Delta^{+-}_{i-1,l}\Theta^{t+1}\left((\mathbf{N}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\mathbf{N}')_{l+1,\cdots,N}\right) \\
= & \Delta^{+-}_{i-1,l}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
\geq & r_l - r_i + \partial^{+}_{i-1}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
& - \partial^{-}_{i}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
= & r_l - r_i + \Delta^{+-}_{i-1,i}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}),
\end{aligned} \tag{A.26}$$

where the second equality follows from the independence property of Θ^{t+1} and the fact that the effective states of $(\mathbf{N}')_{1,\cdots,i-2}$ and $(\hat{\mathbf{N}}')_{1,\cdots,i-2}$ are the same, and the last inequality is because of Lemma A.1.2.

Since $\alpha_{i-1,i} - \alpha_{i-1,l} = r_i + g_i - r_l - g_l$, where $g_i > g_l$, $\hat{n}'_{i-1} > n'_{i-1}$, and Θ^{t+1} is concave, we have

$$\begin{aligned}\alpha_{i-1,i} >&\Delta^{+-}_{i-1,i}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\ \geq&\Delta^{+-}_{i-1,i}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, \hat{n}'_{i-1}, n'_{i-1} - \hat{n}'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\ =&\Delta^{+-}_{i-1,i}\Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, \hat{n}'_{i-1}, \hat{n}'_{i}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}).\end{aligned} \tag{A.27}$$

Thus, the profit can be increased by upgrading demand i with capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$, which contradicts the optimality assumption of the PSR algorithm.

Hence, $(\hat{\mathbf{N}}')_{k,\cdots,N} = (\mathbf{N}')_{k,\cdots,N}$. Similarly, we know $(\mathbf{N}')_{\hat{k},\cdots,N} = (\hat{\mathbf{N}}')_{\hat{k},\cdots,N}$, which concludes the proof. □

A.5.1 Monotonicity

To prove the monotonicity result in Proposition 2.5.4, we start with a basic property.

Under certain conditions, the following lemma states that the marginal values, $\Delta^{+-}_{ij}\Theta^t$ $(i < j)$ and $\Delta^{-+}_{ij}\Theta^t$, remain the same if capacity k $(k < i)$ is used to "optimally" upgrade the back-logged demand i. Note that such an upgrade can go beyond class k as long as there is unmet demand i.

Lemma A.5.3 *Suppose $(\hat{n}_1, \cdots, \hat{n}_{i-1})$ is the effective state of $(n_1, \cdots, n_{i-1})$, and there exists class k $(1 \leq k < i)$ such that $\hat{n}_k > 0$ and $\hat{n}_{k+1} = \cdots = \hat{n}_{i-1} = 0$. If $(n_i, \cdots, n_j) \leq 0$, $\delta > 0$, and $n_i + \delta \leq 0 \leq \hat{n}_k - \delta$, then*

$$\Delta^{+-}_{ij}\Theta^t(n_1, \cdots, n_N) = \Delta^{+-}_{ij}\Theta^t(\hat{n}_1, \cdots, \hat{n}_{k-1}, \hat{n}_k - \delta, 0, \cdots, 0, n_i + \delta, n_{i+1}, \cdots, n_N) \tag{A.28}$$

and

$$\Delta_{ij}^{-+}\Theta^t(n_1,\cdots,n_N) = \Delta_{ij}^{-+}\Theta^t(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N). \tag{A.29}$$

Proof. It is sufficient to prove the equality in (A.28). From Proposition 2.5.1, there is

$$\Delta_{ij}^{+-}\Theta^t(n_1,\cdots,n_N) = \Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N).$$

Thus, for any demand realization $\mathbf{D}$ in period t, we use induction to show

$$\begin{aligned} &\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N|\mathbf{D}) \\ =&\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N|\mathbf{D}) \end{aligned} \tag{A.30}$$

under the conditions given in Lemma A.5.3. To simplify our notations, let

$$\begin{aligned} \mathbf{N} &= (\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N) \\ \bar{\mathbf{N}} &= (\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N). \end{aligned}$$

In period T, let r^* $(1\le r^*\le k)$ be the lowest class such that $n_i+\sum_{s=r^*}^k \hat{n}_s \ge \sum_{s=r^*}^i d_s$, i.e., r^* is the lowest class that satisfies the last unit of demand i. We analyze (A.30) based on following cases.

1. r^* does not exist: Then $n_i+\sum_{s=r}^k \hat{n}_s < \sum_{s=r}^i d_s$ and $n_i-\delta+\sum_{s=r}^k \hat{n}_s < -\delta+\sum_{s=r}^i d_s$ for all class r $(1\le r\le k)$. After applying the PSR, there is unmet demand i in both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$. Thus, given $(n_{i+1},\cdots,n_j)\le 0$, we have

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}|\mathbf{D}) = g_i-g_j.$$

2. r^* does exist: For both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$, since $n_i-\delta+\sum_{s=r^*}^k \hat{n}_s \ge -\delta+\sum_{s=r^*}^i d_s$, the last unit of demand i is fulfilled by capacity r^*. And the states of classes $(r^*,\cdots,i)$ after the last unit of demand i being satisfied are $(n_i+\sum_{s=r^*}^k \hat{n}_s-\sum_{s=r^*}^i d_s,0,\cdots,0)$. Hence,

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}|\mathbf{D}) = g_i-\alpha_{r^*i}+\frac{\partial}{\partial n_i^+}\Theta^T(\tilde{\mathbf{N}}|\tilde{\mathbf{D}})-\frac{\partial}{\partial n_j^-}\Theta^T(\tilde{\mathbf{N}}|\tilde{\mathbf{D}}),$$

where $\tilde{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_{r^*-1}, n_i + \sum_{s=r^*}^{k} \hat{n}_s - \sum_{s=r^*}^{i} d_s, 0, \cdots, 0, n_{i+1}, \cdots, n_N)$ and $\tilde{\mathbf{D}} = ((\mathbf{D})_{1,\cdots,r^*-1}, 0, \cdots, 0, (\mathbf{D})_{i+1,\cdots,N})$.

Hence, (A.30) holds in period T.

In period $t < T$, we apply the PSR algorithm to the general upgrading problem with initial states $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, and denote $\mathbf{N}'$ and $\bar{\mathbf{N}}'$ as the corresponding outcomes. We examine (A.30) based on the states of class i in $\mathbf{N}'$ and $\bar{\mathbf{N}}'$.

1. $(\mathbf{N}')_i = (\bar{\mathbf{N}}')_i = 0$: From the above analysis, the last unit of demand i is satisfied by class r^* in both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, and we assume $r^* = k$ without loss of generality. Hence,

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D}) = \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D}) = g_i - \alpha_{ki} + \frac{\partial}{\partial n_i^+}\Theta^t(\tilde{\mathbf{N}}|\tilde{\mathbf{D}}) - \frac{\partial}{\partial n_j^-}\Theta^t(\tilde{\mathbf{N}}|\tilde{\mathbf{D}}),$$

where $\tilde{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_{k-1}, n_i + \hat{n}_k - \sum_{s=k}^{i} d_s, 0, \cdots, 0, n_{i+1}, \cdots, n_N)$ and $\tilde{\mathbf{D}} = ((\mathbf{D})_{1,\cdots,k-1}, 0, \cdots, 0, (\mathbf{D})_{i+1,\cdots,N})$.

2. $(\mathbf{N}')_i < 0$ and $(\bar{\mathbf{N}}')_i < 0$: If there is no class r^* $(1 \le r^* \le k)$ such that $n_i + \sum_{s=r^*}^{k} \hat{n}_s \ge \sum_{s=r^*}^{i} d_s$, demand i and j will never be satisfied in the remaining periods for both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, which means

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D}) = (T - t + 1)(g_i - g_j) = \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D}).$$

Hence, we only need to consider the case when class r^* does exist. In this case, we assume $r^* = k$ without loss of generality. From Proposition 2.5.1, since $(\mathbf{N}')_i < 0$ and $(\bar{\mathbf{N}}')_i < 0$, MP_{ij} will not affect the optimal allocation decisions in period t under both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$. Thus,

$$\begin{aligned} \Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D}) &= g_i - g_j + \frac{\partial}{\partial n_i^+}\Theta^{t+1}(\mathbf{N}') - \frac{\partial}{\partial n_j^-}\Theta^{t+1}(\mathbf{N}'), \\ \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D}) &= g_i - g_j + \frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}') - \frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}'). \end{aligned} \tag{A.31}$$

By the definition of class k, there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ when applying the PSR under both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$; otherwise, all capacity k should have been depleted before performing the aforementioned upgrade, which means there is no unmet demand i. Since the initial states of classes $(1, \cdots, k-1)$ are the same, the effective states of classes $(1, \cdots, k-1)$ in $\mathbf{N}'$ are the same as those in $\bar{\mathbf{N}}'$. Note that $(\mathbf{N}')_{i+1,\cdots,j} = (\bar{\mathbf{N}}')_{i+1,\cdots,j} = (n_{i+1} - d_{i+1}, \cdots, n_j - d_j) \leq 0$ and $n_i + \hat{n}_k - \sum_{s=k}^{i} d_s > 0$ by assumption. Applying the induction assumption, we have

$$
\begin{aligned}
&\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}') - \frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}') \\
=&\Delta_{ij}^{+-}\Theta^{t+1}\Big((\mathbf{N}')_{1,\cdots,k-1}, n_i + \hat{n}_k + \delta - \sum_{s=k}^{i} d_s, 0, \cdots, 0, -\delta, (\mathbf{N}')_{i+1,\cdots,N}\Big) \\
=&\Delta_{ij}^{+-}\Theta^{t+1}\Big((\bar{\mathbf{N}}')_{1,\cdots,k-1}, n_i + \hat{n}_k + \delta - \sum_{s=k}^{i} d_s, 0, \cdots, 0, -\delta, (\bar{\mathbf{N}}')_{i+1,\cdots,N}\Big) \\
=&\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}') - \frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}'),
\end{aligned}
\tag{A.32}
$$

where $0 < \delta < -\max\big((\mathbf{N}')_i,\ (\bar{\mathbf{N}}')_i\big)$ and the second equality follows from Proposition 2.5.1. This is a contradiction. Hence, $\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D}) = \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D})$ from (A.31) and (A.32).

3. $(\mathbf{N}')_i = 0$ and $(\bar{\mathbf{N}}')_i < 0$: In this case, there exists a class r^*, which can be assumed as $r^* = k$ without loss of generality. Moreover, the last unit of demand i is upgraded by capacity k when the PSR solves $(\mathbf{N}, \mathbf{D})$.

   Given $(\bar{\mathbf{N}}')_i < 0$, we must have $(\bar{\mathbf{N}}')_k > (\mathbf{N}')_k \geq 0$ since the total unmet demand after parallel allocation in classes $(k, \cdots, i)$ is the same for both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$. When the last unit of demand i is upgraded by capacity k in $(\mathbf{N}, \mathbf{D})$, from the PSR, the upgrading decisions between classes $(1, \cdots, k-1)$ and $(i+1, \cdots, N)$ have not been considered yet. At that moment, the effective state of classes $(1, \cdots, k-1)$ in $(\mathbf{N}, \mathbf{D})$ is the same as that in $\bar{\mathbf{N}}'$ because there is also no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ in $(\bar{\mathbf{N}}, \mathbf{D})$ when applying the PSR. Hence, the protection levels between class k and the lower classes are the same for both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$ from Proposition 2.5.1. Let h $(k < h \leq i)$ be the highest class with $(\bar{\mathbf{N}}')_h < 0$. Similar to (A.26) and (A.27) in the proof of Proposition A.5.3, we can show that the profit from solving $(\bar{\mathbf{N}}, \mathbf{D})$

can be increased by upgrading demand h with capacity k. This contradicts the optimality of the PSR. Hence, this case cannot exist.

4. $(\mathbf{N}')_i < 0$ and $(\bar{\mathbf{N}}')_i = 0$: Similar to the previous case, this would lead to a contradiction.

Therefore, (A.30) holds for any demand realization $\mathbf{D}$ in period t, and this completes the induction proof. □

The following lemma shows that the protection level p_{ij} $(1 \le i < j \le N)$ in period $t-1$ is decreasing in the states of classes $(1, \cdots, , i-1)$ if the same monotonicity holds in period t.

Lemma A.5.4 *Consider an N-class upgrading problem in period t $(1 \le t < T)$ with $(n_{i+1}, \cdots, n_j) \le 0$. Let $\bar{\mathbf{N}} = \mathbf{N} + \epsilon \mathbf{e}_r$, where $1 \le r < i$ and $\epsilon > 0$. Then,*

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}) \ge \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}), \qquad \Delta_{ij}^{-+}\Theta^t(\mathbf{N}) \ge \Delta_{ij}^{-+}\Theta^t(\bar{\mathbf{N}}) \tag{A.33}$$

if the same inequality holds for Θ^{t+1}.

Proof. To prove (A.33), it is sufficient to show

$$\Theta^t(\mathbf{N}_{ij}) - \Theta^t(\mathbf{N}) \ge \Theta^t(\bar{\mathbf{N}}_{ij}) - \Theta^t(\bar{\mathbf{N}}), \tag{A.34}$$

where

$$\begin{aligned}
\mathbf{N}_{ij} &= (n_1, \cdots, n_{i-1}, n_i + 1, n_{i+1}, \cdots, n_{j-1}, n_j - 1, n_{j+1}, \cdots, n_N), \\
\bar{\mathbf{N}} &= (n_1, \cdots, n_{r-1}, n_r + 1, n_{r+1}, \cdots, n_N), \\
\bar{\mathbf{N}}_{ij} &= (n_1, \cdots, n_{r-1}, n_r + 1, n_{r+1}, \cdots, n_{i-1}, n_i + 1, n_{i+1}, \cdots, n_{j-1}, n_j - 1, n_{j+1}, \cdots, n_N).
\end{aligned}$$

In each period t, given any demand realization $\mathbf{D} = (d_1, \cdots, d_N)$, we next show

$$\Delta = \Theta^t(\mathbf{N}_{ij}|\mathbf{D}) - \Theta^t(\mathbf{N}|\mathbf{D}) \ge \Theta^t(\bar{\mathbf{N}}_{ij}|\mathbf{D}) - \Theta^t(\bar{\mathbf{N}}|\mathbf{D}) = \bar{\Delta}, \tag{A.35}$$

which proves (A.34).

To compare Δ and $\bar{\Delta}$, we consider upgrading decisions in period t. Denote $\mathbf{R}$ as the resulting states of classes $(1, \cdots, N)$ after applying the PSR under initial state

$(\mathbf{N}, \mathbf{D})$, let h be the highest capacity class which upgrades the demand in classes $(i, \cdots, N)$, and l is the lowest demand class that is upgraded by classes $(1, \cdots, i)$. By the definition of classes h and l, we have $h \leq l$, and $h = l$ only if $h = i$ and $l = i$. From the PSR algorithm, there is neither unmet demand nor remaining capacity between classes h and l in $\mathbf{R}$, i.e. $(\mathbf{R})_{h+1,\cdots,l-1} = \mathbf{0}$. $\bar{\mathbf{R}}$, $\bar{h}$ and $\bar{l}$ are similarly defined under initial state $(\bar{\mathbf{N}}, \mathbf{D})$.

For any classes $1 \leq k < s \leq N$ in period t, the protection level p_{ks} defined in (2.8) are decreasing in $(n_1, \cdots, n_{k-1})$ since (A.33) is true for Θ^{t+1}, thus upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$ rather than $(\mathbf{N}, \mathbf{D})$, i.e., $l \leq \bar{l}$.

Switching from $\mathbf{N}$ ($\bar{\mathbf{N}}$) to $\mathbf{N}_{ij}$ ($\bar{\mathbf{N}}_{ij}$), we not only change the current revenues in period t, but also the result $\mathbf{R}$ ($\bar{\mathbf{R}}$), which is the initial states in period $t+1$. Denote $\mathbf{R}'$ and $\bar{\mathbf{R}}'$ as the outcomes after applying the PSR under $(\mathbf{N}_{ij}, \mathbf{D})$ and $(\bar{\mathbf{N}}_{ij}, \mathbf{D})$, respectively. Then,

$$\Delta = \delta + \Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}), \qquad \bar{\Delta} = \bar{\delta} + \Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}),$$

where δ and $\bar{\delta}$ are the corresponding differences of the current period revenues in period t under $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, respectively.

When the initial states change from $\mathbf{N}$ to $\mathbf{N}_{ij}$, there are four cases which differ in the allocation decisions in period t. Note that the analogy applies when initial states change form $\bar{\mathbf{N}}$ to $\bar{\mathbf{N}}_{ij}$. For simplicity, we assume without loss of generality that $(\mathbf{R})_{l+1} < 0$.

Case 1: An extra unit of demand l is satisfied when $l < j$.

Case 2: An extra unit of capacity h is passed along to period $t+1$.

Case 3: An extra unit of demand $l+1$ is satisfied when $l+1 < j$.

Case 4: An extra unit of demand j is satisfied if $l \geq j$.

Here, we explain the above cases in detail by recalling the "chain reaction" described in the proof of Proposition 2.5.1.

Case 1: There is unmet demand l in $\mathbf{R}$ in this case. Note that capacity h is the highest class that upgrades demand l under $(\mathbf{N}, \mathbf{D})$. And the upgrade between classes h and l is bounded because either there is no capacity h remaining or the protection level p_{hl} is reached. When the initial state changes from $\mathbf{N}$ to $\mathbf{N}_{ij}$, from the chain reaction, there is an additional unit of capacity h which will upgrade the remaining demand l.

Case 2: In this case, class l demand has been fully satisfied in $\mathbf{R}$; otherwise, the analysis of Case 1 gives a contradiction.

Case 3: If $l+1 < j$, similar to *Case 1*, it is possible that an additional unit of demand $l+1$ is upgraded by capacity h under $(\mathbf{N}_{ij}, \mathbf{D})$, in which case all demand l has been satisfied in $\mathbf{R}$.

Case 4: Suppose that k_j is the highest class that upgrades demand j under $(\mathbf{N}, \mathbf{D})$. Because increasing n_i simultaneously decreases n_j by the same amount, there will be an additional unit of both capacity k_j and unmet demand j from the chain reaction. From the PSR, it is optimal to upgrade such an additional demand j by capacity k_j, and the outcome $\mathbf{R}' = \mathbf{R}$ in this case.

To compare Δ and $\bar{\Delta}$, we start with Case 4, where $l \geq j$ and $\mathbf{R}' = \mathbf{R}$ from the above discussion.

1. $n_i < 0$: Suppose the last unit of demand i is upgraded by class k_i, then $\Delta = g_i - g_j - \alpha_{k_i i} + \alpha_{k_i j} = r_j - r_i$;
2. $0 \leq n_i < d_i$: Similar to the previous case, we have $\Delta = -g_j - \alpha_{k_i i} + \alpha_{k_i j} = r_j - r_i - g_i$;
3. $n_i \geq d_i$: Given the chain reaction, the overall effect is equivalent to upgrading demand j with capacity i. Then, $\Delta = -g_j + \alpha_{ij} = r_j - u_i$.

To summarize, if $l \geq j$, we have

$$\Delta = \begin{cases} r_j - r_i, & \text{if } n_i < 0 \\ r_j - r_i - g_i, & \text{if } 0 \leq n_i < d_i \\ r_j - u_i, & \text{if } n_i \geq d_i. \end{cases} \tag{A.36}$$

Note that (A.36) also holds for $\bar{\Delta}$ if $\bar{l} \geq j$. Therefore, $\Delta = \bar{\Delta}$ when $j \leq l \leq \bar{l}$.

Next, we compare Δ and $\bar{\Delta}$ when both $l < j$ and $\bar{l} < j$. We categorize different situations based on Case 1, Case 2 and Case 3 as follows:

1. Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: Notice that class l here is similar to class j in (A.36) in Case 4. Then, we have

$$\delta = \begin{cases} g_i - g_j + \alpha_{k_i l} - \alpha_{k_i i} = g_l - g_j - r_i + r_l, & \text{if } n_i < 0 \\ -g_j + \alpha_{k_i l} - \alpha_{k_i i} = -g_j - g_i + g_l - r_i + r_l, & \text{if } 0 \le n_i < d_i \\ -g_j + \alpha_{il}, & \text{if } n_i \ge d_i, \end{cases} \tag{A.37}$$

and

$$\bar{\delta} = \begin{cases} g_i - g_j + \alpha_{\bar{k}_i \bar{l}} - \alpha_{\bar{k}_i i} = g_{\bar{l}} - g_j - r_i + r_{\bar{l}}, & \text{if } n_i < 0 \\ -g_j + \alpha_{\bar{k}_i \bar{l}} - \alpha_{\bar{k}_i i} = -g_j - g_i + g_{\bar{l}} - r_i + r_{\bar{l}}, & \text{if } 0 \le n_i < d_i \\ -g_j + \alpha_{i\bar{l}}, & \text{if } n_i \ge d_i. \end{cases} \tag{A.38}$$

Thus, $\delta - \bar{\delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}}$.

Furthermore, $\bar{\mathbf{R}}' = \bar{\mathbf{R}}_{\bar{l}j}$ by the assumption of this case. We next show

$$\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}}). \tag{A.39}$$

From the assumption, there is no upgrade between classes $(1, \cdots, \bar{h} - 1)$ and $(\bar{h}, \cdots, N)$ when applying the PSR under $(\bar{\mathbf{N}}, \mathbf{D})$, whose result is $\bar{\mathbf{R}}$. Thus, the effective states of classes $(1, \cdots, \bar{h} - 1)$ in $\bar{\mathbf{R}}$ are the same as those in $\bar{\mathbf{N}}$ by Proposition A.5.2. Moreover, note that $\bar{h}$ is the highest class upgrading demand $\bar{l}$ by assumption. Without loss of generality, we assume $\bar{h}$ is also the lowest class upgrading demand $\bar{l}$, then the effective state of classes $(\bar{h}, \cdots, \bar{l} - 1)$ in $\bar{\mathbf{N}}$ is $((\bar{\mathbf{R}})_{\bar{h}} + y_{\bar{h}\bar{l}}, 0, \cdots, 0)$, where $y_{\bar{h}\bar{l}}$ is the upgrade between classes $\bar{h}$ and $\bar{l}$ under initial state $(\bar{\mathbf{N}}, \mathbf{D})$. Thus, classes $\bar{h}$ and $\bar{l}$ correspond to classes k and i in Lemma A.5.3, which proves (A.39).

Similarly, since $\mathbf{R}' = \mathbf{R}_{lj}$ in this case, there is

$$\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{R}_{lj}) - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{N}_{lj}) - \Theta^{t+1}(\mathbf{N}). \tag{A.40}$$

Moreover,

$$\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{N}_{lj}) - \Theta^{t+1}(\mathbf{N}) \geq \Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}) \quad \text{(A.41)}$$

from the induction assumption.

To complete the proof in this case, from Lemma A.1.2 and the fact that $l \leq \bar{l}$, there is

$$\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) \geq r_{\bar{l}} - r_l,$$

which implies $\Delta - \bar{\Delta} = \delta - \bar{\delta} + r_{\bar{l}} - r_l$ by (A.39) and (A.41). Since $\delta - \bar{\delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}}$ by (A.37) and (A.38), we have $\Delta - \bar{\Delta} = g_l - g_{\bar{l}} \geq 0$.

In the remaining cases, we apply similar arguments to prove (A.35). For simplicity, we will omit some details and only present the primary results.

2. Case 2 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: We have

$$\delta = \begin{cases} g_i - g_j + (r_l + g_l - r_i - g_i) - \alpha_{hl} = -g_j - r_i + u_h, & \text{if } n_i < 0 \\ -g_j + (r_l + g_l - r_i - g_i) - \alpha_{hl} = -g_j - g_i - r_i + u_h, & \text{if } 0 \leq n_i < d_i \\ -g_j + \alpha_{il} - \alpha_{hl} = -g_j - u_i + u_h, & \text{if } n_i \geq d_i \end{cases} \quad \text{(A.42)}$$

and

$$\bar{\delta} = \begin{cases} g_i - g_j + (r_{\bar{l}} + g_{\bar{l}} - r_i - g_i) - \alpha_{\bar{h}\bar{l}} = -g_j - r_i + u_{\bar{h}}, & \text{if } n_i < 0 \\ -g_j + (r_{\bar{l}} + g_{\bar{l}} - r_i - g_i) - \alpha_{\bar{h}\bar{l}} = -g_j - g_i - r_i + u_{\bar{h}}, & \text{if } 0 \leq n_i < d_i \\ -g_j + \alpha_{i\bar{l}} - \alpha_{\bar{h}\bar{l}} = -g_j - u_i + u_{\bar{h}}, & \text{if } n_i \geq d_i. \end{cases} \quad \text{(A.43)}$$

Note that $\delta - \bar{\delta} = u_h - u_{\bar{h}}$ in all cases.

(a) $\bar{l} = l$: From the assumption, all backlogged demands in classes $(i, \cdots, l)$, which are the same for both initial states $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, have been satisfied in period t. Meanwhile, $\mathbf{D}$ is the same for both initial states in period t. Thus, the total demands satisfied are the same for both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, and we have $\bar{h} = h \geq r$ or $h \leq \bar{h} \leq r$.

By assumption, $\mathbf{R}' = \mathbf{R}_{hj}$ and $\bar{\mathbf{R}}' = \bar{\mathbf{R}}_{\bar{h}j}$ in this case, then

$$\begin{aligned}
\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) &= \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}), \\
\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) &= \Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}).
\end{aligned} \tag{A.44}$$

Moreover, we define $\tilde{\mathbf{R}}$ as follows:

$$\tilde{\mathbf{R}} = \begin{cases} \mathbf{R} + \mathbf{e}_r, & \text{if } r < h \\ \mathbf{R} + \mathbf{e}_h, & \text{if } r \geq h. \end{cases}$$

Note that $\tilde{\mathbf{R}} = \bar{\mathbf{R}}$ from the definition. If $r < h$, given $(\mathbf{R})_{h+1,\cdots,j-1} = (\mathbf{R}')_{h+1,\cdots,j-1} \leq 0$, we have

$$\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}) \geq \Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\tilde{\mathbf{R}}) \tag{A.45}$$

from the induction assumption. On the other hand, if $r \geq h$, (A.45) still holds because of the concavity in Proposition 2.4.1.

Since $\tilde{\mathbf{R}} = \bar{\mathbf{R}}$ and $h \leq \bar{h}$, there is $\Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) \geq u_{\bar{h}} - u_h$ by Lemma A.1.1. Therefore, from (A.44) and (A.45), we have

$$\Delta - \bar{\Delta} \geq \delta - \bar{\delta} + \Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) \geq 0,$$

where $\delta - \bar{\delta} = u_h - u_{\bar{h}}$ by (A.42) and (A.43).

(b) $l < \bar{l}$: From the above discussion of Case 2, $\bar{\mathbf{R}}'$ has one more unit of capacity $\bar{h}$ than $\bar{\mathbf{R}}$ after the chain reaction. Note that class $\bar{h}$ would have upgraded demand $\bar{l}$ if there exists unmet demand $\bar{l}$ under $\bar{\mathbf{R}}'$, which implies that the expected value of such a unit of capacity $\bar{h}$ is smaller than $\alpha_{\bar{h}\bar{l}}$. Thus,

$$\begin{aligned}
&\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}) \\
=&\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}) + \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}) - \Theta^{t+1}(\bar{\mathbf{R}}) \\
\leq&\alpha_{\bar{h}\bar{l}} + \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}),
\end{aligned} \tag{A.46}$$

Moreover, similar to (A.39), we can apply Lemma A.5.3 to (A.46) as $(\bar{\mathbf{R}})_{\bar{h}+1,\cdots,\bar{l}} = 0$, then

$$\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}) = \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}})$$

and

$$\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) \leq \alpha_{\bar{h}\bar{l}} + \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}}). \tag{A.47}$$

For initial state $(\mathbf{N}, \mathbf{D})$, since $l < \bar{l}$ and $(\mathbf{R})_{l+1} < 0$, after the chain reaction, there is an additional unit of capacity h in $\mathbf{R}'$ which can be used to upgrade demand $l+1$. However, upgrading demand $l+1$ by capacity h is not optimal under $\mathbf{R}'$, i.e., the expected value of such a unit of capacity h is higher than $\alpha_{h,l+1}$. Then,

$$\begin{aligned} &\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}) \\ =&\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}_{h,l+1}) + \Theta^{t+1}(\mathbf{R}_{h,l+1}) - \Theta^{t+1}(\mathbf{R}) \\ \geq&\alpha_{h,l+1} + \Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}_{h,l+1}). \end{aligned} \tag{A.48}$$

From the definition of $\tilde{\mathbf{R}}$ and the induction assumption, we have

$$\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}_{h,l+1}) \geq \Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\tilde{\mathbf{R}}_{h,l+1})$$

because $(\mathbf{R})_{l+2,\cdots,j-1} = (\tilde{\mathbf{R}})_{l+2,\cdots,j-1} \leq 0$. Moreover, $(\tilde{\mathbf{R}})_{h+1,\cdots,l} = (\mathbf{R})_{h+1,\cdots,l} = 0$ by the assumption of this case, from Lemma A.5.3, we similarly have

$$\Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\tilde{\mathbf{R}}_{h,l+1}) = \Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j}) - \Theta^{t+1}(\bar{\mathbf{N}}).$$

Thus,

$$\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) \geq \alpha_{h,l+1} + \Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j}) - \Theta^{t+1}(\bar{\mathbf{N}}). \tag{A.49}$$

Given $l < \bar{l}$, we have $\Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) \geq r_{\bar{l}} - r_{l+1}$ from Lemma A.1.2. Since $\delta - \bar{\delta} = u_h - u_{\bar{h}}$ by (A.42) and (A.43), (A.47) and (A.49) imply that $\Delta \geq \bar{\Delta}$ as $g_{l+1} \geq g_{\bar{l}}$.

3. Case 3 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: Since $l \leq \bar{l}$, the same proof of "Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$" can be applied.

4. Case 1 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$: Note that (A.41) and (A.47) still hold, meanwhile, δ and $\bar{\delta}$ are given in (A.37) and (A.43), respectively. We have

$$\delta - (\bar{\delta} + \alpha_{\bar{h}\bar{l}}) = r_l + g_l - r_{\bar{l}} - g_{\bar{l}}$$

and

$$
\begin{aligned}
&\big(\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
=&\big(\Theta^{t+1}(\mathbf{R}_{lj}) - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
\geq&\big(\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) - \big(\alpha_{\bar{h}\bar{l}} + \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) \\
=&\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \alpha_{\bar{h}\bar{l}}.
\end{aligned}
$$

Given $l \leq \bar{l}$, $\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) \geq r_{\bar{l}} - r_l$ by Lemma A.1.2. Then, $\Delta \geq \bar{\Delta}$ since $g_l \geq g_{\bar{l}}$.

5. Case 1 for $\mathbf{N}$ and Case 3 for $\bar{\mathbf{N}}$: Note that $l \leq \bar{l} < \bar{l}+1$, the same proof of "Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$" can be applied.

6. Case 2 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$: In this case, class $\bar{l}$ in $(\bar{\mathbf{N}}, \mathbf{D})$ still has unmet demand while demand l is fully satisfied in $(\mathbf{N}, \mathbf{D})$ by assumption. From the induction assumption, upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$, thus $l < \bar{l}$.

   Given that (A.38, A.39, A.42, A.49) all hold, there is

$$
(\delta + \alpha_{h,l+1}) - \bar{\delta} = r_{l+1} + g_{l+1} - r_{\bar{l}} - g_{\bar{l}}.
$$

   Since $l + 1 \leq \bar{l}$,

$$
\begin{aligned}
&\big(\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
=&\big(\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
\geq&\big(\alpha_{h,l+1} + \Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) - \big(\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) \\
\geq&\alpha_{h,l+1} + r_{\bar{l}} - r_{l+1}
\end{aligned}
$$

   by Lemma A.1.2. Then, $\Delta \geq \bar{\Delta}$ since $g_{l+1} \geq g_{\bar{l}}$.

7. Case 2 for $\mathbf{N}$ and Case 3 for $\bar{\mathbf{N}}$: Note that $l \leq \bar{l} < \bar{l}+1$, the same proof of "Case 2 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$" can be applied.

8. Case 3 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$: To apply the same proof of "Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$", we only need to show $l+1 \leq \bar{l}$. Suppose to the contrary that $l+1 > \bar{l}$,

then $l \geq \bar{l}$. Note that upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$ by assumption. Recall the discussions about Case 1 and Case 3, there is unmet demand $\bar{l}$ remaining in $\bar{\mathbf{R}}'$, but all demand l has been satisfied in $\mathbf{R}$. This is a contradiction.

9. Case 3 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$: To apply the same proof of "Case 1 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$", we need to show $l+1 \leq \bar{l}$. Similar to the above discussion, we suppose $l+1 > \bar{l}$. Note that all demand l has been satisfied in $\mathbf{R}$ and some of the lower class demand $l+1$ is also satisfied in $\mathbf{R}'$. Meanwhile, the demand lower than class $\bar{l}$ is not upgraded under both $\bar{\mathbf{R}}$ and $\bar{\mathbf{R}}'$. This is contradiction.

To complete this proof, we need to consider the case when $l < j$ and $\bar{l} \geq j$, where $\bar{l} \geq j$ means $\mathbf{R}' = \mathbf{R}$ and (A.36) is true.

1. Case 1 for $\mathbf{N}$: From (A.40),

$$\Theta^{t+1}(\mathbf{R}_{lj}) - \Theta^{t+1}(\mathbf{R}) \geq r_j - r_l$$

   by Lemma A.1.2. Since $\bar{\Delta}$ is given in (A.36), we have $\Delta \geq \bar{\Delta}$ from δ in (A.37).

2. Case 2 for $\mathbf{N}$: From (A.48) and the fact $l+1 \leq j$, there is

$$\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}_{h,l+1}) \geq r_j - r_{l+1}$$

   by Lemma A.1.2. With $\bar{\Delta}$ in (A.36) and δ in (A.42), we have $\delta + \alpha_{h,l+1} + r_j + g_j - r_{l+1} - g_{l+1} = \bar{\Delta}$. Hence, $\Delta \geq \bar{\Delta}$.

3. Case 3 for $\mathbf{N}$: Note that $l+1 < j$ in this case. Then, the same proof of "Case 1 for $\mathbf{N}$" can be applied.

This completes the proof. $\square$

The next lemma states that the protection level p_{ij} $(1 \leq i < j \leq N)$ in period $T-1$ decease in the states of classes $(1, \cdots, i-1)$.

Lemma A.5.5 *Consider an N-class upgrading problem in period T with $(n_{i+1}, \cdots, n_j) \leq 0$. Let $\bar{\mathbf{N}} = \mathbf{N} + \epsilon \mathbf{e}_r$, where $1 \leq r < i$ and $\epsilon > 0$. Then,*

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}) \geq \Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}), \qquad \Delta_{ij}^{-+}\Theta^T(\mathbf{N}) \geq \Delta_{ij}^{-+}\Theta^T(\bar{\mathbf{N}}).$$

Proof. Following the notations in the proof of Lemma A.5.4, in this proof we only need to consider Case 1, Case 3 and Case 4 for Θ^T since the additional unit of capacity h ($\bar{h}$) will not be passed to the next period. Note that $\Delta = \delta$ and $\bar{\Delta} = \bar{\delta}$ since $\Theta^{T+1} \equiv 0$. Also, the protection levels are zero in period T.

Recall the similarity of Case 1 and Case 3. In the proof of Lemma A.5.4, we have shown that $l + 1 \leq \bar{l}$ if "Case 3 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$". Therefore, we only have three different cases in period T.

1. $j \leq l \leq \bar{l}$: Since (A.36) still holds, we have $\Delta = \bar{\Delta}$.
2. $l \leq \bar{l} < j$: From (A.37) and (A.38), there is $\Delta - \bar{\Delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}} \geq 0$ since $r_l \geq r_{\bar{l}}$ and $g_l \geq g_{\bar{l}}$.
3. $l < j \leq \bar{l}$: From (A.36) and (A.37), we have $\Delta - \bar{\Delta} = r_l + g_l - r_j - g_j > 0$ since $r_l > r_j$ and $g_l > g_j$.

Hence, the desired result holds in period T for any demand realization, which completes the proof. $\square$

With the previous two lemmas, we can prove the monotonicity result.

PROPOSITION 2.5.4 *The optimal protection level p_{ij} $(1 \leq i < j \leq N)$ in period t $(1 \leq t \leq T)$ are decreasing in $(n_1^t, \cdots, n_{i-1}^t)$.*

Proof. Given the definition of the protection level in (2.8), this proposition can be inductively proved using Lemmas A.5.4 and A.5.5. $\square$

Appendix B

Appendices: Upgrading, Product Differentiation, and Heterogeneous Consumers

Appendix B.1: Partition of Region R

The region $\mathbf{R} = \{(p_1, p_2) : \ 0 \le p_1 \le q_1, \ 0 \le p_2 \le q_2 + \delta\}$ is decomposed into subregions $\mathbf{R}_i$ $(i = 1, 2, 3)$ as follows:

$$
\begin{aligned}
\mathbf{R}_1 =& \{(p_1, p_2) \mid p_1 \le p_2 + q_1 - q_2 - \delta\} \cap \mathbf{R} \\
\mathbf{R}_2 =& \{(p_1, p_2) \mid p_2 + q_1 - q_2 - \delta < p_1 \le p_2 + q_1 - q_2\} \cap \mathbf{R} \\
\mathbf{R}_3 =& \{(p_1, p_2) \mid p_1 > p_2 + q_1 - q_2\} \cap \mathbf{R}.
\end{aligned}
\tag{B.1}
$$

And $\mathbf{R}_1$ and $\mathbf{R}_2$ are further decomposed into $\mathbf{R}_{1i}$ $(i = 1, 2, 3)$, $\mathbf{R}_{2i}$ $(i = 1, 2)$, where

$$
\begin{aligned}
\mathbf{R}_{11} =& \{(p_1, p_2) \mid \frac{p_1}{p_2} \ge \frac{q_1}{q_2}, \ p_1 \le p_2 + q_1 - q_2 - \delta\} \cap \mathbf{R}, \\
\mathbf{R}_{12} =& \{(p_1, p_2) \mid \frac{q_1}{q_2 + \delta} \le \frac{p_1}{p_2} < \frac{q_1}{q_2}, \ p_1 \le p_2 + q_1 - q_2 - \delta\} \cap \mathbf{R}, \\
\mathbf{R}_{13} =& \{(p_1, p_2) \mid \frac{p_1}{p_2} < \frac{q_1}{q_2 + \delta}, p_1 \le p_2 + q_1 - q_2 - \delta\} \cap \mathbf{R}, \\
\mathbf{R}_{21} =& \{(p_1, p_2) \mid \frac{p_1}{p_2} \ge \frac{q_1}{q_2}, \ p_2 + q_1 - q_2 - \delta < p_1 \le p_2 + q_1 - q_2\} \cap \mathbf{R}, \\
\mathbf{R}_{22} =& \{(p_1, p_2) \mid \frac{p_1}{p_2} < \frac{q_1}{q_2}, p_2 + q_1 - q_2 - \delta < p_1 \le p_2 + q_1 - q_2\} \cap \mathbf{R}.
\end{aligned}
\tag{B.2}
$$

Appendix B.2: Proofs

LEMMA 3.3.1 *The objective function $\pi(p_1, p_2)$ is continuous in $\mathbf{R}$. Moreover, $\pi(p_1, p_2)$ is continuously differentiable and jointly concave in (p_1, p_2) in $\mathbf{R}_i$ $(i = 1, 2, 3)$, respectively.*

Proof. First, for sub-regions $\mathbf{R}_{1i}$ $(i = 1, 2, 3)$, $\mathbf{R}_{2i}$ $(i = 1, 2)$ and $\mathbf{R}_3$, we define $\pi(p_1, p_2) = \pi_{ij}(p_1, p_2)$ if $(p_1, p_2) \in \mathbf{R}_{ij}$. Specifically,

$$
\begin{aligned}
\pi_{11}(p_1,p_2) =& p_1 \left(r \left(1 - \frac{p_1 - p_2}{q_1 - q_2 - \delta} \right) + (1-r) \left(1 - \frac{p_1 - p_2}{q_1 - q_2} \right) \right) \\
&+ p_2 \left(r \left(\frac{p_1 - p_2}{q_1 - q_2 - \delta} - \frac{p_2}{q_2 + \delta} \right) + (1-r) \left(\frac{p_1 - p_2}{q_1 - q_2} - \frac{p_2}{q_2} \right) \right), \\
\pi_{12}(p_1,p_2) =& p_1 \left(r \left(1 - \frac{p_1 - p_2}{q_1 - q_2 - \delta} \right) + (1-r) \left(1 - \frac{p_1}{q_1} \right) \right) + r p_2 \left(\frac{p_1 - p_2}{q_1 - q_2 - \delta} - \frac{p_2}{q_2 + \delta} \right), \\
\pi_{13}(p_1,p_2) =& p_1 \left(1 - \frac{p_1}{q_1} \right), \\
\pi_{21}(p_1,p_2) =& (1-r) p_1 \left(1 - \frac{p_1 - p_2}{q_1 - q_2} \right) + p_2 \left(r \left(1 - \frac{p_2}{q_2 + \delta} \right) + (1-r) \left(\frac{p_1 - p_2}{q_1 - q_2} - \frac{p_2}{q_2} \right) \right), \\
\pi_{22}(p_1,p_2) =& (1-r) p_1 \left(1 - \frac{p_1}{q_1} \right) + r p_2 \left(1 - \frac{p_2}{q_2 + \delta} \right), \\
\pi_{3}(p_1,p_2) =& p_2 \left(r \left(1 - \frac{p_2}{q_2 + \delta} \right) + (1-r) \left(1 - \frac{p_2}{q_2} \right) \right).
\end{aligned}
\tag{B.3}
$$

Note that all of the above functions are quadratic in (p_1, p_2), and it is easy to verify that they are continuously differentiable and jointly concave in (p_1, p_2) in their respective domains.

From the definition of $\pi(p_1, p_2)$ in (B.3), to prove the continuity of $\pi(p_1, p_2)$ in $\mathbf{R}$, it is straightforward to compare the values on the boundaries which separate two different regions, and the proof is omitted.

To conclude the proof, it is easy to show the following relations

$$\left(\frac{\partial\pi_{11}(p_1,p_2)}{\partial p_1}-\frac{\partial\pi_{12}(p_1,p_2)}{\partial p_1}\right)\Big|_{p_1=\frac{q_1}{q_2}p_2}=0,$$
$$\left(\frac{\partial\pi_{11}(p_1,p_2)}{\partial p_2}-\frac{\partial\pi_{12}(p_1,p_2)}{\partial p_2}\right)\Big|_{p_2=\frac{q_2}{q_1}p_1}=0,$$
$$\left(\frac{\partial\pi_{12}(p_1,p_2)}{\partial p_1}-\frac{\partial\pi_{13}(p_1,p_2)}{\partial p_1}\right)\Big|_{p_1=\frac{q_1}{q_2+\delta}p_2}=0,$$
$$\left(\frac{\partial\pi_{12}(p_1,p_2)}{\partial p_2}-\frac{\partial\pi_{13}(p_1,p_2)}{\partial p_2}\right)\Big|_{p_2=\frac{q_2+\delta}{q_1}p_1}=0,$$
$$\left(\frac{\partial\pi_{21}(p_1,p_2)}{\partial p_1}-\frac{\partial\pi_{22}(p_1,p_2)}{\partial p_1}\right)\Big|_{p_1=\frac{q_1}{q_2}p_2}=0,$$
$$\left(\frac{\partial\pi_{21}(p_1,p_2)}{\partial p_2}-\frac{\partial\pi_{22}(p_1,p_2)}{\partial p_2}\right)\Big|_{p_2=\frac{q_2}{q_1}p_1}=0,$$

which imply the continuously differentiability of $\pi(p_1,p_2)$ with respect to (p_1,p_2) in $\mathbf{R}_i$ $(i=1,2,3)$. $\square$

PROPOSITION 3.3.1 *Consider $x_1\le\frac{1}{2}$. The firm's optimal solution is determined by x_1 and the thresholds k^j $(j=1,\cdots,6)$ defined in (3.4). Specifically,*

Case 1. If $k^1<k^2$, then

$$(p_1^*,p_2^*)\in\begin{cases}\mathbf{R}_{21}, & \textit{if } x_1\in(0,k^1];\\ \mathbf{R}_{11}, & \textit{if } x_1\in(k^1,k^2];\\ \mathbf{R}_{12}, & \textit{if } x_1\in(k^2,\frac{1}{2}).\end{cases}$$

Case 2. If $k^1\ge k^2$, $k^3>k^4$ and $k^6\ge 0$: there exists a threshold $\bar{k}\in[k^4,k^3]$, then

$$(p_1^*,p_2^*)\in\begin{cases}\mathbf{R}_{21}, & \textit{if } x_1\in(0,\bar{k}];\\ \mathbf{R}_{12}, & \textit{if } x_1\in(\bar{k},\frac{1}{2}).\end{cases}$$

Case 3. If $k^1\ge k^2$, $k^3>k^4$ and $k^6<0$: there exists a threshold $\bar{k}\in[0\vee k^5,k^4]$, then

$$(p_1^*,p_2^*)\in\begin{cases}\mathbf{R}_{21}, & \textit{if } x_1\in(0,\bar{k}];\\ \mathbf{R}_{22}, & \textit{if } x_1\in(\bar{k},k^4];\\ \mathbf{R}_{12}, & \textit{if } x_1\in(k^4,\frac{1}{2}).\end{cases}$$

Case 4. If $k^1 \geq k^2$ and $k^3 \leq k^4$: there exists a threshold $\bar{k} \in [k^5, k^3]$, then

$$(p_1^*, p_2^*) \in \begin{cases} \mathbf{R}_{21}, & \textit{if } x_1 \in (0, \bar{k}]; \\ \mathbf{R}_{22}, & \textit{if } x_1 \in (\bar{k}, k^4]; \\ \mathbf{R}_{12}, & \textit{if } x_1 \in (k^4, \frac{1}{2}). \end{cases}$$

Proof. We first consider the optimal solution in each sub-regions $\mathbf{R}_i$ $(i = 1, 2, 3)$ without the capacity constraint x_1 and show that the global unconstrained optimum is an interior point in sub-region $\mathbf{R}_1$. From Lemma 3.3.1, the objective function $\pi(p_1, p_2)$ is jointly concave in each sub-regions and continuous in $\mathbf{R}_1$, the global optimum is in $\mathbf{R}_1$ if

1. The optimum in $\mathbf{R}_3$ is on the boundary between $\mathbf{R}_2$ and $\mathbf{R}_3$;
2. The optimum in $\mathbf{R}_2$ is on the boundary between $\mathbf{R}_1$ and $\mathbf{R}_2$;
3. The optimum in $\mathbf{R}_1$ is an interior solution.

Starting with sub-region $\mathbf{R}_3$, note that $\pi_3(p_1, p_2)$ only depends on p_2 and $d_1(p_1, p_2) = 0$ which clearly satisfies the capacity constraint. From the concavity property, the optimal $p_2 = \frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)}$. Without loss of generality, let $p_1 = q_1 - \frac{q_2(q_2+(1-2r)\delta)}{2(q_2+(1-r)\delta)}$ such that the optimal solution is on boundary $p_1 = p_2 + q_1 - q_2$, which is between sub-regions $\mathbf{R}_2$ and $\mathbf{R}_3$.

Next, we consider sub-region $\mathbf{R}_2$ by sequentially examining sub-regions $\mathbf{R}_{21}$ and $\mathbf{R}_{22}$. Since the objective function $\pi_{21}(p_1, p_2)$ is jointly concave, we consider the following solution

$$p_1 = \frac{q_1}{2} + \frac{q_2 r\delta}{2(q_2 + (1-r)\delta)}, \quad p_2 = \frac{q_2(q_2+\delta)}{2(q_2 + (1-r)\delta)},$$

which satisfies the first-order condition in $\mathbf{R}_{21}$. However, plugging the above solution, we have

$$\frac{p_1}{p_2} - \frac{q_1}{q_2} = -\frac{r\delta(q_2 - q_1)}{q_2(q_2+\delta)} < 0.$$

Since $\pi_{21}(p_1, p_2)$ is jointly concave in (p_1, p_2), the optimum in sub-region $\mathbf{R}_{21}$ is on $p_1 = \frac{q_1}{q_2} p_2$. Similarly, the first-order condition in $\mathbf{R}_{22}$ gives the following solution

$$p_1 = \frac{q_1}{2}, \quad p_2 = \frac{q_2 + \delta}{2},$$

which contradicts the assumption $\frac{p_1-p_2}{q_1-q_2-\delta} \geq 1$ in $\mathbf{R}_{21}$ and implies the optimum in sub-region $\mathbf{R}_{22}$ is on $p_1 = p_2 + q_1 - q_2 - \delta$. Since $\pi_{21}(p_1, p_2) = \pi_{22}(p_1, p_2)$ on $p_1 = \frac{q_1}{q_2}p_2$, the optimum of sub-region $\mathbf{R}_2$ is on the boundary $p_1 = p_2 + q_1 - q_2 - \delta$, which is between $\mathbf{R}_1$ and $\mathbf{R}_2$.

Last, since the objective function is continuous and jointly concave in sub-region $\mathbf{R}_1$, we only need to show that there exists an interior solution in sub-region $\mathbf{R}_1$ that satisfies the first-order condition. The first-order condition of $\pi_{12}(p_1, p_2)$ in sub-region $\mathbf{R}_{12}$ gives the solution $(p_1, p_2) = (\frac{q_1}{2}, \frac{q_2+\delta}{2})$, which is on the boundary $p_1 = \frac{q_1}{q_2+\delta}p_2$. And the above solution also satisfies the first-order condition of $\pi_{13}(p_1, p_2)$ in sub-region $\mathbf{R}_{13}$. Therefore, the interior solution $(p_1, p_2) = (\frac{q_1}{2}, \frac{q_2+\delta}{2})$ is the optimum in sub-region $\mathbf{R}_1$.

Because the objective function $\pi(p_1, p_2)$ is continuous in region $\mathbf{R}$, the above argument implies that $(p_1, p_2) = (\frac{q_1}{2}, \frac{q_2+\delta}{2})$ is the global optimum without the capacity constraint, where the demand $d_1(p_1, p_2) = \frac{1}{2}$, $d_2(p_1, p_2) = 0$ and the unconstrained optimal profit is $\frac{q_1}{4}$.

Now, we consider the impact of the capacity constraint x_1. When $x_1 < \frac{1}{2}$, the global unconstrained optimum $(p_1, p_2) = (\frac{q_1}{2}, \frac{q_2+\delta}{2})$ is not achievable. From the concavity of Lemma 3.3.1, the optimal solution in (3.3) is on the boundary where $d_1(p_1, p_2) = x_1$.

For each sub-region $\mathbf{R}_{ij}$ $((i,j) = (1,1),\ (1,2),\ (1,3),\ (2,1),\ (2,2))$, let $p_1 = h_{ij}(p_2)$ satisfy $d_1(h_{ij}(p_2), p_2) = x_1$, we define $F_{ij}(p_2) = \pi_{ij}(h_{ij}(p_2), p_2)$, each of which is concave in p_2 from Lemma 3.3.1. Let p_2^1 satisfy $h_{21}(p_2^1) = \frac{q_1}{q_2}p_2^1$, p_2^2 satisfy $h_{22}(p_2^2) = q_1 - q_2 - \delta + p_2^2$, p_2^3 satisfy $h_{12}(p_2^3) = \frac{q_1}{q_2+\delta}p_2^3$, and p_2^4 satisfy $h_{11}(p_2^4) = \frac{q_1}{q_2}p_2^4$. From the definition, we have $p_2^1 < p_2^2 < p_2^3$. Furthermore, $h_{21}(p_2^1) = h_{22}(p_2^1) = \frac{q_1}{q_2}p_2^1$, $h_{22}(p_2^2) = h_{12}(p_2^2) = q_1 - q_2 - \delta + p_2^2$, $h_{12}(p_2^3) = h_{13}(p_2^3) = \frac{q_1}{q_2+\delta}p_2^3$, and $h_{11}(p_2^4) = h_{12}(p_2^4) = \frac{q_1}{q_2}p_2^4$. In other words, the curve in which $d_1(p_1, p_2) = x_1$ is continuous across adjacent sub-regions.

Recall that

$$k^1 = \frac{(1-r)\delta}{q_1 - q_2},$$
$$k^2 = \frac{1}{2}\left(1 - \frac{\delta}{q_1 - q_2 - \delta}\sqrt{\frac{rq_1(q_1 - (q_2+\delta)(1-r))}{q_2(q_2 + (1-r)\delta)}}\right),$$
$$k^3 = \frac{(1-r)(q_2 + (1-2r)\delta)}{2(q_2 + (1-r)\delta)},$$
$$k^4 = \frac{(q_2+\delta)(1-r)}{2q_1},$$
$$k^5 = \frac{(q_2-\delta)(1-r)}{2q_2},$$
$$k^6 = \frac{1}{4}(q_2+\delta)\left(\frac{q_2(1-r)}{q_1}\left(\frac{q_2+\delta}{q_1} - 2\right) + \frac{q_2}{q_2 + (1-r)\delta} - r\right).$$

Specifically, we give the definition of each k^i $(i = 1, \cdots, 6)$:

1. k^1: we have $F_{11}(0) < q_1 - q_2 - \delta$ if and only if $x_1 > k^1$, i.e., the sub-region $\mathbf{R}_{11}$ is infeasible if and only if $x_1 < k^1$;
2. k^2: let p_2' satisfy $\frac{\partial}{\partial p_2}F_{11}(p_2) \mid_{p_2=p_2'}= 0$ and p_2'' satisfy $\frac{\partial}{\partial p_2}F_{12}(p_2) \mid_{p_2=p_2''}= 0$, we have $F_{11}(p_2') > F_{12}(p_2'')$ if and only if $x_1 < k^2$;
3. k^3: we have $\frac{\partial}{\partial p_2}F_{21}(p_2) \mid_{p_2=p_2^1}> 0$ if and only if $x_1 > k^3$;
4. k^4: we have $\frac{\partial}{\partial p_2}F_{22}(p_2) \mid_{p_2=p_2^2}> 0$ if and only if $x_1 > k^4$;
5. k^5: we have $\frac{\partial}{\partial p_2}F_{22}(p_2) \mid_{p_2=p_2^1}> 0$ if and only if $x_1 > k^5$;
6. k^6: let p_2' satisfy $\frac{\partial}{\partial p_2}F_{21}(p_2) \mid_{p_2=p_2'}= 0$ and p_2'' satisfy $\frac{\partial}{\partial p_2}F_{22}(p_2) \mid_{p_2=p_2''}= 0$, then
$$k^6 = (F_{21}(p_2') - F_{22}(p_2'')) \mid_{x_1=k^4}.$$

From the definition, it can be shown that:

1. $k^3 > k^5$ since
$$k^3 - k^5 = \frac{(1-r)^2\delta(q_2+\delta)}{2q_2(q_2 + (1-r)\delta)} > 0;$$
2. $k^4 > k^5$. Otherwise, consider $k^4 < x_1 \le k^5$, we have $\frac{\partial}{\partial p_2}F_{22}(p_2) \mid_{p_2=p_2^2}> 0$ and $\frac{\partial}{\partial p_2}F_{22}(p_2) \mid_{p_2=p_2^1}< 0$, which is a contradiction because $F_{22}(p_2)$ is concave in p_2 and $p_2^1 < p_2^2$;

Therefore, there is either $k^5 < k^4 < k^3$ or $k^5 < k^3 < k^4$ for all k^1 and k^2. Furthermore, we only need to consider the constrained solutions in sub-regions $\mathbf{R}_{11}$, $\mathbf{R}_{12}$, $\mathbf{R}_{21}$, and $\mathbf{R}_{22}$. This is because F_{13} is fixed in p_2, which implies $F_{13}(p_2^3) = F_{12}(p_2^3)$ since the objective function $\pi(p_1, p_2)$ is continuous in region $\mathbf{R}$.

Next, we analyze each case as follows.

Case 1. If $k^1 < k^2$: we have $k^2 < \frac{1}{2}$ from the definition.

(a) If $x_1 \in [k^2, \frac{1}{2})$: since $x_1 > k^1$, the optimal solution lies on either $F_{11}(p_2)$ or $F_{12}(p_2)$. Furthermore, let p_2' be the solution of $h_{11}(p_2) = \frac{q_1}{q_2}p_2$, which is the same as the solution of $h_{12}(p_2) = \frac{q_1}{q_2}p_2$. There is

$$\frac{\partial}{\partial p_2}F_{11}\mid_{p_2=p_2'}< 0, \qquad \frac{\partial}{\partial p_2}F_{12}\mid_{p_2=p_2'}> 0,$$

which implies the first-order condition can be satisfied for both $F_{11}(p_2)$ and $F_{12}(p_2)$. From the definition of k^2, the optimum is in sub-region $\mathbf{R}_{12}$, where

$$(p_1^*, p_2^*) = \left(\frac{q_1((1-x_1)(q_1-q_2-\delta)+\frac{1}{2}r(q_2+\delta))}{q_1-(1-r)(q_2+\delta)},\ \frac{q_2+\delta}{2}\right)$$

and

$$d_1^* = x_1,\ d_2^* = \frac{rq_1(1-2x_1)}{2(q_1-(1-r)(q_2+\delta))},\ \pi^* = \frac{q_1(4x_1(1-x_1)(q_1-q_2-\delta)+r(q_2+\delta))}{4(q_1-(1-r)(q_2+\delta))};$$

(b) If $x_1 \in [k^1, k^2)$: since $x_1 < k^2$, similar to the previous scenario, the optimum is in sub-region $\mathbf{R}_{11}$, where

$$(p_1^*, p_2^*) = \left(\frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)} + \frac{(1-x_1)(q_1-q_2)(q_1-q_2-\delta)}{q_1-q_2-(1-r)\delta},\ \frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)}\right)$$

and

$$d_1^* = x_1,\ d_2^* = \frac{1}{2}-x_1,\ \pi^* = \frac{q_2(q_2+\delta)}{4(q_2+(1-r)\delta)}+\frac{x_1(1-x_1)(q_1-q_2)(q_1-q_2-\delta)}{q_1-q_2-(1-r)\delta};$$

(c) If $x_1 \in (0, k^1)$: in this case, we have $k^3 > k^1$. Moreover,

$$\left(\tfrac{\partial}{\partial p_2}F_{12}(p_2)\mid_{p_2=p_2^2}\right)\left(\tfrac{\partial}{\partial p_2}F_{22}(p_2)\mid_{p_2=p_2^2}\right) > 0,$$

i.e., partial derivatives $\frac{\partial}{\partial p_2}F_{12}(p_2)\mid_{p_2=p_2^2}$ and $\frac{\partial}{\partial p_2}F_{22}(p_2)\mid_{p_2=p_2^2}$ have the same sign.

i. If $k^5 < k^4 < k^3$: We have $\frac{\partial}{\partial p_2}F_{21}\mid_{p_2=0}> 0$. And it can be shown that $k^4 < k^1$, thus, we sequentially consider intervals $(0, k^5]$, $(k^5, k^4]$, and (k^4, k^1) for x_1:

A. If $x_1 \in (0, k^5]$: Since $\frac{\partial}{\partial p_2}F_{21}\mid_{p_2=p_2^1}< 0$, $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^1}< 0$, and $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^2}< 0$, the optimal solution is in sub-region $\mathbf{R}_{21}$. The optimal p_2^* can be solved from $\frac{\partial}{\partial p_2}F_{21}(p_2) = 0$, and the optimal $p_1^* = h_{21}(p_2^*)$. Specifically,

$$(p_1^*, p_2^*) = \left(q_1 - \frac{(q_1-q_2)x_1}{1-r} - \frac{q_2(q_2+(1-2r)\delta)}{2(q_2+(1-r)\delta)},\ \frac{q_2(q_2+\delta)}{2(q_2+(1-r)\delta)}\right)$$

and

$$d_1^* = x_1,\ d_2^* = \frac{1}{2} - x_1,\ \pi^* = (q_1-q_2)x_1\left(1-\frac{x_1}{1-r}\right) + \frac{q_2(q_2+\delta)}{4\,(q_2+(1-r)\delta)};$$

B. If $x_1 \in (k^5, k^4]$: Since $\frac{\partial}{\partial p_2}F_{21}\mid_{p_2=p_2^1}< 0$, $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^1}> 0$, and $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^2}< 0$, we need to compare F_{21}^*, the value of the first-order solution from $F_{21}(p_2)$, with F_{22}^*, the value of the first-order solution from $F_{22}(p_2)$.

It can be shown that $F_{21}^* - F_{22}^*$ is convex decreasing in x_1, and $F_{21}^* - F_{22}^* > 0$ when $x_1 = k^4$. Therefore, the optimal solution is also in sub-region $\mathbf{R}_{21}$;

C. If $x_1 \in (k^4, k^1)$: Since $\frac{\partial}{\partial p_2}F_{21}\mid_{p_2=p_2^1}< 0$, $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^1}> 0$, $\frac{\partial}{\partial p_2}F_{22}\mid_{p_2=p_2^2}> 0$, and $\frac{\partial}{\partial p_2}F_{12}\mid_{p_2=p_2^3}< 0$, we need to compare F_{21}^*, the value of the first-order solution from $F_{21}(p_2)$, with F_{12}^*, the value of the first-order solution from $F_{12}(p_2)$.

Similarly, we can show that $F_{21}^* - F_{12}^*$ is convex decreasing in x_1, and $F_{21}^* - F_{12}^* > 0$ when $x_1 = k^1$, therefore, the optimal solution is still in sub-region $\mathbf{R}_{21}$.

ii. If $k^5 < k^3 < k^4$: in this case, we have $k^1 < k^4$, and we need to sequentially consider intervals $(0, k^5]$ and (k^5, k^1) for x_1:

A. If $x_1 \in (0, k^5]$: this is similar to the case when $x_1 \in (0, k^5]$ and $k^5 < k^4 < k^3$, and the optimal solution is in sub-region $\mathbf{R}_{21}$;

B. If $x_1 \in (k^5, k^1)$: this is similar to the case when $x_1 \in (k^5, k^4)$ and $k^5 < k^4 < k^3$, and the optimal solution is in sub-region $\mathbf{R}_{21}$.

Therefore, the optimal solution is always in sub-region $\mathbf{R}_{21}$ when $x_1 \in (0, k^1)$.

Case 2. If $k^1 \geq k^2$, $k^3 > k^4$ and $k^6 \geq 0$: There is $k^4 < k^1$ when $k^1 > k^2$.

(a) If $x_1 \in [k^1, \frac{1}{2})$: Since $x_1 \geq k^1 \geq k^2$, the optimal solution is in sub-region $\mathbf{R}_{12}$;

(b) If $x_1 \in (0, k^1)$:

i. If $k^3 < k^1$: We sequentially consider intervals $(0, k^5]$, (k^5, k^4), $[k^4, k^3]$, and (k^3, k^1) for x_1:

A. $x_1 \in (0, k^5]$: Similar to the case when $x_1 \in (0, k^5]$, $k^5 < k^4 < k^3$ and $k^1 < k^2$, we can show that the optimal solution is in sub-region $\mathbf{R}_{21}$;

B. $x_1 \in (k^5, k^4)$: Similar to the case when $x_1 \in (k^5, k^4]$, $k^5 < k^4 < k^3$ and $k^1 < k^2$, we compare F^*_{21}, the value of the first-order solution from $F_{21}(p_2)$, with F^*_{22}, the value of the first-order solution from $F_{22}(p_2)$. As $F^*_{21} - F^*_{22}$ is convex decreasing in $x_1 \in (k^5, k^4)$, and $k^6 = (F^*_{21} - F^*_{22}) \mid_{x_1 = k^4} \geq 0$, there is $F^*_{21} \geq F^*_{22}$, and the optimal solution is in sub-region $\mathbf{R}_{21}$;

C. $x_1 \in [k^4, k^3]$: Similarly, we need to compare F^*_{21} with F^*_{12}, which is the value of the first-order solution from $F_{12}(p_2)$.
Since $F^*_{21} - F^*_{12}$ is decreasing in $x_1 \in [k^4, k^3]$, and

$$(F^*_{21} - F^*_{22}) \mid_{x_1 = k^4} = (F^*_{21} - F^*_{12}) \mid_{x_1 = k^4} = k^6 \geq 0,$$

there exists a threshold $\bar{k} \in [k^4, k^3]$ such that $F^*_{21} - F^*_{12} \geq 0$ if $x_1 \in [k^4, k^3] \cap (-\infty, \bar{k}]$, which means the optimal solution is in sub-region $\mathbf{R}_{21}$, and $F^*_{21} - F^*_{12} < 0$ if $x_1 \in [k^4, k^3] \cap (\bar{k}, \infty)$, which means the optimal solution is in sub-region $\mathbf{R}_{12}$;

D. $x_1 \in (k^3, k^1)$: Since $x_1 > k^3$, we have $\frac{\partial}{\partial p_2} F_{21} \mid_{p_2=p_2^1} > 0$, $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2=p_2^1} > 0$, and $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2=p_2^2} > 0$. This implies that the optimal solution is in sub-region $\mathbf{R}_{12}$.

ii. If $k^3 \geq k^1$: The result is the same as the case $k^3 < k^1$ since we can still sequentially consider intervals $(0, k^5]$, (k^5, k^4), and $[k^4, k^1)$ for x_1.

Therefore, there exists a threshold $\bar{k} \in [k^4, k^3]$, then the optimal solution is in sub-region $\mathbf{R}_{21}$ if $x_1 \in (0, \bar{k}]$, and the optimal solution is in sub-region $\mathbf{R}_{12}$ if $x_1 \in (\bar{k}, \frac{1}{2})$.

Case 3. If $k^1 \geq k^2$, $k^3 > k^4$ and $k^6 < 0$: Similar to Case 2, we sequentially consider intervals $(0, k^5)$, $[k^5, k^4]$, $(k^4, k^1]$, and $(k^1, \frac{1}{2})$. Since $k^6 = (F_{21}^* - F_{22}^*) \mid_{x_1=k^4} < 0$, the analysis of intervals $[k^5, k^4]$ and $(k^4, k^1]$ is slightly different from that in Case 2. Specifically,

(a) There exists a threshold $\bar{k} \in [0 \vee k^5, k^4]$, the optimal solution is in sub-region $\mathbf{R}_{21}$ if $x_1 \in [k^5, \bar{k}]$ and in sub-region $\mathbf{R}_{22}$ if $x_1 \in (\bar{k}, k^4]$. Specifically, the optimal solution in sub-region $\mathbf{R}_{22}$ is

$$(p_1^*, p_2^*) = \left(q_1(1 - \frac{x_1}{1-r}),\ \frac{q_2 + \delta}{2} \right)$$

and

$$d_1^* = x_1,\ d_2^* = \frac{r}{2},\ \pi^* = \frac{r(q_2 + \delta)}{4} + q_1 x_1 (1 - \frac{x_1}{1-r});$$

(b) Since

$$(F_{21}^* - F_{22}^*) \mid_{x_1=k^4} = (F_{21}^* - F_{12}^*) \mid_{x_1=k^4} = k^6 < 0$$

for all $x_1 \in (k^4, k^3]$, the optimal solution is in sub-region $\mathbf{R}_{12}$. Meanwhile, for all k^3, the optimal solution is in sub-region $\mathbf{R}_{12}$ for $x_1 \in (k^3 \vee k^1, k^1]$ from our previous analysis. Thus, the optimal solution is in sub-region $\mathbf{R}_{12}$ if $x_1 \in (k^4, k^1]$.

Therefore, there exists a threshold $\bar{k} \in [0 \vee k^5, k^4]$, then the optimal solution is in sub-region $\mathbf{R}_{21}$ if $x_1 \in (0, \bar{k}]$, in sub-region $\mathbf{R}_{22}$ if $x_1 \in (\bar{k}, k^4]$, and in sub-region $\mathbf{R}_{12}$ if $x_1 \in (k^4, \frac{1}{2})$.

Case 4. If $k^1 \geq k^2$ and $k^3 \leq k^4$: In this case, $k^5 < k^3 < k^4 < k^1$, and we sequentially consider intervals $(0, k^5)$, $[k^5, k^3]$, $(k^3, k^4]$, $(k^4, k^1]$, and $(k^1, \frac{1}{2})$ for x_1.

(a) If $x_1 \in (0, k^5)$: From the analysis in $x_1 \in (0, k^5]$, $k^5 < k^4 < k^3$ and $k^1 < k^2$, the optimal solution is in sub-region $\mathbf{R}_{21}$;

(b) If $x_1 \in [k^5, k^3]$: From the definition, we have $\frac{\partial}{\partial p_2} F_{21} \mid_{p_2 = p_2^1} < 0$, $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^1} > 0$, and $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^2} < 0$, we need to compare F_{21}^*, the value of the first-order solution from $F_{21}(p_2)$, with F_{22}^*, the value of the first-order solution from $F_{22}(p_2)$. Since,

$$
\begin{aligned}
(F_{21}^* - F_{22}^*) \mid_{x_1 = k^5} &= \frac{(1-r)^2 \delta^2 (q_2 + \delta)}{4 q_2 (q_2 + (1-r)\delta)} > 0, \\
(F_{21}^* - F_{22}^*) \mid_{x_1 = k^3} &= -\frac{(1-r)^2 \delta^2 r (q_2 + \delta)}{4 (q_2 + (1-r)\delta)^2} < 0,
\end{aligned}
$$

and $F_{21}^* - F_{22}^*$ is convex decreasing in $[k^5, k^3]$, there exists a threshold $\bar{k} \in [k^5, k^3]$ such that the optimal solution is in sub-region $\mathbf{R}_{21}$ if $x_1 \in [k^5, \bar{k}]$ and in sub-region $\mathbf{R}_{22}$ if $x_1 \in (\bar{k}, k^3]$;

(c) If $x_1 \in (k^3, k^4]$: Since $\frac{\partial}{\partial p_2} F_{21} \mid_{p_2 = p_2^1} > 0$, $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^1} > 0$, and $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^2} < 0$, the optimal solution is in sub-region $\mathbf{R}_{22}$;

(d) If $x_1 \in (k^4, k^1]$: We have $\frac{\partial}{\partial p_2} F_{21} \mid_{p_2 = p_2^1} > 0$, $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^1} > 0$, and $\frac{\partial}{\partial p_2} F_{22} \mid_{p_2 = p_2^2} > 0$. Moreover,

$$
\frac{\partial}{\partial p_2} F_{12} \mid_{p_2 = p_2^3} = -\frac{q_1 r (1 - 2x_1)}{q_1 - (1-r)(q_2 + \delta)} < 0.
$$

Thus, there exists a solution satisfying the first-order condition of F_{12}, and the optimal solution is in sub-region $\mathbf{R}_{12}$;

(e) If $x_1 \in (k^1, \frac{1}{2})$: Similar to the case when $x_1 \in [k^1, \frac{1}{2})$, $k^1 \geq k^2$, $k^3 > k^4$ and $k^6 \geq 0$, the optimal solution is in sub-region $\mathbf{R}_{12}$.

Therefore, there exists a threshold $\bar{k} \in [k^5, k^3]$, then the optimal solution is in $\mathbf{R}_{21}$ if $x_1 \in (0, \bar{k}]$, in sub-region $\mathbf{R}_{22}$ if $x_1 \in (\bar{k}, k^4]$, and in sub-region $\mathbf{R}_{12}$ if $x_1 \in (k^4, \frac{1}{2})$.

$\square$

Proposition 3.4.1 *For any capacity level* $x_1 \leq \frac{1}{2}$*, there exists a threshold* $\bar{\delta}^1 \in (0, q_1 - q_2)$ *such that the optimal solution of* (3.3) *belongs to* $\mathbf{R}_{11}$ *if and only if* $\delta \in [0, \bar{\delta}^1]$.

Proof. First, k^1 is increasing in δ, and k^2 is decreasing in δ since

$$\frac{\partial k^1}{\partial \delta} = \frac{1-r}{q_1 - q_2} > 0,$$
$$\frac{\partial k^2}{\partial \delta} = \frac{z(q_1) q_1 r \delta}{4 q_2 \delta (q_1 - q_2 - \delta)^2 (q_2 + (1-r)\delta)^2 \sqrt{\frac{q_1 r (q_1 - (1-r)(q_2+\delta))}{q_2 (q_2 + (1-r)\delta)}}} < 0,$$

where

$$z(q_1) = q_1^2(-(\delta + 2q_2 - \delta r)) + q_1 \left(-2q_2^2(r-2) + \delta q_2 \left(r^2 - 6r + 5\right) + \delta^2 \left(2r^2 - 3r + 1\right)\right) + q_2(r-1)(\delta + q_2)(2q_2 - \delta(r-2)).$$

In particular, $z(q_1)$ is concave in q_1 as

$$\frac{\partial^2 z(q_1)}{\partial q_1^2} = -4q_2 + 2(-1+r)\delta < 0.$$

Since

$$\frac{\partial z(q_1)}{\partial q_1} \mid_{q_1 = q_2 + \delta} = -2q_2^2 r + \delta q_2((r-4)r - 1) + \delta^2 (r-1)(2r+1) \leq 0$$

and

$$z(q_2 + \delta) = -2\delta r(\delta + q_2)(\delta + q_2 - \delta r) < 0,$$

we have $z(q_1) < 0$ for $q_1 > q_2 + \delta$ and $\frac{\partial k^2}{\partial \delta} \leq 0$.

Next, we have

$$k^1 \mid_{\delta=0} = 0, \qquad k^2 \mid_{\delta=0} = \frac{1}{2}.$$

From the proof of Proposition 3.3.1, the optimal solution is in sub-region $\mathbf{R}_{11}$ if and only if $k^1 \leq x_1 \leq k^2$. As δ increases from 0 to $q_1 - q_2$, k^1 is increasing from 0, and k^2 is decreasing from $\frac{1}{2}$. Let $\bar{\delta}_1^1$ be the solution to $k^1 = x_1$ with respect to δ, and $\bar{\delta}_2^1$ the solution to $k^2 = x_1$ with respect to δ. Then, we can define $\bar{\delta}^1 = \min(\bar{\delta}_1^1, \bar{\delta}_2^1)$ such that the optimal solution of (3.3) belongs to $\mathbf{R}_{11}$ if and only if $\delta \in [0, \bar{\delta}^1]$. □

Lemma 3.4.1 *The firm's optimal profit* π_{11}^* *in* (3.5) *is concave in* δ, π_{12}^* *in* (3.6) *is convex and increasing in* δ, π_{21}^* *in* (3.7) *is concave and increasing in* δ, *and* π_{22}^* *in* (3.8) *is linearly increasing in* δ.

Proof. π_{11}^* in (3.5) is concave in δ since

$$\frac{\partial^2 \pi_{11}^*}{\partial \delta^2} = -\frac{1}{2} r(1-r)\left(\frac{4(q_1-q_2)^2 x_1(1-x_1)}{(q_1-q_2-\delta+r\delta)^3} + \frac{q_2^2}{(q_2+(1-r)\delta)^2}\right) \leq 0.$$

π_{12}^* in (3.6) is convex and increasing in δ since

$$\frac{\partial \pi_{12}^*}{\partial \delta} = \frac{r q_1^2(1-2x_1)^2}{4(q_1-(1-r)(q_2+\delta))^2} \geq 0$$
$$\frac{\partial^2 \pi_{12}^*}{\partial \delta^2} = \frac{r(1-r) q_1^2(1-2x_1)^2}{2(q_1-(1-r)(q_2+\delta))^3} \geq 0.$$

π_{21}^* in (3.7) is concave and increasing in δ since

$$\frac{\partial \pi_{21}^*}{\partial \delta} = \frac{r q_2^2}{4(q_2+(1-r)\delta)^2} \geq 0$$
$$\frac{\partial^2 \pi_{21}^*}{\partial \delta^2} = -\frac{r(1-r) q_2^2}{2(q_2+(1-r)\delta)^3} \leq 0.$$

π_{22}^* in (3.8) is linearly increasing in δ since $\frac{\partial \pi_{22}^*}{\partial \delta} = \frac{r}{4}$. □

Proposition 3.4.2 *If* $x_1 \in (0, \frac{1}{2})$, *then*

1. $\frac{\partial}{\partial \delta}\pi^*(\delta) \mid_{\delta=0} > 0$;
2. *There exists* $\bar{\delta}^2 \in (0, \bar{\delta}^1)$ *and a threshold* $k^* < \frac{1}{5}$ *such that* $\pi^*(\delta)$ *is decreasing in* $\delta \in [\bar{\delta}^2, \bar{\delta}^1]$ *if and only if* $x_1 > k^*$, *and* $\pi^*(\delta)$ *is increasing in* δ *otherwise. Moreover, the threshold* k^* *does not depend on* x_1 *and* r.

Proof. From Lemma 3.4.1, in order to prove $\frac{\partial}{\partial \delta}\pi^*(\delta) \mid_{\delta=0} > 0$, we only need to verify π_{11}^* at $\delta = 0$, which can be seen from

$$\frac{\partial}{\partial \delta}\pi_{11}^* \mid_{\delta=0} = \frac{r}{4}(1-2x_1)^2 > 0.$$

Next, we define $\bar{\delta}^2$ as the solution of the first-order condition $\frac{\partial \pi_{11}^*}{\partial \delta} = 0$. Since π_{11}^* is concave in δ and $\frac{\partial}{\partial \delta}\pi_{11}^* \mid_{\delta=0} > 0$, then $\bar{\delta}^2 > 0$.

Recall the definition of $\bar{\delta}^1$, which is the minimum between $\bar{\delta}_1^1$, the solution to $k^1 = x_1$ with respect to δ, and $\bar{\delta}_2^1$, the solution to $k^2 = x_1$ with respect to δ. It can be shown that there exists a threshold $k^* < \frac{1}{5}$ such that $\bar{\delta}^2 < \bar{\delta}_1^1$ if and only if $x_1 > k^*$, in which case the optimal solution is in sub-region $\mathbf{R}_{11}$ and the optimal value π_{11}^* is decreasing in δ. Note that it is possible that $\bar{\delta}^1 = \bar{\delta}_2^1 < \bar{\delta}^2$, which implies that $[\bar{\delta}^2, \bar{\delta}^1]$ is an empty set.

If $x_1 < k^*$, we have $\bar{\delta}^2 > \bar{\delta}_1^1$, which means the optimal solution is always increasing even if the optimal solution is in sub-region $\mathbf{R}_{11}$.

This concludes the proof. □

LEMMA 3.4.2 *Suppose (p_1^*, p_2^*) is in region $\mathbf{R}_{11}$ as (3.5), then p_1^* is concave and decreasing in δ, and p_2^* is concave and increasing in δ.*

Proof. p_1^* is concave and decreasing in δ since

$$\begin{aligned}
\frac{\partial p_1^*}{\partial \delta} &= -\frac{r(1-x_1)(q_1-q_2)^2}{(q_1-q_2-\delta+r\delta)^2} + \frac{rq_2^2}{2(q_2+(1-r)\delta)} \leq 0, \\
\frac{\partial^2 p_1^*}{\partial \delta^2} &= -r(1-r)\left(\frac{2(1-x_1)(q_1-q_2)^2}{(q_1-q_2-\delta+r\delta)^3} + \frac{q_2^2}{(q_2+(1-r)\delta)^3}\right) \leq 0.
\end{aligned}$$

p_2^* is concave and increasing in δ since

$$\begin{aligned}
\frac{\partial p_2^*}{\partial \delta} &= \frac{rq_2^2}{2(q_2+(1-r)\delta)^2} \geq 0, \\
\frac{\partial^2 p_2^*}{\partial \delta^2} &= -\frac{r(1-r)q_2^2}{(q_2+(1-r)\delta)^3} \leq 0.
\end{aligned}$$

□

LEMMA 3.5.1 *π_{11}^* in (3.5) is convex in r, and π_{12}^* in (3.6) is concave and strictly increasing in r.*

Proof. π_{11}^* in (3.5) is convex in r since

$$\frac{\partial^2 \pi_{11}^*}{\partial r^2} = \frac{\delta^2}{2}\left(\frac{4(q_1-q_2)(q_1-q_2-\delta)x_1(1-x_1)}{(q_1-q_2-\delta+r\delta)^3} + \frac{q_2(q_2+\delta)}{(q_2+(1-r)\delta)^2}\right) \geq 0.$$

π_{12}^* in (3.6) is concave and strictly increasing in r since

$$\begin{aligned}
\frac{\partial \pi_{12}^*}{\partial r} &= \frac{q_1(q_2+\delta)(q_1-q_2-\delta)(1-2x_1)^2}{4(q_1-(1-r)(q_2+\delta))^2} > 0 \\
\frac{\partial^2 \pi_{12}^*}{\partial r^2} &= -\frac{q_1(q_1-q_2-\delta)(q_2+\delta)^2(1-2x_1)^2}{2(q_1-(1-r)(q_2+\delta))^3} \leq 0.
\end{aligned}$$

$\square$

PROPOSITION 3.5.1 *Suppose* $\frac{\delta}{q_1-q_2} < \frac{1}{2}$. *There exists* $\bar{r} \in (0,1)$ *and a threshold* $\tilde{k} \in [\frac{\delta}{q_1-q_2}, \frac{1}{2})$ *such that:*

1. *If* $x_1 \in \left(\tilde{k}, \frac{1}{2}\right)$, *then the firm's optimal profit is decreasing in* $r \in [0, \bar{r}]$ *and increasing in* $r \in (\bar{r}, 1]$.

2. *If* $x_1 \in \left[\frac{\delta}{q_1-q_2}, \tilde{k}\right]$, *then the firm's optimal profit is increasing in* $r \in [0,1]$.

Proof. First, we have both k^1 and k^2 are decreasing in r since

$$\begin{aligned}
\frac{\partial k^1}{\partial r} &= -\frac{\delta}{q_1-q_2} < 0; \\
\frac{\partial k^2}{\partial r} &= -\frac{\delta(q_2+\delta)(q_1+q_2(-1+2r)-(1-r)^2\delta)}{4r(q_1-(q_2+\delta))(q_2+(1-r)\delta)}\sqrt{\frac{rq_1}{q_2(q_2+(1-r)\delta)(q_1-(1-r)(q_2+\delta))}} \leq 0,
\end{aligned}$$

where the second inequality is from

$$\begin{aligned}
q_1+q_2(-1+2r)-(1-r)^2\delta &= q_1-q_2-\delta+2rq_2+2r\delta-r^2\delta \\
&\geq q_1-q_2-\delta+2rq_2+2r\delta-r\delta = q_1-q_2-\delta+2rq_2+r\delta \geq 0.
\end{aligned}$$

Furthermore,

$$k^1 \mid_{r=0} = \frac{\delta}{q_1-q_2} > 0, \qquad k^2 \mid_{r=0} = \frac{1}{2},$$

and

$$\frac{\partial k^1}{\partial r} \mid_{r=0} < 0, \text{ if and only if } x_1 > \frac{1}{2}\left(1-\sqrt{\frac{\delta q_1}{(q_1-q_2)(q_2+\delta)}}\right).$$

If $\frac{\delta}{q_1-q_2} < \frac{1}{2}$, let $\tilde{k} = \min\left(\frac{1}{2}\left(1-\sqrt{\frac{\delta q_1}{(q_1-q_2)(q_2+\delta)}}\right), \frac{\delta}{q_1-q_2}\right)$. Then, $\tilde{k} \in [\frac{\delta}{q_1-q_2}, \frac{1}{2})$. Moreover, let $\bar{r}_1$ be the threshold that $k^2(\bar{r}_1) = x_1$, and $\bar{r}_2$ be the solution of the first-order condition $\frac{\partial k^1}{\partial r} = 0$. Since

$$\frac{\partial k^1}{\partial r} \mid_{r=1} = \frac{\delta q_1 + q_2(q_1-q_2-\delta)(1-2x_1)^2}{4q_2(q_1-q_2)} > 0,$$

we must have $\bar{r}_2 \in (0,1)$. Let $\bar{r} = \min(\bar{r}_1, \bar{r}_2)$, then $\bar{r} \in (0,1)$.

Under the assumption in this proposition, we have $x_1 \geq k^1$ for all $r \in [0,1]$. And if $x_1 \in (\tilde{k}, \frac{1}{2})$, we have $k^1 \leq x_1 \leq k^2$ when $r \in [0, \bar{r}]$, which means that the optimal solution is in sub-region $\mathbf{R}_{11}$. Moreover, the optimal profit π^*_{11} is decreasing in r from the definition of $\bar{r}$. When $r \in (\bar{r}, 1]$, the optimal solution is in either sub-region $\mathbf{R}_{11}$ or $\mathbf{R}_{12}$. However, the optimal profit function π^*_{11} and π^*_{12} are increasing in $r \in (\bar{r}, 1]$ from Lemma 3.5.1 and the definition of $\bar{r}$.

If $x_1 \in \left[\frac{\delta}{q_1 - q_2}, \tilde{k}\right]$, we have either the optimal solution lies only in sub-region $\mathbf{R}_{12}$ or the optimal profit function π^*_{11} and π^*_{12} are increasing in $r \in [0, 1]$, thus firm's optimal profit is increasing in $r \in [0, 1]$. □

Lemma 3.5.2 *Suppose* (p_1^*, p_2^*) *is in region* $\mathbf{R}_{11}$ *as* (3.5)*, then* p_1^* *is non-monotone convex in* r*, and* p_2^* *is convex increasing in* r*.*

Proof. p_1^* is convex in r since

$$\frac{\partial^2 p_1^*}{\partial r^2} = \frac{q_2(q_2+\delta)\delta^2}{(q_2+(1-r)\delta)^3} + \frac{2(q_1-q_2)(q_1-q_2-\delta)(1-x_1)\delta^2}{(q_1-q_2-\delta+r\delta)^3} \geq 0.$$

p_2^* is convex and increasing in r since

$$\frac{\partial p_2^*}{\partial r} = \frac{\delta q_2(q_2+\delta)}{2(q_2+(1-r)\delta)^2} \geq 0,$$
$$\frac{\partial^2 p_2^*}{\partial r^2} = \frac{\delta^2 q_2(q_2+\delta)}{(q_2+(1-r)\delta)^3} \geq 0.$$

□