Optimal Production Control under Uncertainty with Revenue Management Considerations

by

Oben Ceryan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering and Industrial and Operations Engineering)
in The University of Michigan
2010

Doctoral Committee:

Professor Izak Duenyas, Co-Chair
Professor Yoram Koren, Co-Chair
Professor Panos Y. Papalambros
Associate Professor Göker Aydın
Assistant Professor Özge Şahin

© Oben Ceryan 2010
All Rights Reserved

To my parents with deepest gratitude
for their love and continuous support

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my thesis supervisors Professor Izak Duenyas and Professor Yoram Koren. Professor Koren has continually shared with me his visionary perspectives on a broad range of topics and research directions and has given me freedom to explore diverse subject areas. I thank him for his guidance, suggestions and the many opportunities he has provided me. I am especially grateful for his exceedingly generous support. I learned much of what I know in stochastic control and optimization from Professor Duenyas. He has introduced me to an exciting research area which ultimately help shape the topic of this dissertation. I owe a deep debt of gratitude for his mentoring, guidance, advice, insights and continuous support. I would like to thank Professor Panos Y. Papalambros for his valuable and intriguing discussions and for serving on my committee. His book gave me the first and substantial insights into optimization theory. I am greatly thankful to Professor Özge Şahin for her advice, support, and ideas. It has always been extremely delightful to collaborate with her. I also would like to thank Professor Göker Aydın for his advice, support, suggestions, and for serving on my committee. I feel privileged to having known and be working with each of my committee members. Finally, I would like to acknowledge the National Science Foundation Engineering Research Center for Reconfigurable Manufacturing Systems, the Department of Mechanical Engineering, the Department of Industrial and Operations Engineering, and the Ross School of Business for providing funding for this research.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

CHAPTER I

Introduction

1.1 Motivation

The unifying theme of this dissertation is supply chain and revenue management. This area of research aims to enhance firms' profitability by aligning supply with demand through integration of marketing decisions (e.g., pricing) that influence the demand process and strategic and operational decisions (e.g., capacity installations and production planning) that govern the supply process.

Within this broad area, this dissertation focuses on stochastic optimal control problems related to joint pricing, demand management, and production control decisions for multiple products under capacity limitations and demand uncertainties.

Manufacturing and service firms across various industries face uncertainties in their demand and supply processes. These uncertainties may result in demand losses and excess inventories, lowering profitability and competitiveness in the long run. Traditionally, firms countered variability in demand and supply by either building extra capacity or keeping reserve inventories. In addition, over the last decade, many industries have seen investments in reconfigurable and flexible manufacturing systems that enable the production of multiple variations of products in the same factory. This enables the product mix to be easily altered if demand for one product increases while demand for another decreases, hence providing a risk-pooling benefit. Devising optimal capacity investment, production control and inventory management strategies under uncertainty has been among the foremost interests of supply chain and operations management research.

More recently however, revenue management has emerged as a powerful tool to endogenize the demand process. Strategies such as dynamic pricing in which prices respond to demand and availability of products, or customer segmentation and prioritization in which different customer classes may be offered different service availability levels have been widely used in service industries such as airline and hotel management. With the advent of e-commerce and the ability to frequently change and advertise prices, these strategies have also increasingly been adapted by manufacturing enterprises in industries such as electronics and automotive.

Consequently, gaining insight into optimal production control decisions within a multiple product setting where firms also influence demand constitutes an interesting research question and is the main motivation for this dissertation. The problems addressed in this thesis were motivated by actual business concerns and apply to a wide array of industry practices. Through a rigorous and theoretical analysis of each research question set forth in the subsequent chapters, the focal point has been providing managers significant insights and implementable policies throughout their dynamic decision making processes regarding production, pricing and demand prioritization.

1.2 Research Objectives and Methodologies

Manufacturing firms often produce multiple variations of products that are substitutable from a customer's perspective. For example, an automotive manufacturer may produce several types of vehicles of the same model with varying engine displacements. In such a setting, the relative price of each product is a key factor that determines the consumer demand for a specific product.

A manufacturer that employs flexible resources to produce multiple products and that implements a dynamic pricing strategy thus has the following choices to respond to a change in demand. It may either individually increase or decrease the prices of items to stimulate, restrict, or shift demand from one item to another, it may assign more of the flexible capacity to a product that faces shortages, or it may use a combination of the two policies. To prevent impairing consumers' perception of product valuations in the long run, an important consideration for the manufacturer is to maintain a reasonable price gap among the different models.

How a firm under this setting should manage its joint pricing and production policies using flexibility, how the availability of a flexible resource influences the firm's pricing strategy, and the circumstances under which dynamic pricing contributes to profitability more than capacity flexibility (and vice versa) are among the main research questions addressed in Chapter 2.

Next, we consider a business setting that consists of multiple selling channels for a product for different purposes and at different prices. For example, in addition to assembling end products, a firm may also sell some intermediate products separately in order to sustain an after-sales service operation or to supply another firm through a component sharing agreement. If a firm operating within this setting has sufficient inventories of a certain product, it may choose to sell the item through a low revenue and/or low priority channel. Besides bringing in revenues, this sale will also reduce inventory levels and generate additional cost savings. However, when the inventories of a specific item are low, the firm faces a tradeoff between whether to sell the item through the secondary channel or reserve it for assembly purposes that could bring in a higher revenue. Resolving this tradeoff is a difficult task when the final product requires coordination of availabilities of other items and when both the demand and production/assembly processes exhibit uncertainties.

In such a setting, determining efficient production control and demand prioritization decisions may contribute significantly to profitability. Among the decisions the firm faces at any point in time are whether to accept or reject individual demands for intermediate products, the production quantities for each product and, determining whether an intermediate product is more valuable individually or as part of an assembled end product. These questions are the main motivating factors for the problem studied in Chapter 3.

The final consideration of this dissertation corresponds to the make-to-order and mass customization paradigm. Business models such as make-to-stock, which may usually be preferred if the number of products offered is limited, lead to very significant inventory costs for a high variety of end products especially under both production and demand uncertainties. On the other hand, a make-to-order system keeps inventory only at the component level and products are assembled after a customer order is received.

As many firms increasingly implement a make-to-order strategy, the challenges faced by firms in this setting to effectively coordinate the production of components, allocate assembly line capacity shared across many different products, and set demand admission decisions for products that bring in diverse revenues constitute the research questions investigated in Chapter 4.

Besides the common theme of jointly determining production control and demand management strategies under a variety of problem settings, the analysis in each subsequent chapter of this dissertation also share several elements of the following research objectives and outcomes. These are: (a) the modeling and formulation of a multiple period optimization problem, (b) characterization of optimal policy structures, (c) investigating the sensitivity of the optimal policy to various problem parameters, (d) providing managerial insights, (e) performing numerical studies, and (f) development of heuristic solution approaches and algorithms to facilitate implementations in large scale and practical problems.

Characterization of the optimal policy structure for each of the problems studied in this thesis is especially pivotal. The structural properties enable us to gain managerial insights on the nature of the optimal actions. They also facilitate sensitivity analysis, furthering our understanding of how various problem parameters influence the optimal decisions. Moreover, knowing the structure of the optimal policies allows us to perform efficient computations to determine the optimal decisions for a particular problem instance. Finally, the structural properties also enable the construction of algorithms that search among only specific types of decision rules.

Due to the highly interdisciplinary nature of this research, the theoretical analysis within this dissertation draws from many tools and methodologies from disciplines such as engineering, operations research, applied mathematics, statistics, and economics. Specifically, the formulation and analysis of the problems set forth in the following chapters apply methodologies related to convex optimization, optimal control theory, stochastic dynamic programming, Markov decision processes, and queueing theory.

1.3 Organization of the Dissertation

The dissertation is presented in a multiple manuscript format. The results in Chapters 2, 3, and 4 have appeared as individual research papers [10, 11, 13]. The organization of the dissertation is as follows.

Chapter 2 considers a firm that utilizes both dynamic pricing and capacity flexibility to manage the demand and supply for multiple products. Specifically, the setting consists of a firm that employs a capacity portfolio of product-dedicated and flexible resources and produces two substitutable products for which it sets the prices dynamically. The structure of the optimal production and pricing policies are characterized. In addition, the sensitivity of the optimal policy to various problem parameters (e.g., production costs, capacity levels and the demand model) is investigated. Further, several numerical studies are presented to visualize the benefits of the joint strategy as well as the circumstances under which each strategy is most beneficial.

Chapter 3 studies a manufacturing firm that has a two-stage operation where several intermediate products are produced in the first stage which are then assembled into an end-product through a second stage assembly operation. The manufacturer experiences demands for both the end-product and any of the intermediate items. We provide structural results regarding the optimal demand admission, production and assembly decisions. In addition, we investigate the sensitivity of the optimal policy to product prices. Further, the model is also extended to take into account multiple customer classes based on their willingness to pay and to a more general revenue collecting scheme where only an upfront partial payment for an item is received if a customer demand is accepted for future delivery with the remaining revenue received upon delivery. Finally, an effective heuristic policy is proposed.

Chapter 4 also examines a two-stage make-to-order production system where products are assembled only after an order is received. In this study, we allow customers to choose among several versions of the same product to be assembled rather than a single end-product. The structure of the optimal policies regarding the firm's decisions on how many components of each type to produce and how to set demand admission and rejection rules to prioritize orders for various products that compete for a shared capacity is discussed. In addition, a heuristic algorithm is devised that is robust with respect to the number of product alternatives offered.

Finally, Chapter 5 concludes the thesis by summarizing major contributions and presenting future research directions.

CHAPTER II

Managing Demand and Supply for Multiple Products through Dynamic Pricing and Capacity Flexibility

2.1 Overview

Firms that offer multiple products are often susceptible to periods of inventory mismatches where one product may face shortages while the other has excess inventories. This chapter studies a joint mechanism of dynamic pricing and capacity flexibility to alleviate the level of such inventory disparities. The setting consists of a firm producing two products with correlated demands utilizing capacitated product dedicated and flexible resources. The first objective is to characterize the structure of the optimal production and pricing decisions followed by an exploration on how changes in various problem parameters affect this optimal policy structure.

The results in this chapter show that the availability of a flexible resource helps maintain stable price differences across items over time even though the price of each item may fluctuate over time. This result has favorable ramifications from a marketing standpoint as it suggests that even when a firm applies a dynamic pricing strategy, it may still establish consistent price positioning among multiple products if it can employ a flexible replenishment resource.

In addition, the economic benefits of a joint strategy is compared to applying each tool individually. The results indicate that dynamic pricing and capacity flexibility can be viewed as substitute, but not fully interchangeable approaches and that the former is a more powerful tool if demands are positively correlated while the latter provides much of the benefits when demands are negatively correlated.

2.2 Introduction

Virtually all manufacturing and service industries are susceptible to periods of supply and demand mismatches. Due to capacity limitations and demand uncertainties, firms producing multiple products may frequently encounter instances where one of their products faces shortages while the other has excess inventories. In order to alleviate the level of such inventory mismatches, firms may utilize several tools to either alter supply or demand. Our focus in this paper will be a joint analysis of two of these mechanisms, namely, dynamic pricing and capacity flexibility.

In the last decade, firms in many industries have invested in flexible manufacturing systems that enable the production of multiple variations of products in the same factory. This enables the firm to easily alter its product mix if demand for one product increases while demand for another decreases. However, firms can also dynamically decrease or increase prices in response to demand fluctuations. For example, many LCD manufacturers make multiple sizes of LCDs in the same factory. The facilities are flexible so the firm can alter its mix fairly easily and demand is subject to tremendous variability.

During the “great recession” of 2009, demand for larger sized (42 inches and above) LCD TVs have slowed down in the U.S. as consumers trimmed their budgets and preferred smaller sized and lower priced models, according to the market research firm DisplaySearch. Thus an LCD TV manufacturer that produces multiple models of different sizes has the following choices to respond to this change in demand: 1)

It can decrease the price of larger sized models to stimulate more demand, 2) it can switch more of their production to smaller sized models (e.g., 32, 37 and 40 inch) or a combination of the two policies. Further, DisplaySearch estimates that the increased demand for smaller sized TVs as a result of the economic downturn is temporary and as the world emerges from the recession, demand for larger sizes will again outpace the smaller size TVs. Therefore, an important consideration is that the LCD TV manufacturers would like to maintain a reasonable price difference between the different size models (e.g., it may not be a good strategy to drastically reduce the price of 46 inch TVs below those of 37 inch TVs to respond to short term demand fluctuations and inventory excess as this will influence customers' perceptions of product valuations in the long run). This motivates the problem addressed in this chapter: How should a firm manage its simultaneous production and pricing policies for multiple products using flexibility?

Dynamic pricing in which prices respond to demand and availability of products has long been used in airline management. More recently, with the advent of e-commerce and the ability to frequently change and advertise prices, dynamic pricing has also been increasingly used in many other industries such as electronics and automobiles. As discussed by Biller et al in [6], several companies in various industries, notably Dell Computer, implement a Direct-to-Customer model in which dynamic pricing is used based on inventory levels and competition. As another example from the automotive industry, Copeland et al [18] provide empirical observations on whether vehicle prices are correlated with inventory fluctuations and they conclude that a significant negative relationship exists between inventories and prices. Through price discounts or price surcharges that may stimulate or reduce the overall demand or shift demand from one item to another, dynamic pricing may enable reductions in both high inventories and long customer backlogs. As a result, dynamic pricing may help firms to achieve higher profits. According to a recent study, if managed well, dynamic pricing can improve revenues and profits by up to 8% and 25%, respectively [49].

On the supply side, flexible manufacturing systems may also be utilized to align supply with demand. By shifting additional resources to a product with deficient inventory, flexible resources enable reductions in costs associated with production delays and customer backlogs. Goyal et al [28] analyze empirically how flexibility is utilized in the automotive industry where they consider flexibility as the ability of the general assembly line to manufacture different car platforms. Their data indicate that the share of flexible capacity is increasing over time and constituted approximately 40% and 30% of the overall capacity portfolio for GM and Ford, respectively in 2004. They also find that flexibility deployment is positively associated with demand uncertainty and negatively associated with demand correlation among different models.

Several interesting questions arise when dynamic pricing and capacity flexibility are considered simultaneously. First, we are interested in answering (*i*) how should the firm decide on the price charged for each item, (*ii*) how much of each product should the firm produce and (*iii*) how should the flexible resource be allocated among products in a given period. Hence, the first goal in this study is to characterize the optimal dynamic pricing and replenishment policy for multiple products over multiple periods in the presence of capacity limitations and the availability of a flexible resource. Second, we are interested in understanding the influence of the availability of a flexible resource on the firm's pricing decision. That is, we would like to compare the optimal pricing policy of a firm which may utilize flexible resources to that of a firm which employs only product dedicated resources. Third, we aim to identify the economic benefits obtained by applying each tool jointly and separately and understand (i) whether dynamic pricing and capacity flexibility are substitute approaches, i.e. if the economic benefits obtained by one tool diminishes with the utilization of the other, (ii) whether applying one tool dominates the other, and (iii) the circumstances under which dynamic pricing may contribute to profitability more than capacity flexibility, and vice versa.

The first contribution of this chapter is therefore providing a full characterization of joint optimal production and pricing decisions for two substitutable products with limited production capacities in the form of product dedicated and flexible resources. Assuming a linear additive stochastic demand model that is commonly used in the literature, this study shows that the optimal production policy can be characterized by modified base-stock levels that exhibit distinct forms across two broad regions of the state-space. To assist in the representation of the optimal policy, the initial inventory level of a product is classified as overstocked if the item requires no further replenishment, as moderately understocked if the available capacity is adequate to bring the inventory to a desired level, and as critically understocked if capacity is restrictive to reach the desired inventory level. The analysis shows that when at most one item is critically understocked, the modified base-stock level for each product is described by a decreasing function of the inventory level of the other item. However, when both items are critically understocked, it is shown that the modified base-stock level for a product is characterized by an increasing function of the inventory position of both products.

Regarding the optimal pricing policy, the results indicate that a list price is charged for an item if it is moderately understocked. If an item is critically understocked, then a price markup that depends on both inventory levels is applied. When an item is overstocked, a price discount that depends on both inventory levels is given. Furthermore, the analysis reveals that when inventory levels for both items are critically understocked and when the flexible capacity is simultaneously shared between products, the existence of the flexible resource leads to an optimal pricing scheme that maintains a constant price difference between products. At such instances, dynamic pricing only adjusts the overall level of demand for both products but does not attempt to shift demand from one product to another while mismatches between the desired and actual inventory level of products is restored solely by the availability of flexible capacity.

Hence, the second major finding in this chapter is that the availability of a flexible resource helps maintain stable price differences across items over time even though the price of each item may fluctuate over time. This result has favorable ramifications from a marketing standpoint as it suggests that even when a firm applies a dynamic pricing strategy, it may still establish consistent price positioning among multiple products if it can employ a flexible replenishment resource.

On the economic benefits of implementing dynamic pricing and capacity flexibility individually or simultaneously, this study shows that the two mechanisms may be viewed as substitute, but not fully interchangeable approaches. Through numerical examples, it is demonstrated that dynamic pricing is a more effective tool when both items are either under- or over-stocked. Such instances may be observed frequently when demand uncertainties for the products are positively correlated. On the other hand, the results indicate flexible capacity to be the more effective tool when there is a negative correlation between the demand uncertainties which yields to instances with inventory mismatches where one item is well stocked and the other having limited inventories.

The remainder of this chapter is organized as follows. In Section 2.3, the related literature is reviewed. The model framework and the problem formulation is provided in Section 2.4. In Section 2.5, the structure of the optimal pricing and production policies is characterized while in Section 2.6 the sensitivity properties of the optimal policy with respect to various demand, cost, and capacity parameters are analytically investigated. Section 2.7, numerical studies are performed to evaluate the benefits of flexibility and compare the performances of joint strategies to applying each tool individually. Section 2.8 summarizes the conclusions and main results. Finally, 2.9 provides the proofs of all results.

2.3 Literature Review

There exists a vast literature on dynamic pricing. Due to the positioning of the research question addressed in this chapter, only those studying joint pricing and replenishment decisions are referenced. Extensive reviews on the interplay of pricing and production decisions have been provided by Elmaghraby and Keskinocak [21], Bitran and Caldentey [8], and Chan et. al. [14].

Single product settings have been the focus of much of the earlier work in this area. Whitin [56] is among the first to consider joint pricing and inventory control for single period problems under both deterministic and stochastic demand models. For a finite horizon, periodic review model, Federgruen and Heching [23] show that the optimal policy is of a base-stock, list-price type. When it is optimal to order, the inventory is brought to a base-stock level and a list-price is charged. For inventory levels where no ordering takes place, the optimal policy assigns a discounted price. In a subsequent work, Li and Zheng [39] extend the setting studied by Federgruen and Heching to include yield uncertainty for replenishments. Chen and Simchi-Levi [15] further extend the results of Federgruen and Heching to include fixed ordering costs and show that a stationary (s,S,p) policy is optimal for both the discounted and average profit models with general demand functions. In such a policy, the period inventory is managed based on the classical (s,S) type policy, and price is determined based on the inventory position at the beginning of each period.

Recently, settings consisting of multiple substitutable products have received more attention. Aydin and Porteus [3] study a single period inventory and pricing problem for an assortment consisting of multiple products. They investigate various demand models and show that a price vector accompanied by corresponding inventory stocking levels constitute the unique solution to the profit maximization problem although the profit function may not necessarily be quasi-concave in product prices. Song and Xue [52] extend the setting studied by Aydin and Porteus to multiple periods and characterize the optimal policy structure and develop algorithms.

Zhu and Thonemann [58] study a periodic review, infinite capacity, joint production and pricing problem with two substitutable products assuming a linear additive demand model. They show that production for each item follows a base stock policy which is nonincreasing in the inventory position of the other item. They also show that the optimal pricing decisions do not necessarily exhibit monotonicites with respect to inventory positions except for settings where the demand process for both products are influenced by identical cross-price elasticities. They find that a list price is optimal whenever an order is placed for a product, regardless of the inventory position of the other product and a discount is given for any product that is not ordered. Ye [57] extends their results to an assortment of more than two products and shows that under a similar linear additive demand model and identical cross-price elasticities, a base-stock, list-price policy extends to an arbitrary number of products. Both of these papers assume infinite production capacity. If production capacity is limited, charging list prices for an item whenever an order is placed for that item is no longer optimal. Intuitively, one would expect to charge a higher price when the desired production quantity is restricted by a limited capacity. In this chapter, it is show that this expectation is indeed true. Consequently, as opposed to the results for the infinite capacity setting, whenever an order is placed for a product, its price is no longer independent of the inventory position of the other item.

On the flexible capacity side, a major research area has been determining the optimal portfolio of flexible and dedicated capacities under demand uncertainty. We refer the reader to the pioneering works by Fine and Freund [24] and Van Mieghem [42] for the analysis of optimal capacity investments as well as the more recent works [44, 12, 35] and the references therein for extensions to discrete capacity choices. Rather than the optimal investment problem, the setting studied in this chapter considers the optimal allocation problem. In one of the earliest works, Evans [22] studies a periodic review problem with two products produced by a single shared resource and characterizes the optimal allocation policy for the flexible resource. DeCroix and Arreola-Risa [19] study extensions to multiple products. For an infinite horizon problem with homogenous products where all products have identical cost parameters and resource requirements, they derive structural results regarding the optimal allocation of the flexible capacity. Besides these periodic review models, continuous time formulations and corresponding results may also be found in works such as the ones by Glasserman [27] and Ha [29]. However, these papers on flexible capacity allocation treat the demand process as exogenous whereas our focus is to also consider dynamic pricing that influences the demand for each item.

There has also been prior interest in combining these two streams of research. Chod and Rudi [16] study the effects of resource flexibility and price-setting in a single period model. In their model, the firm first decides on the capacity investments prior to demand realizations. After product demands are realized, capacity allocations and product pricing decisions are given. Hence the major differences in the setting discusses in this chapter are that 1) here we consider a multiple period model requiring price selections and production decisions every period whereas they consider a single period model and 2) they assume that allocation decisions can be made after demand is realized which implicitly means zero lead times, whereas in this study, the assumption is that the allocation decisions are made prior to demand realization.

2.4 Problem Formulation

Consider a firm that produces two products where prices and replenishment quantities for both items are dynamically set at the beginning of each period over a finite planning horizon of length T. Let x_i^t, y_i^t, and d_i^t denote the initial inventory position at the beginning of period t, the produce-up-to-level in period t, and the demand in period t for product i, $i = \{1, 2\}$, respectively. The sequence of events is given in Figure 2.1. At the beginning of period t, the manufacturer reviews the current inventory positions $(x_1^t, x_2^t) \in \Re^2$ and decides on (i) the optimal order up to levels (y_1^t, y_2^t) and (ii) the prices, (p_1^t, p_2^t) to charge during the period.

The demands for both items are assumed to be correlated by the following linear additive demand model which has been prevalent in related literature.

$$\begin{aligned} d_1^t(p_1^t, p_2^t, \epsilon_1^t) &= b_1^t - a_{11}^t p_1^t - a_{12}^t p_2^t + \epsilon_1^t \\ d_2^t(p_1^t, p_2^t, \epsilon_2^t) &= b_2^t - a_{21}^t p_1^t - a_{22}^t p_2^t + \epsilon_2^t \end{aligned} \tag{2.1}$$

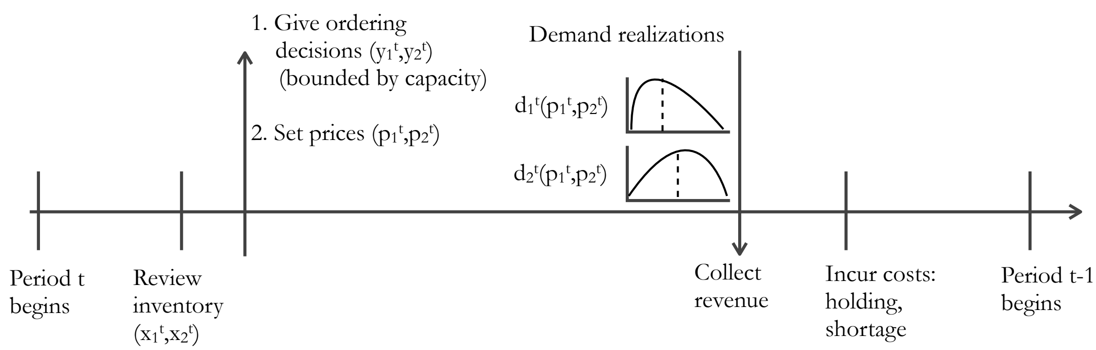


Figure 2.1: Sequence of events

In (2.1), b_i denotes the demand intercept whereas a_{ii}^t and a_{ij}^t for $i,j = \{1,2\}$ and $i \neq j$ refer to the individual and cross-price elasticities for product type-i. We let ϵ_1^t and ϵ_2^t refer to independent random variables having continuous probability distributions with mean zero and nonnegative support on the product demands. For future reference, the mean demand for product type-i is denoted by $\bar{d}_i^t(p_1^t, p_2^t)$ where $\bar{d}_1^t(p_1^t, p_2^t) = b_1^t - a_{11}^t p_1^t - a_{12}^t p_2^t$ and $\bar{d}_2^t(p_1^t, p_2^t) = b_2^t - a_{21}^t p_1^t - a_{22}^t p_2^t$.

We assume that the square matrix $\mathbf{A}^t$ with elements a_{ij}^t for $i,j = \{1,2\}$ has positive diagonal elements and negative off-diagonal elements, that is $a_{ii}^t > 0$ and $a_{ij}^t < 0$ for $i \neq j$. This assumption reflects the substitutable nature of the products and that the demand for an item is decreasing in its own price and increasing with the price of the other item. It is also assumed that $\mathbf{A}^t$ possesses diagonal dominance property, i.e., $a_{11}^t \geq |a_{12}^t|$ and $a_{22}^t \geq |a_{21}^t|$. This implies that the income effect is at least as strong as the substitution effect, i.e., a price change on an item influences its demand at least as strongly as it influences the demand for the other item. These assumptions on demand parameters, besides their economic justification, also result in a concave revenue function. Further, we impose another assumption on $\mathbf{A}^t$, that $\mathbf{A}^t$ is symmetric. A symmetric $\mathbf{A}^t$ is equivalent to settings where the demands for both items may be influenced by different individual price elasticities but they experience identical cross-price elasticities. In other words, the derivative of the expected demand for an item with respect to the price of the other item is equivalent for both products. Albeit restrictive in modeling more diverse demand structures, this assumption has been incorporated in a number of related works and is also essential in our derivations to fully characterize the structure of the optimal policy. Furthermore, the same property is also inherently present in Multinomial Logit (MNL) type demand models that is described in Section 2.7. Finally, no restrictions are imposed on the price decisions with $\mathbf{p^t} \in \Re^2$ as non-negativity of optimal prices may be guaranteed within a set of demand parameters reflecting a practical setting.

Production decisions are made at the beginning of period t, and prices are set before the demand is realized. The firm utilizes fixed dedicated capacities $K_1, K_2 \geq 0$ for the production of each item exclusively, as well as a limited flexible resource, $K_0 \geq 0$, that may be assigned partially or entirely for the production of both items. A unit of flexible resource may be used towards producing a unit of either product. At each period, the optimal production quantities are bounded by the corresponding available flexible and product-dedicated capacity levels. We let $\mathcal{D}(\mathbf{x^t})$ denote the set of admissable values for $\mathbf{y^t}$, i.e., $\mathbf{y^t} \in \mathcal{D}(\mathbf{x^t})$ where $\mathcal{D}(\mathbf{x^t}) := \{\mathbf{y^t} | x_i^t \leq y_i^t \leq x_i^t + K_0 + K_i \ \forall i = 1,2 \text{ and } y_1^t + y_2^t \leq x_1^t + x_2^t + K_0 + K_1 + K_2\}$. We let c_i^t denote the unit production cost for product type-i in period t and assume that this unit cost is applicable to both dedicated and flexible production systems when producing the same item. Consequently, this allows incurring separate production costs corresponding to each item at instances when both items are produced on the same flexible resource. This assumption is especially applicable when the production cost for an item constitutes mostly of the raw materials or when the processing costs differ across products yet remain constant across types of resources. All unsatisfied demands are allowed to be backordered. At the end of period t, the firm incurs holding and backorder costs of h_i^t and π_i^t per unit of product type-i that is kept in inventory or backordered, respectively. To simplify the notation throughout the subsequent analysis, we suppress the superscript t on demand and cost parameters $a_{ij}^t, b_i^t, c_i^t, h_i^t$, and π_i^t. The results in this chapter do not assume that these parameters are stationary over the planning horizon.

Letting $V^t(\mathbf{x^t})$ denote the expected discounted profit-to-go function under the optimal policy starting at state $\mathbf{x^t}$ with t periods remaining until the end of the horizon, the problem can be expressed as a stochastic dynamic program satisfying the following recursive relation:

$$V^t(\mathbf{x^t}) = \max_{\mathbf{y^t} \in \mathcal{D}(\mathbf{x^t}), \mathbf{p^t}} G^t(\mathbf{y^t}, \mathbf{p^t})$$

where

$$\begin{aligned} G^t(\mathbf{y^t}, \mathbf{p^t}) &= R(\mathbf{p^t}) - \mathbf{c}(\mathbf{y^t} - \mathbf{x^t}) + \mathrm{E}_{\boldsymbol{\epsilon}^t}\big\{ - \mathbf{h}(\mathbf{y^t} - \mathbf{\bar{d}^t} - \boldsymbol{\epsilon}^t)^+ - \boldsymbol{\pi}(\mathbf{\bar{d}^t} + \boldsymbol{\epsilon}^t - \mathbf{y^t})^+ \\ &\qquad + \beta V^{t-1}(\mathbf{y^t} - \mathbf{\bar{d}^t} - \boldsymbol{\epsilon}^t)\big\}, \\ R(\mathbf{p^t}) &= \mathbf{p^t}(\mathbf{b} - \mathbf{A}\mathbf{p^t}), \end{aligned} \tag{2.2}$$

and β is the discount factor. $V^0(\mathbf{x})$ denotes the terminal value function and is set at $V^0(\mathbf{x}) = 0$. In order to facilitate the analysis, a change of variables is performed by defining $\mathbf{z^t}$ such that $\mathbf{z^t} = \mathbf{y^t} - \mathbf{\bar{d}^t}$, i.e., $\mathbf{z^t} = \mathbf{y^t} - \mathbf{b} + \mathbf{A}\mathbf{p^t}$. Therefore, if we let $\mathcal{D}'(\mathbf{x^t}, \mathbf{p^t})$ denote the set of admissable decisions for $\mathbf{z^t}$, we can write $\mathcal{D}'(\mathbf{x^t}, \mathbf{p^t}) = \{\mathbf{z^t} | x_i^t \le z_i^t + b_i - a_{i1}p_1^t - a_{i2}p_2^t \le x_i^t + K_0 + K_i \ \forall i = 1,2 \text{ and } z_1^t + z_2^t + b_1 + b_2 - (a_{11} + a_{21})p_1^t - (a_{12} + a_{22})p_2^t \le x_1^t + x_2^t + K_0 + K_1 + K_2\}$. Then, the dynamic programming formulation given in (2.2) may be written as:

$$V^t(\mathbf{x^t}) = \max_{\mathbf{z^t}\in\mathcal{D}'(\mathbf{x^t},\mathbf{p^t}),\mathbf{p^t}} J^t(\mathbf{z^t},\mathbf{p^t})$$

where

$$\begin{aligned} J^t(\mathbf{z^t},p^t) &= R'(\mathbf{p^t}) + \mathbf{cx^t} - \mathbf{cz^t} + \mathrm{E}_{\boldsymbol{\epsilon}^t}\left\{-\mathbf{h}(\mathbf{z^t}-\boldsymbol{\epsilon}^t)^+ - \boldsymbol{\pi}(\boldsymbol{\epsilon}^t-\mathbf{z^t})^+ + \beta V^{t-1}(\mathbf{z^t}-\boldsymbol{\epsilon}^t)\right\}, \\ R'(\mathbf{p^t}) &= (\mathbf{p^t}-\mathbf{c})(\mathbf{b}-\mathbf{Ap^t}) \end{aligned} \tag{2.3}$$

In this reconstructed formulation, the new decision variables are $\mathbf{z^t}$ and $\mathbf{p^t}$, where $\mathbf{z^t}$ corresponds to a target inventory level reached after the current inventory position is augmented by the replenishment quantity and depleted by the selected mean demand. The profit-to-go function, $V^{t-1}(\mathbf{z^t}-\boldsymbol{\epsilon}^t)$, only depends on the set of variables $\mathbf{z^t}$ and proves useful in deriving several structural results on the value function that we require in the analysis of the optimal policy.

The next section explores the effects of the presence of a flexible resource and the limitations in production capacity on the optimal pricing and production policy structure.

2.5 Characterization of the Optimal Policy Structure

In this section, we first establish several structural properties on the value function and prove that these properties are preserved under the dynamic programming recursions. Under the assumptions outlined in the preceding section, Lemma 2.1 shows that the single period objective function and the optimal value function are strictly concave throughout the planning horizon.

Lemma 2.1. *$J^t(\mathbf{z^t},\mathbf{p^t})$ and $V^t(\mathbf{x^t})$ are jointly strictly concave for all $t = 1, 2, \cdots, T$.*

Proof: The proof of Lemma 2.1 is provided in Section 2.9.1.

Strict concavity of the objective function $J^t(\mathbf{z^t}, \mathbf{p^t})$ implies the uniqueness of an optimal solution and thus strict complementary slackness holds almost everywhere except for a set of points with measure zero on $\Re^2$. Following Fiacco (1976), Lagrange multipliers are differentiable in decision variables, hence $J^t(\mathbf{z^t}, \mathbf{p^t})$ and $V^t(\mathbf{x^t})$ are twice continuously differentiable almost everywhere. The analysis is based on the first-order optimality conditions (provided in Section 2.9.1) which are necessary and sufficient due to the concavity of the problem.

While joint concavity established in Lemma 2.1 implies that the production policy will be of base-stock type and that there is a price pair that maximizes the profits, determining the complete structure of the optimal production and pricing policies requires additional properties on $J^t(\mathbf{z}^t, \mathbf{p}^t)$ which are summarized in Lemma 2.2.

Lemma 2.2. *For all* $t = 1, 2, \cdots T$*,*

(a) $J^t(\mathbf{z^t}, \mathbf{p^t})$ *is submodular in* $(\mathbf{z^t})$*,*

(b) $J^t(\mathbf{z^t}, \mathbf{p^t})$ *possesses the following diagonal-dominance property:*

$\frac{\partial^2 J^t}{\partial z_i^t \partial z_i^t} \leq \frac{\partial^2 J^t}{\partial z_i^t \partial z_j^t} \ \forall \ i, j; \quad i \neq j$

Proof: The proof of Lemma 2.2 is provided in Section 2.9.1.

We next characterize the optimal production and pricing policies which exhibit distinct forms across several regions of the state space.

2.5.1 Optimal Production Policy

In order to establish the optimal policy, we segment the state space into two broad regions based on the initial inventory levels of the items. The first region corresponds to instances for which there remains some resource (either dedicated or flexible) that is not fully utilized, and is denoted as Region A. The second, denoted as Region B, corresponds to initial inventory levels for which all resources are fully utilized.

The boundaries of these two regions are described by two monotone functions $\gamma_1^t(x_2^t)$ and $\gamma_2^t(x_1^t)$ (as specified in Theorem 2.1) which also subdivide Region A into several subregions with respect to the inventory position of each product and capacity limitations according to the following definition.

Definition 2.1. *Consider initial inventory levels* (x_1^t, x_2^t) *and the functions* $\gamma_1^t(x_2^t)$ *and* $\gamma_2^t(x_1^t)$ *and let* $(\bar{x}_1^t, \bar{x}_2^t) := \{(x_1^t, x_2^t),$ *s.t.* $x_1^t = \gamma_1^t(x_2^t)$ *and* $x_2^t = \gamma_2^t(x_1^t)\}$*. Further, let* $\hat{\gamma}_1^t(x_2^t)$ *(and* $\hat{\gamma}_2^t(x_1^t)$ *in a similar fashion) be given by*

$$\hat{\gamma}_1^t(x_2^t) := \begin{cases} \gamma_1^t(x_2^t) - K_1 & \text{if} \quad x_2^t \le \bar{x}_2^t - K_0 - K_2 \\ \gamma_1^t(x_2^t) + \bar{x}_2^t - x_2^t - K_0 - K_1 - K_2 & \text{if} \quad \bar{x}_2^t - K_0 - K_2 < x_2^t \le \bar{x}_2^t - K_2 \\ \gamma_1^t(x_2^t) - K_0 - K_1 & \text{if} \quad \bar{x}_2^t - K_2 < x_2^t \end{cases}$$

Then, product 1 (and likewise, product 2) is classified as: (a) "overstocked" if $x_1^t > \gamma_1^t(x_2^t)$*, (b) "moderately understocked" if* $\gamma_1^t(x_2^t) \ge x_1^t > \hat{\gamma}_1^t(x_2^t)$*, and (c) "critically understocked" if* $\hat{\gamma}_1^t(x_2^t) \ge x_1^t$*.*

Defining an item as overstocked means the item requires no further replenishment. A moderately understocked product requires production for which the available capacity is adequate to reach the desired base stock level whereas a critically understocked product may not be brought to the desired base stock level due to capacity restrictions.

Region A collectively represents all states in which at most one product is critically understocked whereas Region B corresponds to initial inventory levels for which both items are critically understocked. The segmentation of the state space is illustrated in Figures 2.2 and 2.3 and formally derived by accompanying lemmas within the proof of Theorem 2.1 which describes the optimal production policy.

Theorem 2.1. (Production Policy)*: The optimal production policy is a state dependent modified base stock policy characterized by three monotone functions* $\gamma_1^t(x_1^t)$, $\gamma_2^t(x_1^t)$, *and* $\alpha^t(x_1^t)$ *such that*

(i) In states corresponding to initial inventory levels for which at most one product is critically understocked (i.e. in Region A),

- *(a) the optimal production policy for product* i $(i = 1, 2)$ *is to produce up to the modified base stock level* $\min\left(x_i + K_0 + K_i, \gamma_i^t(x_{3-i}^t)\right)$.
- *(b) the modified base stock level for product* i *is non-decreasing with* x_i^t *and non-increasing with* x_j^t, $j \neq i$.

(ii) In states corresponding to initial inventory levels for which both products are critically understocked (i.e. in Region B),

- *(a) the optimal production policy for product 1 and product 2 is to produce up to the modified base stock level* $x_1^t + K_1 + l^t(x_1^t, x_2^t)$ *and* $x_2^t + K_2 + K_0 - l^t(x_1^t, x_2^t)$, *respectively, where* $l^t(x_1^t, x_2^t)$ *denotes the amount of flexible capacity allocated to product 1.*
- *(b)* $l^t(x_1^t, x_2^t) = 0$ *if* $x_2^t \leq \alpha^t(x_1^t) - K_0$, $l^t(x_1^t, x_2^t) = K_0$ *if* $x_2^t \geq \alpha^t(x_1^t + K_0)$. *Otherwise,* $l^t(x_1^t, x_2^t)$ *satisfies* $l^t(x_1^t, x_2^t) + \alpha^t(x_1^t + l^t(x_1^t, x_2^t)) = x_2^t + K_0$ *and the modified base stock levels for either product is a function of the starting inventory levels through their sum.*
- *(c)* $l^t(x_1^t, x_2^t)$ *is decreasing with* x_1^t *and increasing with* x_2^t.
- *(d) The modified base stock levels for product* i *is nondecreasing with either product's inventory level.*

Proof: The proof of Theorem 2.1 is provided in Section 2.9.2.

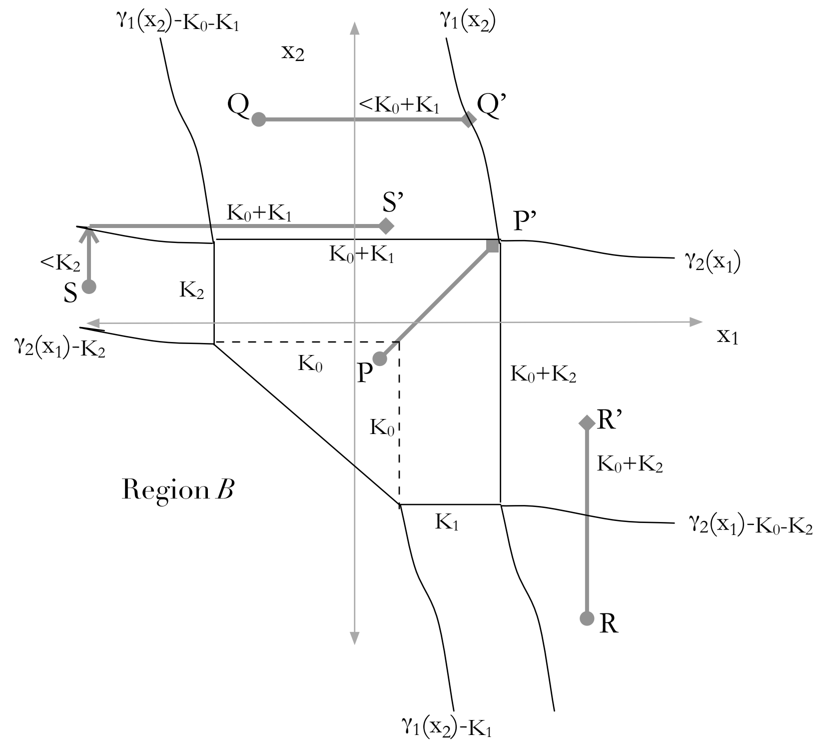


Figure 2.2: Optimal production policy in Region A

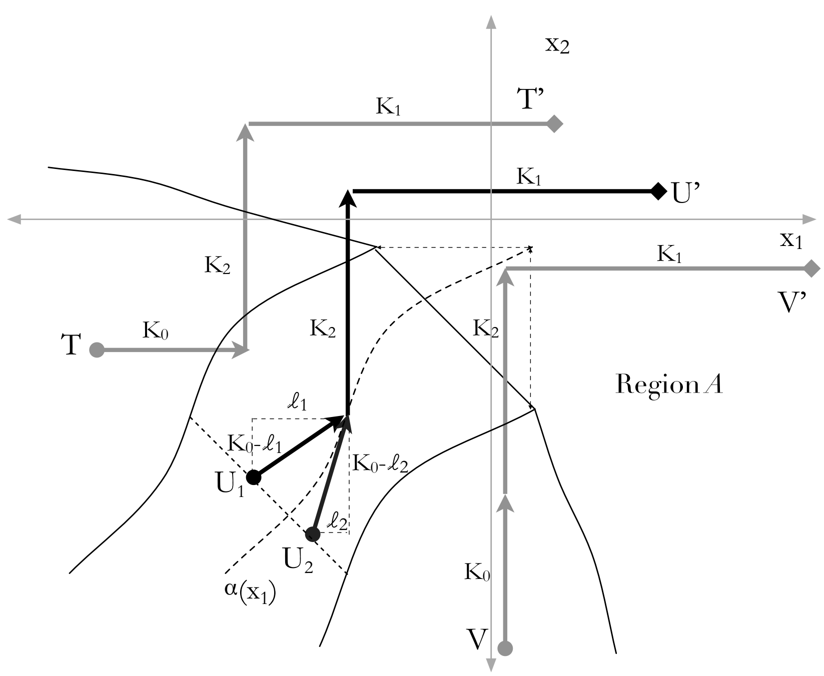


Figure 2.3: Optimal production policy in Region B

As Theorem 2.1 suggests, the optimal production policy has a number of properties depending on the inventory state at the beginning of a period. Figure 2.2 illustrates the optimal production policy in region A. When both products are moderately understocked, as shown by the starting inventory level P on Figure 2.2, it is optimal to produce both products up to the uniquely defined point $(\bar{x}_1^t, \bar{x}_2^t)$ which is depicted by point P' in Figure 2.2 and the optimal order-up-to levels in this region are independent of initial inventories.

Initial inventory levels Q and R in Figure 2.2 are examples of states where one item is overstocked and the other is understocked. Point Q illustrates an instance where item 2 is overstocked and item 1 is moderately understocked. Thus, starting at Q, with a base-stock level of $\gamma_1^t(x_2^t)$ for item 1 and no production for item 2, the optimal policy is to move to point Q'. We note that, point Q' refers to a base-stock level for item 1 which is lower than the one suggested by P'. The reason is twofold. First, as we will see in the discussion of the optimal pricing policy, the overstocked item 2 results in a price decrease for that item which in turn increases its demand and decreases the demand for item 1 which further decreases the base stock for item 1. Second, an overstocked item 2 reduces the potential work load on the flexible resource for that item and increases the availability of the flexible capacity for item 1 in future periods. This allows for fewer units of item 1 to be produced in the current period.

An initial inventory position such as point R on the other hand, shows an instance when the available capacity is not sufficient to bring the inventory position of a critically understocked item 2 to $\gamma_2^t(x_1^t)$. Hence, with no production for item 1 and using all available capacity to produce $K_0 + K_2$ units of item 2, the optimal policy is to move to point $(x_1^t, x_2^t + K_0 + K_2)$ which is depicted by point R'.

Point S refers to a state where item 1 is critically and item 2 is mildly understocked. In this case, Theorem 2.1 states that it is optimal to produce $K_0 + K_1$ units of item 1 and to bring the inventory of item 2 to the desired order-up-to level of $\gamma_2^t(x_1^t)$, as shown by point S'. Again, we note that point S' corresponds to a base stock level higher than the one implied by P' with similar but reverse dynamics as discussed previously. That is, a critically understocked item 1 not only results in a price increase for this item which increases the demand for item 2, but also potentially requires a higher share of the flexible capacity in future periods to be allocated to item 1 and reduces its availability for item 2. Consequently, the base stock level for item 2 is set higher in the current period. Finally, in the region where both items are sufficiently stocked, i.e., $x_i^t > \gamma_i^t(x_j^t)$, it is optimal not to produce either product.

Part 2 of Theorem 2.1 corresponds to the states in region B where both products are critically understocked and thus all production resources are fully utilized. As illustrated in Figure 2.3, Theorem 2.1 part 2 states that when the initial inventory levels for both products fall within a band defined by $\{(x_1^t, x_2^t)$, s.t. $(x_1^t, x_2^t) \in$ Region B, and $\alpha(x_1^t + K_0) > x_2^t > \alpha(x_1^t) - K_0\}$, the optimal policy allocates $l^t(x_1^t, x_2^t) > 0$ units of the flexible resource to product 1 and the remaining $K_0 - l^t(x_1^t, x_2^t) > 0$ units to product 2. Moreover, for any two inventory states corresponding to the same total inventory, the intermediate inventory positions after the flexible resource is utilized are identical. From this point on, additional units of each item is produced to the full extent of their dedicated resources. In Figure 2.3, points U_1 and U_2 refer to two states with equivalent total inventories and point U' corresponds to the inventory level reached by the optimal policy.

For initial inventory levels that fall outside this band, the flexible resource is fully assigned to the product which experiences the most severe shortage. For example, in

Figure 2.3, point T refers to an instance where all flexible capacity is used towards item 1 whereas point V shows an instance where the flexible capacity is used entirely for the production of item 2.

Part 2 (c) of Theorem 2.1 states that the share of flexible resource an item receives is decreasing with its own inventory and increasing with the other item's inventory. Referring to Figure 2.2, since the initial inventory level of product 1 corresponding to point U_1 is less than that of corresponding to U_2, the amount of flexible capacity allocated to item 1 when starting at U_1 is larger than the one starting at U_2. This reflects how flexible capacity is able to shift resources towards the product experiencing more severe shortages. As will be discussed next, the optimal prices charged for each product have a specific relationship within this band.

2.5.2 Optimal Pricing Policy

When making pricing decisions, it is often helpful to think in terms of markdowns and markups where a markdown (markup) corresponds to a price discount (surcharge) relative to a current period list price. Earlier results in the literature focused on infinite capacity settings, hence the optimal pricing policy was characterized by a list price, markdown policy. In such a policy, whenever an item is produced, a list price is charged regardless of the inventory position of the other item and a discount is given otherwise. In the presence of capacity limitations however, we find that, unlike the infinite capacity setting, charging list prices whenever production takes place for an item is no longer optimal. Consequently, the characterization of the optimal pricing policy relies on a third component, namely price markups.

In this section, we let $m_i^t(x_1^t, x_2^t)$ to denote the price markup/markdown for item i in period t with $m_i^t < 0$ corresponding to markdowns and $m_i^t > 0$ corresponding to markups in reference to a current period list price p_{iL}^t. Thus, in period t we have,

$$p_i^t(x_1^t, x_2^t) = p_{iL}^t + m_i^t(x_1^t, x_2^t) \tag{2.4}$$

The following theorem defines the optimal pricing policy.

Theorem 2.2. (Pricing Policy)*: For all $i = 1, 2$, in period t, we have the following:*

a) In Region A, if an item i is moderately understocked, then $m_i^t(x_1^t, x_2^t) = 0$ and it is optimal to charge a list price, p_{iL}^t, for that item where

$$p_{i,L}^t = \frac{a_{3-i,3-i}^t b_i^t - a_{12}^t b_{3-i}^t}{2(a_{11}^t a_{22}^t - {a_{12}^t}^2)} + \frac{c_i^t}{2}$$

If an item i is overstocked, then $m_i^t(x_1^t, x_2^t) < 0$, i.e., it is optimal to give a price discount to that item. If on the other hand, item i is critically understocked, then $m_i^t(x_1^t, x_2^t) > 0$, indicating that it is optimal to give a price markup to that item.

b) In Region B, $m_i^t(x_1^t, x_2^t) > 0$, hence the optimal policy marks up the price of both items. Furthermore, if (x_1^t, x_2^t) is such that $0 < l^t(x_1^t, x_2^t) < K_0$, then $m_1^t(x_1^t, x_2^t) = m_2^t(x_1^t, x_2^t)$ resulting in

$$p_2^t(x_1^t, x_2^t) = p_1^t(x_1^t, x_2^t) + C^t$$

where $C^t = p_{2L}^t - p_{1L}^t$.

c) The optimal price $p_i^t(x_1^t, x_2^t)$, $i = 1, 2$ is decreasing with respect to x_1^t and x_2^t.

Proof: The proof of Theorem 2.2 is provided in Section 2.9.2.

Figure 2.4 illustrates the optimal pricing policy in terms of markups (markdowns) for each product. It is optimal to give discounts on a product if it is overstocked, apply the list price on the product if it is moderately understocked and to markup the price of the item if it is critically understocked.

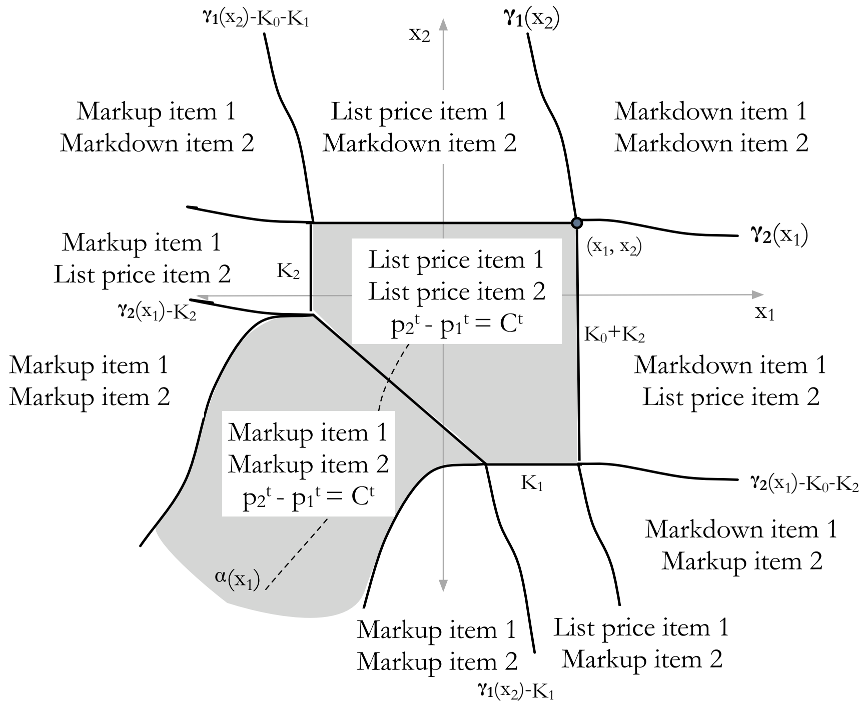


Figure 2.4: Optimal pricing policy with the shaded area indicating constant price difference between products

Part (b) of Theorem 2.2 suggests an interesting fact about the pricing policy when the inventory level falls within the band in region B where both products use a positive share of the flexible capacity. In states corresponding to this region, both products are marked up by exactly the same amount. This results in the price difference between items to remain identical to the difference between their list prices. (Although the exact price difference being maintained is due to the linear additive stochastic demand model with uniform cross price elasticities, in section 2.7, we show that even when a different demand model is used, the price difference between the products remains in a very narrow range.) This is due to the fact that the availability of flexible capacity enables us to direct capacity where it is most urgently needed and relieves the use of drastic changes in prices to shift demand.

This special structure of the optimal price policy has favorable consequences.

Capacity flexibility may be viewed as a significantly beneficial tool when firms use dynamic pricing and are sensitive to maintaining consistent price gaps among items in order to preserve product price positioning over different items.

2.6 Sensitivity of the Optimal Policy

Having characterized the optimal policy, we now discuss how changes in various problem parameters affect the optimal policy structure. Specifically, we explore the sensitivity of the optimal policy to (i) cost parameters including the production, holding and backorder costs, (ii) capacity parameters, and (iii) demand parameters including demand intercepts, individual and cross price elasticities. (In the following, the terms increasing and decreasing are used in the weak sense to denote nondecreasing and nonincreasing, respectively.)

2.6.1 Sensitivity to Cost Parameters

First, we are interested in how the production, holding or backorder costs for an item influence the optimal price and modified base-stock levels. Intuitively, as it becomes more costly to produce product 1, the selling price for this item would increase, reducing the demand and therefore its modified base-stock level. However, it is not obvious how the price of item 2 is affected. On the one hand, a resulting price increase for item 1 strengthens the demand for item 2 which may drive the prices for this item higher. On the other hand, the cost increase and the resulting price increase for item 1 decreases item 1's demand, potentially allowing more flexible capacity to be assigned to item 2, increasing item 2's availability and thereby decreasing its price.

Our first result in Theorem 2.3 shows that the former argument dominates unless both items are critically understocked and are receiving a share of the flexible capacity, for which the second argument prevails.

Theorem 2.3. *(a) If the current period production cost c_1 for product 1 increases, the optimal price charged for product 1 increases. The price for product 2 decreases if both items are critically understocked and the flexible resource is shared between the items, otherwise the price for product 2 increases. The modified base stock level for product 1 decreases while the modified base stock level for product 2 increases.*

(b) If the current period holding (backorder) cost h_1 (π_1) for product 1 increases, the price for item 1 decreases (increases) and the base stock level for product 1 decreases (increases).

Proof: The proof of Theorem 2.3 is provided in Section 2.9.3.

When the unit holding cost for item 1 increases, in order to reduce the number of unsold items, we would intuitively increase the demand for the product by decreasing its price. In addition, we would lower the modified base stock level for item 1. Part (b) of Theorem 2.3 verifies that this intuition is indeed correct. Similar but reverse reasoning applies for an increase in the backorder cost. That is, when the unit backorder cost for item 1 increases, the price of item 1 increases to lower its demand, and the modified base-stock level for the item increases. Both changes reduce the possibility of facing backorders. An increase in the holding or backorder cost for item 1 however, does not necessarily result in uniform monotonicities regarding the price and modified base stock level for product 2.

2.6.2 Sensitivity to Capacity Parameters

Next, we consider how changes in capacity parameters influence the optimal policy. Theorem 2.4 parts (a) and (b) correspond to a capacity increase in either dedicated resource and the flexible resource, respectively.

Theorem 2.4. *(a) If the current period dedicated capacity for product* 1*,* K_1*, increases, the prices charged for both products decrease. The modified base stock level for product* 1 *increases. The modified base stock for product* 2 *increases if both items are critically understocked and the flexible resource is shared between the items, otherwise the modified base stock level for product* 2 *decreases.*

(b) If the current period flexible capacity K_0 *increases, the prices charged for both products decrease. The modified base stock level for item 1 (item 2) increases if item 1 (item 2) is critically understocked. Otherwise, the modified base-stock level for item 1 (item 2) decreases.*

Proof: The proof of Theorem 2.4 is provided in Section 2.9.3.

When capacity increases, one would expect that the price for both products would decrease. In Theorem 2.4 parts (a) and (b), we show that this expectation is true. A capacity increase in either the dedicated resource or the flexible resource helps reduce instances where products are critically understocked which limits price markups and hence reduces prices.

Regarding the modified base stock levels, Theorem 2.4 part (a) shows that an increase in the dedicated capacity for product 1 leads to an increase in the modified base stock level for item 1. When both items share the flexible resource and are critically understocked, an increase in the dedicated capacity for item 1 allows more flexible capacity to be allocated to item 2 increasing item 2's modified base stock level. In all other instances, the modified base stock level for item 2 decreases. As an example, consider the instance when item 1 is critically understocked while item 2 is moderately understocked. An increase in the dedicated capacity and thus the modified base stock level for item 1 results in less price surcharge for item 1 as we do not have to decrease product 1 demand as much by pricing which in turn results in less inventory of item 2 needed, i.e. a reduced modified base stock level. The logic of Theorem 2.4 part (b) is similar.

2.6.3 Sensitivity to Demand Parameters

Finally, we examine the effect of demand parameters on the optimal policy. Theorem 2.5 part (a) corresponds to an increase in the demand intercept for a product whereas part (b) and part (c) states the sensitivity with respect to individual and cross price elasticities, respectively.

Theorem 2.5. *(a) If the current period demand intercept* b_1 *for product* 1 *increases, the price for both items increase. The modified base stock level for product* 1 *increases. The modified base stock level for product* 2 *increases except when both items are critically understocked and the flexible resource is shared between the items, in which case the modified base stock level for product* 2 *decreases.*

(b) If the current period individual price elasticity a_{11} *for product* 1 *increases, the price and the modified base stock level for item* 1 *decreases.*

(c) If magnitude of the current period cross price elasticity a_{12} *increases, when both products are moderately understocked, the prices and base-stock levels for both products increase.*

Proof: The proof of Theorem 2.5 is provided in Section 2.9.3.

An increase in the demand intercept for product 1 can be viewed as an exogenous factor (such as an increase in the perceived quality) that makes product 1 more desirable. Theorem 2.5 part (a) shows that this increased demand allows a higher price to be charged for both products. (A similar relationship was also recently observed by Aydin and Porteus [3] for a one period problem without capacity considerations). As the demand and price for product 1 increases, its modified base stock level also increases. We find that a joint increase in prices also yields an increase in the modified base stock level for item 2 except when both items are critically understocked and requesting a share of the flexible capacity. In that case, the increase in the demand intercept for item 1 necessitates more of the flexible capacity to be allocated to item 1, and thus decreases the modified base stock level for item 2.

A higher individual price elasticity means fewer customers will demand the product at the current prices. Theorem 2.5 part (b) shows that when the demand for an item is more sensitive to its own price, it is optimal for the firm to counter this demand reduction by setting a lower price and modified base-stock level for that item. However, the price and modified base stock level of item 2 do not necessarily have uniform monotonicities with respect to an increase in item 1's individual price elasticity. Finally, part (c) corresponds to a setting where the magnitude of the cross price elasticity is higher. Such a setting results in a higher demand for both products and we show that it results in an increase in both prices and modified base stock levels when both products are moderately understocked. When either product is critically understocked or overstocked, the behavior of base stock levels and prices do not appear to possess uniform monotonicity.

2.7 Numerical Study

In this section, we first investigate how the availability of a flexible resource influences the firm's optimal pricing strategy. We then compare the economic benefits obtained by using dynamic pricing and capacity flexibility jointly and individually and explore settings under which dynamic pricing may be more valuable than capacity flexibility to improve profitability, and vice versa.

2.7.1 Impact of Capacity Flexibility on Optimal Pricing Policy

In order to explore the impact of flexible resources on the optimal pricing policy, we construct three problem instances with a gradually increasing share of a flexible resource in the firm's capacity portfolio. In the first setting, the firm implements dynamic pricing using only dedicated production capacities for each item with capacity installations $K_0 = 0$ and $K_1 = K_2 = 15$. In the second setting, the firm employs a portfolio of dedicated and flexible resources where $K_0 = K_1 = K_2 = 10$. Finally, in the third problem setting, the firm utilizes full flexibility in the production line with $K_0 = 30$ and $K_1 = K_2 = 0$.

For a $T = 15$ period problem initialized at state $\mathbf{x}^{15} = (0, 0)$, we run 500 randomly generated sample paths which follow the optimal production and pricing policies at each period until the end of the planning horizon. We observe the optimal price selections for both products at each period for the following demand model.

$$
\begin{aligned}
d_1^t(p_1^t, p_2^t, \epsilon_1^t) &= 35 - 0.75p_1^t + 0.25p_2^t + \epsilon_1^t \\
d_2^t(p_1^t, p_2^t, \epsilon_2^t) &= 30 + 0.25p_1^t - 0.5p_2^t + \epsilon_2^t
\end{aligned} \tag{2.6}
$$

The remaining problem parameters are set as $c_1 = 15, c_2 = 20, h_1 = 7.5, h_2 = 10, \pi_1 = 30, \pi_2 = 40$, and $\beta = 0.8$. We let ϵ_1^t and ϵ_2^t be randomly drawn from a uniform distribution over the interval [-10,10] with a positive support on the realized demand. The model yields list prices of $p_{1,L}^t = 47.5$ and $p_{2,L}^t = 60.0$ over the time horizon with list-price mean demands of $\bar{d}_1^t(p_1^t, p_2^t) = 14.4$ and $\bar{d}_2^t(p_1^t, p_2^t) = 11.9$ units per period for products 1 and 2. (This corresponds to dedicated capacity utilizations of 96% and 79%, respectively.) Table 2.1 reports the average and standard deviation of the prices and price differences observed along the planning horizon of 15 periods for the 500 randomly generated problem instances. We observe that the standard

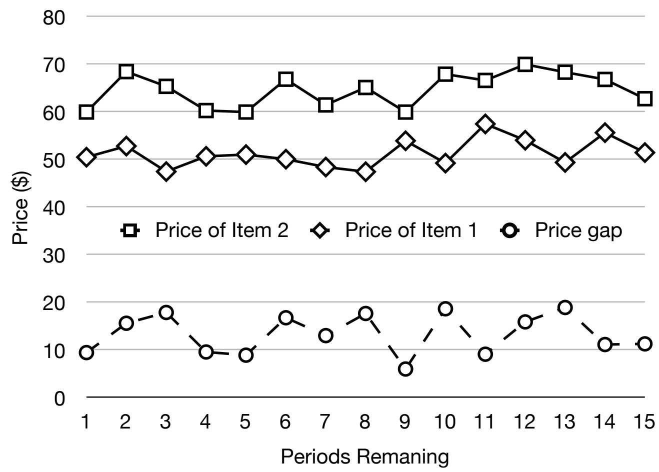
Table 2.1: Price statistics for systems with (i) only dedicated resources, (ii) a portfolio of dedicated and flexible resources and (iii) a fully flexible resource

	(i) $K_0 = 0$, $K_1 = K_2 = 15$	(ii) $K_0 = 10$, $K_1 = K_2 = 10$	(iii) $K_0 = 30$, $K_1 = K_2 = 0$
Average Price 1	49.83	49.00	48.85
Average Price 2	61.75	61.34	61.33
Std. Dev. of Price 1	2.62	2.10	2.05
Std. Dev. of Price 2	2.57	2.10	2.07
Std. Dev. of Price Difference	3.05	0.76	0.19

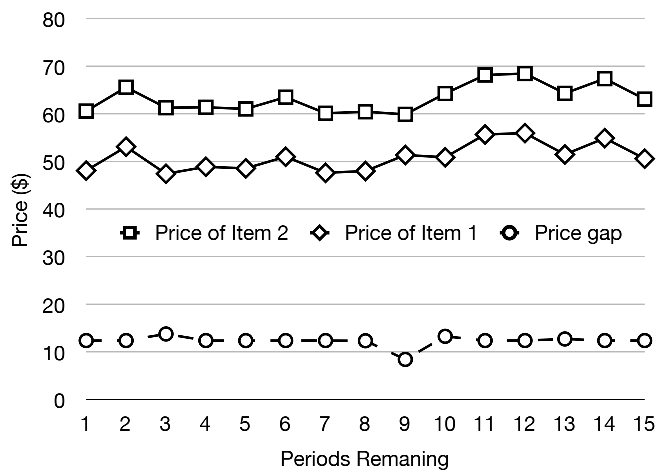
deviation of the prices charged for each item over time decreases as the share of the flexible resource in the capacity portfolio increases. Moreover, we see that the standard deviation of the price difference between products also decreases, and rather significantly, as the firm utilizes more flexible resources.

To visualize the effect of flexible capacity on the optimal pricing policy, we next illustrate a particular sample path over the 15-period horizon. Figure 2.5 depicts the optimal price selection at each decision period for the three settings and highlights the advantages of flexible resources. First, as implied by the reduction of the standard deviation of prices presented in Table 2.1, we observe that the prices for items 1 and 2 have somewhat smoother fluctuations across periods when capacity is more flexible. We calculate that the standard deviation of prices over the 15-period horizon for items 1 and 2 are, respectively, 3.48 and 3.31 for the dedicated capacity setting, 2.94 and 2.79 for the hybrid capacity portfolio setting, and 2.75 and 2.75 for the fully flexible capacity setting. This small sample of data which shows that there is a slight reduction in price fluctuations for an item over time but not a complete price smoothing also suggests that the presence of capacity flexibility does not entirely eliminate the need for dynamic pricing.

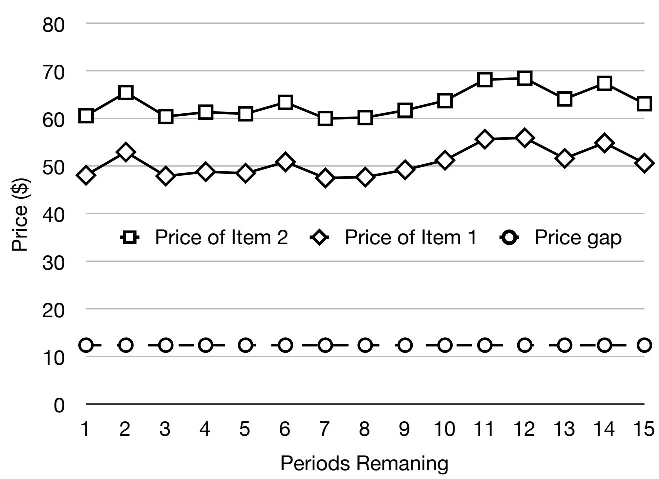
The most interesting aspect displayed by Figure 2.5 is that when flexible systems are used instead of product dedicated resources, the difference between the prices



(a) Only dedicated resources, K_0=0, K_1=K_2=15



(b) Portfolio of dedicated and flexible resources, K_0=K_1=K_2=10



(c) A fully flexible resource, K_0=30, K_1=K_2=0

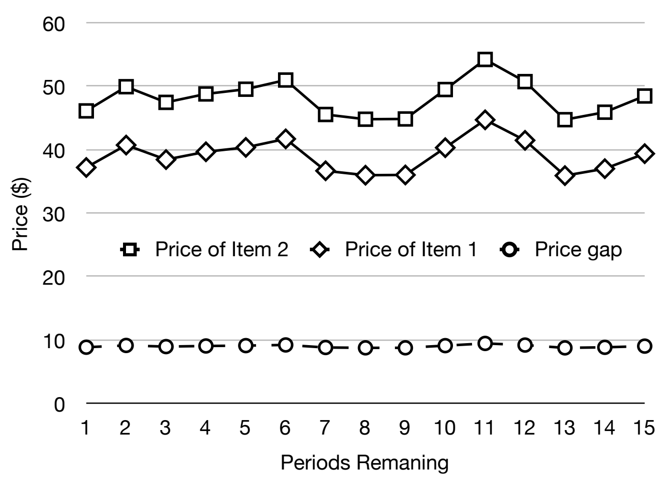
Figure 2.5: Optimal price selections for products 1 and 2 for a 15-period problem.

charged for products 1 and 2 remain almost constant across periods (compare the corresponding dotted lines in Figure 2.5). For this particular instance, we find that the standard deviations of the price difference between items 1 and 2 over the 15-period horizon are 4.44, 1.57, and 0 for the dedicated only, hybrid, and fully flexible capacity settings, respectively. This is in line with the statement in Theorem 2.2 that the price difference between the two products will be constant when both items are either moderately understocked or critically understocked and sharing the flexible resource. Table 2.1 and Figure 2.5 indicate that the availability of flexible capacity actually results in the price difference between the products to be fairly constant in most instances, aiding in the consistent price positioning of the products.

The extended constant price difference region set forth in Theorem 2.2 is based on the linear additive stochastic demand model with uniform cross price elasticities. However, we are also interested to explore whether flexible capacity continues to enable more stable price differences between items over time for nonuniform cross-price elasticities and for other possible demand models. Figure 2.6 illustrates a particular sample path based on a modified version of the previously described demand model in (2.6) with $b_1 = 45$, $b_2 = 35$, $a_{11} = 1.2$, $a_{12} = -0.3$, $a_{21} = -0.4$, and $a_{22} = 0.8$ and preserving the same capacity and cost parameters. We find that the standard deviation of the price difference between items 1 and 2 over the 15-period horizon is 3.47%, and 0.21% for the dedicated only, and fully flexible capacity settings, respectively. Thus, for this setting, the flexible capacity continues to help maintain stable price differences.

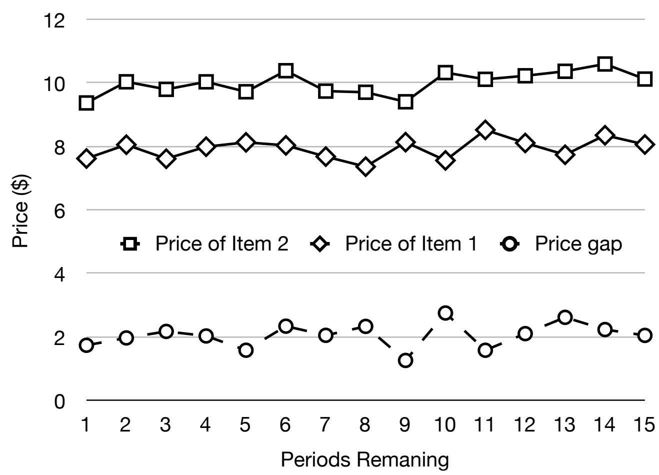
Next, we consider a Multinomial Logit (MNL) demand model based on consumer choice behavior. For a detailed discussion of MNL demand models in this context, we refer the reader to a study by Aydin and Porteus [3]. Following [3],

(a) Only dedicated resources, K_0=0, K_1=K_2=15

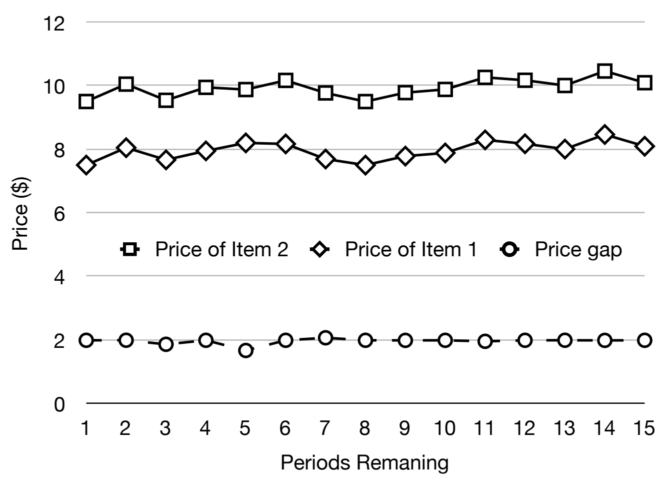


(b) A fully flexible resource, K_0=30, K_1=K_2=0

Figure 2.6: Non-uniform cross-price elasticities: Optimal price selections for products 1 and 2 for a 15-period problem.



(a) Only dedicated resources, K_0=0, K_1=K_2=15



(b) A fully flexible resource, K_0=30, K_1=K_2=0

Figure 2.7: Multinomial Logit (MNL) demand model: Optimal price selections for products 1 and 2 for a 15-period problem.

we let $u_i^t - p_i^t$ denote the surplus utility of a customer who purchases product i. We let the deterministic part of the demand for product i be given by $\bar{d}_i^t(p_1^t, p_2^t) = \Theta\left(\exp(u_i^t - p_i^t)\right) / \left(1 + \sum_j \exp(u_j^t - p_j^t)\right)$ where Θ denotes the market size. Figure 2.7 displays a sample path of product prices for a setting where $u_1 = 8$, $u_2 = 10$, $\Theta = 30$ and with $c_1 = 3, c_2 = 5, h_1 = 1.5, h_2 = 2.5, \pi_1 = 6, \pi_2 = 10$, and $\beta = 0.8$. We find that the insights we have gained by the linear demand model regarding the price gap stabilizing effects of flexible capacity continue to hold under the MNL demand model.

2.7.2 Economic Benefits of Dynamic Pricing and Capacity Flexibility

To explore the economic benefits obtained by dynamic pricing and capacity flexibility, we study several numerical examples where each strategy may be utilized jointly or individually. Specifically, we consider problem settings where the following strategies are implemented: fixed list prices with dedicated resources, fixed list priced with a fully flexible resource, dynamic pricing with dedicated resources, and dynamic pricing with a fully flexible resource.

For the example problem given in (2.6), we analyze how each strategy performs under instances with high or low demand uncertainty, and negatively or positively correlated product demands. We let the uncertain component of demand for the high and low variability settings be drawn from a uniform distribution over the interval [-10,10] and [-4,4], respectively with a positive support on the period's demand. For the positive and negative correlation settings, we use the high demand variability parameters with a correlation coefficient of $\rho = 1$ and $\rho = -1$, respectively. Table 2.2 displays the profit obtained by each strategy for a 5-period problem starting from an initial state of (x_1^{15}, x_2^{15})=$(0, 0)$.

Table 2.2: Economic benefits of dynamic pricing and/or capacity flexibility

	Fixed list price	Dynamic price
High demand variability		
Dedicated capacity	5,253	6,039
Flexible capacity	5,950	6,119
Low demand variability		
Dedicated capacity	6,511	6,637
Flexible capacity	6,619	6,670
Positively correlated demand		
Dedicated capacity	5,361	5,979
Flexible capacity	5,522	5,987
Negatively correlated demand		
Dedicated capacity	5,113	6,044
Flexible capacity	6,158	6,181

The profits reported in Table 2.2 shows that, for a system with high demand variability, dynamic pricing and capacity flexibility provide profit gains of 15.0% and 13.2%, respectively, compared to a base setting of fixed list prices and dedicated capacities. A joint strategy improves the profits by 16.5%. Furthermore, we calculate that, for a system with flexible capacity, the additional economic benefit derived by dynamic pricing is 2.8%. Similarly, the profit gain by flexibility for a system which already uses a dynamic pricing strategy is 1.3%. These results imply that, although dynamic pricing and capacity flexibility may be viewed as substitute approaches, neither strategy dominates the other and a joint strategy may still provide significant economic benefits. For a system with low demand variability, we observe that dynamic pricing and capacity flexibility provides economic benefits of 1.9% and 0.5%, respectively. Thus, the value of both tools increase significantly as the demand variability increases.

Regarding the correlation among demand uncertainties, we see that dynamic pricing is a more valuable tool than capacity flexibility in settings where demand uncertainties are positively correlated. We calculate that the respective profit gains are

11.5% and 3.0%. For negatively correlated demand uncertainties, on the other hand, we find that capacity flexibility is a slightly more powerful tool than dynamic pricing implied by a profit gain of 20.4% compared to 18.2%. Positively correlated demands may result in occasions where both items are overstocked or critically understocked. At such instances, the products will either jointly require or do not require additional capacity. Hence, capacity flexibility can only offer a marginal benefit, much less than the one obtained by dynamic pricing which can raise or reduce the demand for both products to prevent excessive holding or backorder costs. Negatively correlated demands may frequently result in instances when one item is critically understocked while the other is either moderately understocked or overstocked. Shifting demand by dynamic pricing at such instances alleviates the shortage costs arising due to the critically understocked item. However, this benefit comes at the expense of losing some of the overall product revenue due to the concavity of the revenue function. A flexible resource, on the other hand, may be used to shift production to the item with deficient supply enabling a reduction in the shortage cost without having an impact on the revenue.

2.8 Conclusions

In this chapter, we studied a joint mechanism of dynamic pricing and capacity flexibility to mitigate demand and supply mismatches. We considered a firm producing two products with correlated demands utilizing capacitated product dedicated and flexible resources and characterized the structure and sensitivity of the optimal production and pricing decisions.

Under a linear additive stochastic demand model that is commonly adapted in existing literature, we showed that the optimal production policy can be characterized by modified base-stock levels that exhibit distinct forms across two broad regions of the state-space. We presented the optimal policy by classifying the initial inventory level of a product as overstocked if the item requires no further replenishment, as moderately understocked if the available capacity is adequate to bring the inventory to a desired level, and as critically understocked if capacity is restrictive to reach the desired inventory level. Our analysis showed that when at most one item is critically understocked, the modified base-stock level for each product is described by a decreasing function of the inventory level of the other item. However, when both items are critically understocked, we showed that the modified base-stock level for a product is characterized by an increasing function of the inventory position of both products.

In terms of the pricing policy, our results showed that a list price is charged for an item if it is moderately understocked. If an item is critically understocked, then a price markup that depends on both inventory levels is applied. When an item is overstocked, a price discount that depends on both inventory levels is given. Our analysis also indicated that when inventory levels for both items are critically understocked and when the flexible capacity is simultaneously shared between products, the flexible resource led to an optimal pricing scheme that maintained a constant price difference between products. At such instances, dynamic pricing only served to adjust the overall level of demand for both products and not to attempt to shift demand from one product to another. The flexible capacity was the sole factor in restoring the mismatches between the desired and actual inventory levels of products.

We found that the presence of a flexible resource may significantly reduce the fluctuations of price differences across items over time. Thus, the existence of a flexible resource in the firm's capacity portfolio helps maintain stable price differences across items over time. This enables the firm to establish consistent price positioning among multiple products even if it uses a dynamic pricing strategy. Regarding the contribution to a firm's profitability, we find that the two mechanisms may be viewed as substitute, but not fully interchangeable approaches and that using a joint strategy may improve profits significantly. We find dynamic pricing to be the dominant contributor to increased profits when demand uncertainties for the products are positively correlated and flexible capacity to be the more powerful contributor when there is a negative correlation between demand uncertainties.

Our focus in this chapter has been limited to studying price based substitutions between the products. It remains an interesting question for future research to identify how stockout based substitutions such as upgrading a customer to a higher quality item when the low quality item experiences shortages would effect the firm's optimal pricing and production policy. In addition, the flexible resource may be viewed as a dual source option with a higher production or ordering cost. Earlier work in such models with exogenous demand processes indicate that a two-tier base stock policy is optimal. Incorporating pricing decisions to investigate how the dual sourcing option influences the firm's pricing policy and vice versa may constitute another interesting research question.

2.9 Appendix

2.9.1 Proofs of Preserved Structural Properties

Proof of Lemma 2.1: (Concavity)

The proof is by induction. $J^1(z_1^1, z_2^1, p_1^1, p_2^1)$ is separable in (z_1^1, z_2^1) and (p_1^1, p_2^1) and it is straightforward to check that $R'(\mathbf{p}^1)$ is strictly concave in (p_1^1, p_2^1) and the terms associated with holding and backorder costs are concave in (z_1^1, z_2^1). Since

$J^1(z_1^1, z_2^1, p_1^1, p_2^1)$ is formed by the addition of strictly concave, concave and linear functions, itself is strictly concave. Next, note that the capacity constraints result in a convex domain over which the maximization is performed. Since concavity is preserved under maximization in a convex domain [33], we have $V^1(x_1^1, x_2^1)$ strictly concave. Now, assume that $V^t(x_1^t, x_2^t)$ is strictly concave which implies, through a similar argument, that $J^{t+1}(z_1^{t+1}, z_2^{t+1}, p_1^{t+1}, p_2^{t+1})$ is strictly concave. Again, due to concavity preservation under maximization in a convex domain, we have $V^{t+1}(x_1^{t+1}, x_2^{t+1})$ strictly concave. $\square$

The Karush-Kuhn-Tucker Optimality Conditions:

To construct the KKT optimality conditions, we first introduce Lagrange multipliers $\lambda_{ij}^t > 0$ for $i, j = \{1, 2\}$ and $\mu^t > 0$ where $\lambda_{i1}^t > 0$ and $\lambda_{i2}^t > 0$ are associated with constraints $z_i^t + b_i - a_{i1}p_1^t - a_{i2}p_2^t \geq x_i^t$ and $z_i^t + b_i - a_{i1}p_1^t - a_{i2}p_2^t \leq x_i^t + K_0 + K_i$, respectively and μ^t corresponds to the constraint $z_1^t + z_2^t + b_1 + b_2 - (a_{11} + a_{21})p_1^t - (a_{12} + a_{22})p_2^t \leq x_1^t + x_2^t + K_0 + K_1 + K_2$. Together with the complementary slackness conditions, we then have for i={1,2},

$$\frac{\partial J^t}{\partial p_i^t} = a_{1i}(\lambda_{11}^t - \lambda_{12}^t) + a_{2i}(\lambda_{21}^t - \lambda_{22}^t) - (a_{1i} + a_{2i})\mu^t \tag{2.7a}$$

$$\frac{\partial J^t}{\partial z_i} = \mu^t - (\lambda_{i1}^t - \lambda_{i2}^t) \tag{2.7b}$$

Several pairs of constraints form "box constraints" and may not be simultaneously active for positive capacity parameters. As the following observation suggests, we can exploit this complementary sparsity pattern arising from the special structure of constraints to represent the first-order optimality conditions in simpler notation.

Observation 2.1. *For $i = 1, 2$, let λ_i^t be defined such that $\lambda_i^t := \lambda_{i1}^t - \lambda_{i2}^t$. Then, λ_i^t uniquely determines λ_{ij}^t for $j = 1, 2$ where (a) $\lambda_i^t < 0$ implies $\lambda_{i1}^t = 0$ and $\lambda_{i2}^t > 0$, (b) $\lambda_i^t > 0$ implies $\lambda_{i1}^t > 0$ and $\lambda_{i2}^t = 0$; and (c) $\lambda_i^t = 0$ implies $\lambda_{i1}^t = \lambda_{i2}^t = 0$.*

In addition, for $= \{1,2\}$*, the conditions given in the set of equations (2.7) may be represented as:*

$$\frac{\partial J^t}{\partial p_i^t} = a_{1i}\lambda_1^t + a_{2i}\lambda_2^t - (a_{1i}+a_{2i})\mu^t \tag{2.8a}$$

$$\frac{\partial J^t}{\partial z_i} = \mu^t - \lambda_i^t \tag{2.8b}$$

Proof: We first observe that having $\lambda_{11}^t > 0$ and $\lambda_{12}^t > 0$ simultaneously, implies that both $z_1^t + b_1 - a_{11}p_1^t - a_{12}p_2^t - x_1^t = 0$ and $z_1^t + b_1 - a_{11}p_1^t - a_{12}p_2^t - x_1^t - K_0 - K_1 = 0$. Since this is not possible for any $K_0, K_1 > 0$, we conclude that λ_{11}^t and λ_{12}^t cannot be simultaneously positive. Thus, if we define $\lambda_1^t := \lambda_{11}^t - \lambda_{12}^t$, any value of λ_1^t uniquely determines the values of λ_{11}^t and λ_{12}^t. We note that with this definition, λ_1^t is no longer sign restricted. Specifically, we have $\lambda_1^t < 0$ for the case where $\lambda_{11}^t = 0$, $\lambda_{12}^t > 0$, and we have $\lambda_1^t > 0$ for the case where $\lambda_{11}^t > 0$ and $\lambda_{12}^t = 0$. For the case where $\lambda_{11}^t = \lambda_{12}^t = 0$, we have $\lambda_1^t = 0$. An analogous argument holds for λ_{21}^t and λ_{22}^t, hence a corresponding $\lambda_2^t := \lambda_{21}^t - \lambda_{22}^t$ can be similarly defined. $\square$

Following Observation 2.1, λ_i^t is no longer sign restricted and is associated with two constraints where its sign - negative, positive or zero - identifies which of the corresponding constraints, if any, is binding.

Proof of Lemma 2.2: (Submodularity, Diagonal Dominance)

The proof is by induction. To simplify the notation, recalling that J^t is separable in (z_1^t, z_2^t) and (p_1^t, p_2^t), we let $J_{ij}^t := \frac{\partial^2 J^t}{\partial z_i^t \partial z_j^t}$. Similarly, we also let $V_{ij}^t = \frac{\partial^2 V^t}{\partial x_i^t \partial x_j^t}$. For $J^1(z_1^1, z_2^1, p_1^1, p_2^1)$, both cross partials are zero, thus part (1) follows. For part (2), we note that, by Lemma 2.1, $J^1(z_1^1, z_2^1, p_1^1, p_2^1)$ is strictly concave. Therefore, $J_{11}^1, J_{22}^1 < 0$. Hence part (2) follows. We now assume that the Lemma holds for t and show that it continues to hold for $t+1$. It is sufficient to show that $\mathbb{E}V^t(z_1^t - \epsilon_1^t, z_2^t - \epsilon_2^t)$ preserves these properties. It can be verified recursively that the first and second derivatives of $V^t(x_1^t, x_2^t)$ are bounded. Through the interchangeability of differentiation and expectation, it is sufficient to show that $V^t(x_1^t, x_2^t)$ has the required properties. From Envelope Theorem, we have

$$\frac{\partial V^t(x_1^t, x_2^t)}{\partial x_1^t} = \frac{\partial J^t}{\partial x_1^t} - \lambda_1^t + \mu^t = c_1 - \lambda_1^t + \mu^t \tag{2.9a}$$

$$\frac{\partial V^t(x_1^t, x_2^t)}{\partial x_2^t} = \frac{\partial J^t}{\partial x_2^t} - \lambda_2^t + \mu^t = c_2 - \lambda_2^t + \mu^t \tag{2.9b}$$

At this point, it is helpful to partition the state space in two broad regions where $\mu^t = 0$ and $\mu^t > 0$. We first treat the cases associated with $\mu^t = 0$. For these cases, we have

$$V_{12}^t(x_1^t, x_2^t) = -\frac{\partial \lambda_1^t}{\partial x_2^t} \tag{2.10}$$

From the KKT conditions, we further have

$$-\frac{\partial \lambda_1^t}{\partial x_2^t} = \frac{\partial}{\partial x_2^t}\left(\frac{\partial J^t}{\partial z_1^t}\right)$$

Therefore,

$$V_{12}^t(x_1^t, x_2^t) = \frac{\partial}{\partial x_2^t}\left(\frac{\partial J^t}{\partial z_1^t}\right) = J_{11}^t \frac{\partial z_1^t}{\partial x_2^t} + J_{12}^t \frac{\partial z_2^t}{\partial x_2^t} \tag{2.11}$$

We implicitly assume $V_{12}^t(x_1^t, x_2^t) = V_{21}^t(x_1^t, x_2^t)$ which requires continuity of the second partial derivatives. This is fulfilled since J^t is strictly concave and twice continuously differentiable in (z_1^t, z_2^t) and z_1^t, z_2^t are differentiable in (x_1^t, x_2^t).

There are four cases: (1) $\lambda_1^t = 0$ or $\lambda_2^t = 0$, (2) $\lambda_1^t > 0$ and $\lambda_2^t > 0$, (3) $\lambda_1^t > 0$ and $\lambda_2^t < 0$, and (4) $\lambda_1^t < 0$ and $\lambda_2^t > 0$. Note that, the case $\lambda_1^t < 0$ and $\lambda_2^t < 0$ is not feasible since this case would have implied that the flexible capacity is fully utilized for both products simultaneously.

Case 1: When $\lambda_1^t = 0$, we have $\frac{\partial \lambda_1^t}{\partial x_2^t} = 0$. Thus $V_{12}^t(x_1^t, x_2^t) = -\frac{\partial \lambda_1^t}{\partial x_2^t} = 0$. A similar argument for $\lambda_2^t = 0$ also yields $V_{12}^t(x_1^t, x_2^t) = -\frac{\partial \lambda_2^t}{\partial x_1^t} = 0$. This establishes the result for part (1), i.e., that $V^t(x_1^t, x_2^t)$ is submodular.

For part (2), since $V^t(x_1^t, x_2^t)$ is strictly concave, we have $V_{11}^t(x_1^t, x_2^t) < 0$ and $V_{22}^t(x_1^t, x_2^t) < 0$, hence $V_{11}^t(x_1^t, x_2^t) \leq V_{12}^t(x_1^t, x_2^t)$ and $V_{22}^t(x_1^t, x_2^t) \leq V_{12}^t(x_1^t, x_2^t)$ hold.

Case 2: When $\lambda_1^t > 0$ and $\lambda_2^t > 0$, from KKT conditions we have

$$p_1^t = p_{1L}^t + \frac{1}{2}\frac{\partial J^t}{\partial z_1^t}$$
$$p_2^t = p_{2L}^t + \frac{1}{2}\frac{\partial J^t}{\partial z_2^t}$$

where

$$p_{1L}^t = \frac{a_{22}b_1 - a_{12}b_2}{2(a_{11}a_{22} - a_{12}^2)} + \frac{c_1}{2}$$
$$p_{2L}^t = \frac{a_{11}b_2 - a_{12}b_1}{2(a_{11}a_{22} - a_{12}^2)} + \frac{c_2}{2}$$

Complementary slackness yields $x_1^t = z_1^t + b_1 - a_{11}p_1^t - a_{12}p_2^t$ and $x_2^t = z_2^t + b_2 - a_{21}p_1^t - a_{22}p_2^t$. Combining these we get,

$$\begin{aligned} x_1^t &= z_1^t + b_1 - a_{11}p_{1L}^t - a_{12}p_{2L}^t - \frac{a_{11}}{2}\frac{\partial J^t}{\partial z_1^t} - \frac{a_{12}}{2}\frac{\partial J^t}{\partial z_2^t} \\ x_2^t &= z_2^t + b_2 - a_{21}p_{1L}^t - a_{22}p_{2L}^t - \frac{a_{21}}{2}\frac{\partial J^t}{\partial z_1^t} - \frac{a_{22}}{2}\frac{\partial J^t}{\partial z_2^t} \end{aligned} \tag{2.12}$$

Taking partial derivatives with respect to x_2^t and solving for $\frac{\partial z_1^t}{\partial x_2^t}$ and $\frac{\partial z_2^t}{\partial x_2^t}$, we get

$$\begin{aligned} \frac{\partial z_1^t}{\partial x_2^t} &= \frac{1}{\Lambda}\left(a_{11}J_{12}^t + a_{12}J_{22}^t\right) \\ \frac{\partial z_2^t}{\partial x_2^t} &= \frac{1}{\Lambda}\left(2 - a_{11}J_{11}^t - a_{12}J_{21}^t\right) \end{aligned} \tag{2.13}$$

where $\Lambda = 2 - (a_{11}J_{11}^t + 2a_{12}J_{12}^t + a_{22}J_{22}^t) + \frac{1}{2}\left[(a_{11}a_{22} - a_{12}^2)(J_{11}^t J_{22}^t - J_{12}^{2t})\right]$. We note that $\Lambda > 0$ by first observing that the terms in the brackets are strictly positive since $a_{11}a_{22} - a_{12}^2 > 0$ by the assumptions on demand parameters and $J_{11}^t J_{22}^t - J_{12}^{2t} > 0$ is strictly concave as shown in Lemma 2.1. We only need to show that $a_{11}J_{11}^t + 2a_{12}J_{12}^t + a_{22}J_{22}^t \leq 0$. We have,

$$\begin{aligned} a_{11}J_{11}^t + 2a_{12}J_{12}^t + a_{22}J_{22}^t &\leq (a_{11} + 2a_{12} + a_{22})\, J_{12}^t \quad \text{(by diagonal dominance)} \\ &\leq 0 \quad (\text{since } \ a_{11} + a_{12} > 0,\ a_{12} + a_{22} > 0 \text{ and } \ J_{12}^t \leq 0) \end{aligned}$$

Substituting (2.13) into (2.11) establishes submodularity as follows:

$$\begin{aligned}
V^t_{12}(x^t_1, x^t_2) &= \frac{1}{\Lambda}\left(J^t_{11}\left(a_{11}J^t_{12} + a_{12}J^t_{22}\right) + J^t_{12}\left(2 - a_{11}J^t_{11} - a_{12}J^t_{21}\right)\right) \\
&= \frac{1}{\Lambda}\left(a_{12}\left(J^t_{11}J^t_{22} - J^{2t}_{12}\right) + 2J^t_{12}\right) \\
&\leq 0 \ \ (\text{since } \ J^t_{12} \leq 0, J^t_{11}J^t_{22} - J^{2t}_{12} > 0, \text{ and } \ a_{12} < 0)
\end{aligned}$$

To show part (2), we first state the expressions for $V^t_{11}(x^t_1, x^t_2)$ and $V^t_{22}(x^t_1, x^t_2)$ as well as evaluate $\frac{\partial z^t_1}{\partial x^t_1}$ and $\frac{\partial z^t_2}{\partial x^t_1}$. Following similar steps as in (2.10)-(2.11), we have;

$$V^t_{11}(x^t_1, x^t_2) = \frac{\partial}{\partial x^t_1}\left(\frac{\partial J^t}{\partial z^t_1}\right) = J^t_{11}\frac{\partial z^t_1}{\partial x^t_1} + J^t_{12}\frac{\partial z^t_2}{\partial x^t_1} \tag{2.14}$$

$$V^t_{22}(x^t_1, x^t_2) = \frac{\partial}{\partial x^t_2}\left(\frac{\partial J^t}{\partial z^t_2}\right) = J^t_{21}\frac{\partial z^t_1}{\partial x^t_2} + J^t_{22}\frac{\partial z^t_2}{\partial x^t_2} \tag{2.15}$$

and with analogous arguments as in (2.12)-(2.13), we get

$$\begin{aligned}
\frac{\partial z^t_1}{\partial x^t_1} =& \frac{1}{\Lambda}\left(2 - a_{12}J^t_{12} - a_{22}J^t_{22}\right) \\
\frac{\partial z^t_2}{\partial x^t_1} =& \frac{1}{\Lambda}\left(a_{12}J^t_{11} + a_{22}J^t_{12}\right)
\end{aligned} \tag{2.16}$$

We now consider $V^t_{11}(x^t_1, x^t_2) \leq V^t_{12}(x^t_1, x^t_2)$. Substituting (2.16) into (2.14), we get

$$\begin{aligned}
V^t_{11}(x^t_1, x^t_2) &= \frac{1}{\Lambda}\left(J^t_{11}\left(2 - a_{12}J^t_{12} - a_{22}J^t_{22}\right) + J^t_{12}\left(a_{12}J^t_{11} + a_{22}J^t_{12}\right)\right) \\
&= \frac{1}{\Lambda}\left(2J^t_{11} - a_{22}\left(J^t_{11}J^t_{22} - J^{2t}_{12}\right)\right)
\end{aligned}$$

Therefore we get

$$\begin{aligned}
V^t_{11}(x^t_1, x^t_2) - V^t_{12}(x^t_1, x^t_2) &= \frac{1}{\Lambda}\left(\left(2J^t_{11} - a_{22}\left(J^t_{11}J^t_{22} - J^{2t}_{12}\right)\right)\right. \\
&\qquad \left. - \left(2J^t_{12} + a_{12}\left(J^t_{11}J^t_{22} - J^{2t}_{12}\right)\right)\right) \\
&= \frac{1}{\Lambda}\left(\left(2J^t_{11} - 2J^t_{12}\right) - (a_{12} + a_{22})\left(J^t_{11}J^t_{22} - J^{2t}_{12}\right)\right) \\
&\leq 0 \quad (\text{by } \ J^t_{11} - J^t_{12} \leq 0, a_{12} + a_{22} > 0, \text{and concavity})
\end{aligned}$$

The fact that $V^t_{22}(x^t_1, x^t_2) \leq V^t_{12}(x^t_1, x^t_2)$ may be shown similarly by substituting (2.16) into (2.15).

The analysis for Cases 3 and 4 are very similar to the analysis of Case 2 and are omitted for brevity.

We now consider the regions corresponding to $\mu^t > 0$. By the definition of the multipliers and their relationships among each other, this region is subdivided into three subregions such that (1) $\mu^t > 0$, $\lambda_1^t < 0$, $\lambda_2^t = 0$; (2) $\mu^t > 0$, $\lambda_1^t = 0$, $\lambda_2^t = 0$; (3) $\mu^t > 0$, $\lambda_1^t = 0$, $\lambda_2^t < 0$.

Case 1 corresponds to the regions where the flexible capacity is used solely and fully to produce item 1. Once again, the Envelope Theorem yields

$$V_{12}^t(x_1^t, x_2^t) = \frac{\partial}{\partial x_2^t}\left(\frac{\partial J^t}{\partial z_1^t}\right) = J_{11}^t \frac{\partial z_1^t}{\partial x_2^t} + J_{12}^t \frac{\partial z_2^t}{\partial x_2^t} \tag{2.17}$$

Complementary slackness conditions yield $x_1^t = z_1^t + b_1 - a_{11}p_1^t - a_{12}p_2^t - K_0 - K_1$ and $x_2^t = z_2^t + b_2 - a_{21}p_1^t - a_{22}p_2^t - K_2$. Combining these, we get a similar expression as in (2.12):

$$\begin{aligned} x_1^t =& z_1^t + b_1 - a_{11}p_{1L}^t - a_{12}p_{2L}^t - K_0 - K_1 - \frac{a_{11}}{2}\frac{\partial J^t}{\partial z_1^t} - \frac{a_{12}}{2}\frac{\partial J^t}{\partial z_2^t} \\ x_2^t =& z_2^t + b_2 - a_{21}p_{1L}^t - a_{22}p_{2L}^t - K_2 - \frac{a_{21}}{2}\frac{\partial J^t}{\partial z_1^t} - \frac{a_{22}}{2}\frac{\partial J^t}{\partial z_2^t} \end{aligned}$$

The same arguments as presented in the analysis of the previous case yields the desired result. Further, the analysis for Case 3 is also symmetric to the analysis of Case 1 and hence omitted. Case 2 defines the only remaining region and it corresponds to $\mu^t > 0$, $\lambda_1^t = 0$, $\lambda_2^t = 0$, where the flexible capacity is used fully and by both products simultaneously.

In this region we have

$$x_1^t + x_2^t = z_1^t + z_2^t + b_1 + b_2 - (a_{11} + a_{21})p_1^t - (a_{12} + a_{22})p_2^t - K_0 - K_1 - K_2 \tag{2.18}$$

Differentiating (2.18) with respect to x_2^t, we get:

$$\begin{aligned} 1 = &\left(1 - \frac{(a_{11}+a_{21})}{2} J_{11}^t - \frac{(a_{12}+a_{22})}{2} J_{21}^t\right) \frac{\partial z_1^t}{\partial x_2^t} \\ &+ \left(1 - \frac{(a_{11}+a_{21})}{2} J_{12}^t - \frac{(a_{12}+a_{22})}{2} J_{22}^t\right) \frac{\partial z_2^t}{\partial x_2^t} \end{aligned} \tag{2.19}$$

Through the KKT conditions, in this region we also have $J_1^t = J_2^t$, hence differentiating with respect to x_2^t we get

$$J_{11}^t \frac{\partial z_1^t}{\partial x_2^t} + J_{12}^t \frac{\partial z_2^t}{\partial x_2^t} = J_{21}^t \frac{\partial z_1^t}{\partial x_2^t} + J_{22}^t \frac{\partial z_2^t}{\partial x_2^t} \tag{2.20}$$

Combining (2.19) and (2.20), we get

$$\begin{aligned} \frac{\partial z_1^t}{\partial x_2^t} =& \frac{1}{\Lambda'}(J_{12}^t - J_{22}^t) \\ \frac{\partial z_2^t}{\partial x_2^t} =& \frac{1}{\Lambda'}(J_{12}^t - J_{11}^t) \end{aligned} \tag{2.21}$$

where $\Lambda' = -J_{12}^t + 2J_{12}^t - J_{12}^t + (a_{11}+a_{12})(J_{11}^t J_{22}^t - J_{12}^{2t})/2 + (a_{22}+a_{12})(J_{11}^t J_{22}^t - J_{12}^{2t})/2$. As in the previous discussion for Λ, it can easily be shown that $\Lambda' > 0$. Substituting (2.21) into (2.11), we get

$$\begin{aligned} V_{12}^t(x_1^t, x_2^t) &= \frac{1}{\Lambda}\left(J_{11}^t\left(J_{12}^t - J_{22}^t\right) + J_{12}^t\left(J_{12}^t - J_{11}^t\right)\right) \\ &= \frac{1}{\Lambda}\left(J_{12}^{2t} - J_{11}^t J_{22}^t\right) \\ &\leq 0 \end{aligned}$$

To show part (2); we first consider $V_{11}^t(x_1^t, x_2^t) \leq V_{12}^t(x_1^t, x_2^t)$. Using the above argument, we find $V_{11}^t(x_1^t, x_2^t) - V_{12}^t(x_1^t, x_2^t) = \frac{1}{\Lambda}\left(J_{12}^{2t} - J_{11}^t J_{22}^t\right) \leq 0$. Similarly, $V_{22}^t(x_1^t, x_2^t) - V_{12}^t(x_1^t, x_2^t) = \frac{1}{\Lambda}\left(J_{12}^{2t} - J_{11}^t J_{22}^t\right) \leq 0$. □

2.9.2 Proofs of Optimal Policy Structure

Proof of Theorem 2.1 (Optimal Production Policy):

We utilize the shadow prices on capacity constraints by both determining the pricing and production decisions in terms of these Lagrangian variables and partitioning the state space with respect to their signs. We start by solving (2.8a) for p_1^t and p_2^t and find

$$p_i^t = p_{i,L}^t - \frac{1}{2}(\lambda_i^t - \mu^t) \tag{2.22}$$

where we define $p_{i,L}^t$ as the list-price in period t for product i, $i = \{1, 2\}$, given by the following expressions.

$$p_{1,L}^t = \frac{a_{22}b_1 - a_{12}b_2}{2(a_{11}a_{22} - a_{12}^2)} + \frac{c_1}{2} \tag{2.23a}$$

$$p_{2,L}^t = \frac{a_{11}b_2 - a_{12}b_1}{2(a_{11}a_{22} - a_{12}^2)} + \frac{c_2}{2} \tag{2.23b}$$

Further, (2.8b) imply implicit functions ϕ_1^t and ϕ_2^t such that $z_1^t = \phi_1^t(\lambda_1^t, \lambda_2^t, \mu^t)$ and $z_2^t = \phi_2^t(\lambda_1^t, \lambda_2^t, \mu^t)$ as stated in Lemma 2.3 below.

Lemma 2.3. *There exists implicit functions ϕ_1 and ϕ_2 such that $z_1^t = \phi_1(\lambda_1^t, \lambda_2^t, \mu^t)$ and $z_2^t = \phi_2(\lambda_1^t, \lambda_2^t, \mu^t)$. Furthermore, $\phi_1(\lambda_1^t, \lambda_2^t, \mu^t)$ is increasing in λ_1^t, and decreasing in λ_2^t, μ^t whereas $\phi_2(\lambda_1^t, \lambda_2^t, \mu^t)$ is increasing in λ_2^t, and decreasing in λ_1^t, μ^t.*

Proof: We first introduce two functions $F_1(\mathbf{L}^t, \mathbf{z}^t)$ and $F_2(\mathbf{L}^t, \mathbf{z}^t)$ where $\mathbf{L}^t = (\lambda_1^t, \lambda_2^t, \mu^t)$ and $\mathbf{z}^t = (z_1^t, z_2^t)$. We define these functions to represent KKT conditions (2.8b).

$$F_1(\mathbf{L}^t, \mathbf{z}^t) = J_1^t(z_1^t, z_2^t) + \lambda_1^t - \mu^t \tag{2.24a}$$

$$F_2(\mathbf{L}^t, \mathbf{z}^t) = J_2^t(z_1^t, z_2^t) + \lambda_2^t - \mu^t \tag{2.24b}$$

Differentiating (2.24a) and (2.24b), and letting $D_{\mathbf{L}}F$ and $D_{\mathbf{z}}F$ to denote partial Jacobians, we have

$$D_{\mathbf{L}}F = \begin{bmatrix} \frac{\partial F_1}{\partial \lambda_1} & \frac{\partial F_1}{\partial \lambda_2} & \frac{\partial F_1}{\partial \mu} \\ \frac{\partial F_2}{\partial \lambda_1} & \frac{\partial F_2}{\partial \lambda_2} & \frac{\partial F_2}{\partial \mu} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}, \quad D_{\mathbf{z}}F = \begin{bmatrix} \frac{\partial F_1}{\partial z_1} & \frac{\partial F_1}{\partial z_2} \\ \frac{\partial F_2}{\partial z_2} & \frac{\partial F_2}{\partial z_2} \end{bmatrix} = \begin{bmatrix} J_{11}^t & J_{12}^t \\ J_{21}^t & J_{22}^t \end{bmatrix}$$

Since $J^t(z_1^t, z_2^t)$ is strictly concave by Lemma 2.1, $D_{\mathbf{z}}F$ is invertible. Thus, there exists implicit functions ϕ_1 and ϕ_2 such that $z_1^t = \phi_1(\lambda_1^t, \lambda_2^t, \mu^t)$ and $z_2^t = \phi_2(\lambda_1^t, \lambda_2^t, \mu^t)$. Moreover, by the Implicit Function Theorem, we have

$$D\phi = -D_{\mathbf{z}}F^{-1}D_{\mathbf{L}}F$$

that is,

$$\begin{aligned}
\begin{bmatrix} \frac{\partial\phi_1}{\partial\lambda_1} & \frac{\partial\phi_1}{\partial\lambda_2} & \frac{\partial\phi_1}{\partial\mu} \\ \frac{\partial\phi_2}{\partial\lambda_1} & \frac{\partial\phi_2}{\partial\lambda_2} & \frac{\partial\phi_2}{\partial\mu} \end{bmatrix} &= -\begin{bmatrix} J_{11}^t & J_{12}^t \\ J_{21}^t & J_{22}^t \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \\
&= \frac{1}{J_{11}^t J_{22}^t - (J_{12}^t)^2} \begin{bmatrix} -J_{22}^t & J_{12}^t & J_{22}^t - J_{12}^t \\ J_{12}^t & -J_{11}^t & J_{11}^t - J_{12}^t \end{bmatrix} \qquad (2.25)
\end{aligned}$$

The strict concavity established in Lemma 2.1 yields $J_{11}^t < 0, J_{22}^t < 0$, and $J_{11}^t J_{22}^t - (J_{12}^t)^2 > 0$. The submodularity and diagonal dominance properties in Lemma 2.2, gives $J_{12}^t \leq 0, J_{11}^t - J_{12}^t \leq 0, J_{22}^t - J_{12}^t \leq 0$. Therefore, the monotonicity results follow immediately. □

By Lemma 2.3, we can rewrite the capacity constraints as follows.

$$x_1^t \leq \phi_1^t(\lambda_1^t, \lambda_2^t, \mu^t) - a_{11}p_1^t(\lambda_1^t, \mu^t) - a_{12}p_2^t(\lambda_2^t, \mu^t) + b_1 \leq x_1^t + K_0 + K_1 \qquad (2.26a)$$

$$x_2^t \leq \phi_2^t(\lambda_1^t, \lambda_2^t, \mu^t) - a_{21}p_1^t(\lambda_1^t, \mu^t) - a_{22}p_2^t(\lambda_2^t, \mu^t) + b_2 \leq x_2^t + K_0 + K_2 \qquad (2.26b)$$

$$\begin{aligned}
&\phi_1^t(\lambda_1^t, \lambda_2^t, \mu^t) + \phi_2^t(\lambda_1^t, \lambda_2^t, \mu^t) - (a_{11} + a_{21})p_1^t(\lambda_1^t, \mu^t) \\
&\qquad - (a_{12} + a_{22})p_2^t(\lambda_2^t, \mu^t) + b_1 + b_2 \leq x_1^t + x_2^t + K_0 + K_1 + K_2
\end{aligned} \qquad (2.26c)$$

The inventory state space may be partitioned into several regions based on the signs of λ_1^t, λ_2^t, and μ^t. In order to clarify the portrayal of state space segmentation, we define two broad regions: region A and region B, corresponding to initial inventory levels for which $\mu^t = 0$ and $\mu^t > 0$, respectively. In words, region A represents the initial inventory levels for which there remains some resource, either dedicated or flexible, that is not fully utilized. Region B, on the other hand, corresponds to inventory levels for which all resources are fully utilized.

A specific point is of certain interest in our partitioning of the state space. When none of the constraints are binding, we have $\lambda_1^t = \lambda_2^t = \mu^t = 0$ hence, $(\boldsymbol{\phi}^t(0,0,0), \mathbf{p_L^t})$ is the optimal solution to the unconstrained problem of $\max J^t(\mathbf{z}^t, \mathbf{p}^t)$. If we define $\bar{\mathbf{x}}^t$ such that

$$\bar{x}_1^t = \phi_1^t(0,0,0) - a_{11}p_{1L}^t - a_{12}p_{2L}^t + b_1 \tag{2.27a}$$

$$\bar{x}_2^t = \phi_2^t(0,0,0) - a_{21}p_{1L}^t - a_{22}p_{2L}^t + b_2 \tag{2.27b}$$

then, $(\bar{\mathbf{x}}^t, \mathbf{p}_L^t)$ is the optimal solution for $\max G^t(\mathbf{y}^t, \mathbf{p}^t)$, the unconstrained original problem.

Lemma 2.4. *The boundaries of the state space Region A are defined by two monotone functions:*

i. $\gamma_1^t(x_2^t) : \Re \rightarrow \Re$ *with* $\gamma_1^t(x_2^t) = \bar{x}_1^t$ *for* $x_2^t \in [\bar{x}_2^t - K_0 - K_2,\ \bar{x}_2^t]$ *and* $\gamma_1^t(x_2^t)$ *strictly decreasing with respect to* x_2^t *for* $x_2^t \in \Re \setminus [\bar{x}_2^t - K_0 - K_2,\ \bar{x}_2^t]$.

ii. $\gamma_2^t(x_1^t) : \Re \rightarrow \Re$ *with* $\gamma_2^t(x_1^t) = \bar{x}_2^t$ *for* $x_1^t \in [\bar{x}_1^t - K_0 - K_1,\ \bar{x}_1^t]$ *and* $\gamma_2^t(x_1^t)$ *strictly decreasing with respect to* x_1^t *for* $x_1^t \in \Re \setminus [\bar{x}_1^t - K_0 - K_1,\ \bar{x}_1^t]$.

that further partitions Region A into the following eight subregions:

$$A_{(0,0)} := \{(x_1^t, x_2^t) : \bar{x}_i^t - K_0 - K_i \leq x_i^t < \bar{x}_i^t \ \ \forall i = 1,2 \ \textit{and}$$
$$x_1^t + x_2^t > \bar{x}_1^t + \bar{x}_2^t - K_0 - K_1 - K_2\}$$

$$A_{(0,j)} := \left\{ (x_1^t, x_2^t) : \begin{array}{lr} \gamma_1^t(x_2^t) - K_0 - K_1 \leq x_1^t < \gamma_1^t(x_2^t) \ \ \textit{and} \ \ x_2^t \geq \bar{x}_2^t & \textit{if} \quad j{=}1 \\ \gamma_1^t(x_2^t) - K_1 \leq x_1^t < \gamma_1^t(x_2^t) \ \ \textit{and} \ \ x_2^t < \bar{x}_2^t & \textit{if} \quad j{=}{-}1 \end{array} \right\}$$

$$A_{(1,j)} := \left\{ (x_1^t, x_2^t) : \begin{array}{lll} x_1^t \geq \bar{x}_1^t \;\; and \;\; \gamma_2^t(x_1^t) - K_0 - K_2 \leq x_2^t < \gamma_2^t(x_1^t) & if & j{=}0 \\ x_1^t \geq \gamma_1^t(x_2^t) \;\; and \;\; x_2^t \geq \gamma_2^t(x_1^t) & if & j = 1 \\ x_1^t \geq \gamma_1^t(x_2^t) \;\; and \;\; x_2^t < \gamma_2^t(x_1^t) - K_0 - K_2 & if & j{=}{-1} \end{array} \right\}$$

$$A_{(-1,j)} := \left\{ (x_1^t, x_2^t) : \begin{array}{lll} x_1^t \leq \bar{x}_1^t - K_0 - K_1 \;\; and & & \\ \qquad\qquad \gamma_2^t(x_1^t) - K_2 \leq x_2^t < \gamma_2^t(x_1^t) & if & j{=}0 \\ x_1^t \leq \gamma_1^t(x_2^t) - K_0 - K_1 \;\; and \;\; x_2^t \geq \gamma_2^t(x_1^t) & if & j{=}1 \end{array} \right\}$$

Proof: The subscripts of A reflect the sign of the Lagrange variables and imply which, if any, of the constraints are binding. As an example, consider the region defined by $A_{\{k_1,k_2\}}$. The index $k_i = 1$ if $\lambda_i^t > 0$, $k_i = 0$ if $\lambda_i^t = 0$, and $k_i = -1$ if $\lambda_i^t < 0$. We prove the results associated with regions $A_{\{0,0\}}, A_{\{0,1\}}, A_{\{0,-1\}}, A_{\{1,1\}}$, and $A_{\{1,-1\}}$. The analysis for regions $A_{\{1,0\}}, A_{\{-1,0\}}$, and $A_{\{-1,1\}}$ are symmetric to the ones for regions $A_{\{0,1\}}, A_{\{0,-1\}}$, and $A_{\{1,-1\}}$, respectively. An illustration of the state space segmentation is provided in Figure II.8(a).

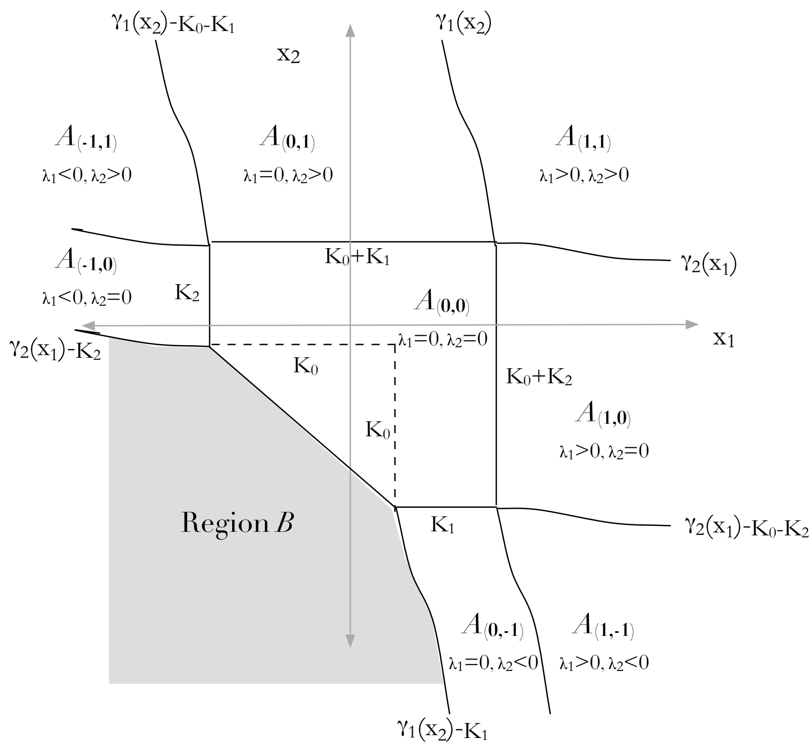
We first consider region $A_{\{0,0\}}$ that corresponds to $\lambda_1^t = \lambda_2^t = 0$. Following (2.26a) - (2.26c), in this region we have

$$x_1^t < \phi_1^t(0,0,0) - a_{11}p_{1L}^t - a_{12}p_{2L}^t + b_1 < x_1^t + K_0 + K_1 \tag{2.28a}$$

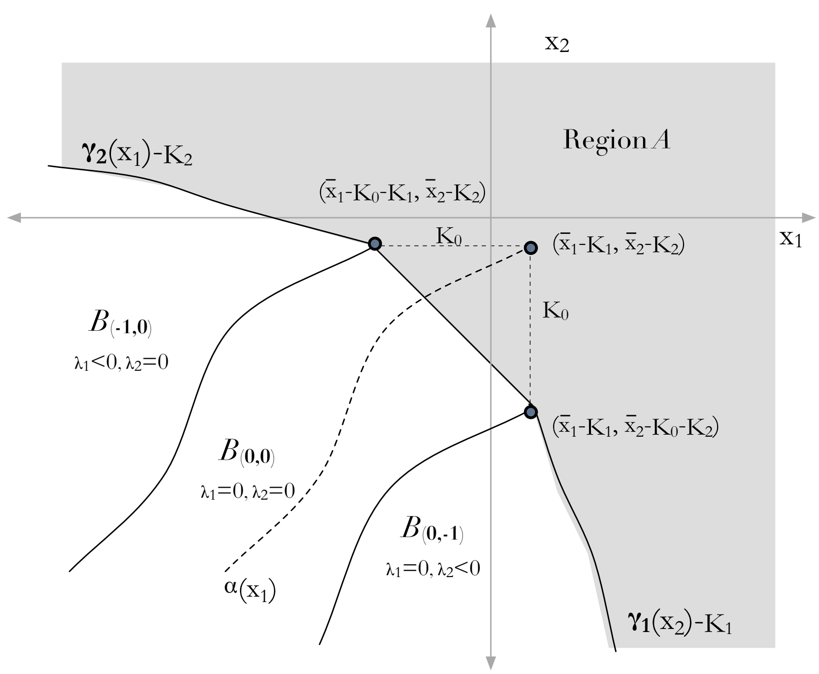
$$x_2^t < \phi_2^t(0,0,0) - a_{21}p_{1L}^t - a_{22}p_{2L}^t + b_2 < x_2^t + K_0 + K_2 \tag{2.28b}$$

$$\begin{aligned} &\phi_1^t(0,0,0) + \phi_2^t(0,0,0) - (a_{11} + a_{21})p_{1L}^t - (a_{12} + a22)p_{2L}^t + b_1 + b_2 \\ &\qquad < x_1^t + x_2^t + K_0 + K_1 + K_2 \end{aligned} \tag{2.28c}$$

Thus, by substituting (2.27a) and (2.27b) into (2.28a)-(2.28c), we can define this region as $\{(x_1^t, x_2^t) : \bar{x}_i^t - K_0 - K_i \leq x_i^t < \bar{x}_i^t \;\; \forall i = 1,2 \text{ and } x_1^t + x_2^t > \bar{x}_1^t + \bar{x}_2^t - K_0 - K_1 - K_2\}$.



(a) Partitioning of region A



(b) Partitioning of region B

Figure 2.8: Segmentation of the state space

We now consider region $A_{\{0,1\}}$ that corresponds to $\lambda_1^t = 0$ and $\lambda_2^t > 0$. Since $\lambda_2^t > 0$, after substituting in (2.22), (2.27a) and (2.27b), constraints (2.26a) - (2.26c) reduce to the following:

$$x_1^t < \phi_1^t(0, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{12}}{2}\lambda_2^t < x_1^t + K_0 + K_1 \tag{2.29a}$$

$$x_2^t = \phi_2^t(0, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0,0,0) + \frac{a_{22}}{2}\lambda_2^t \qquad \text{(equality due to } \lambda_2^t > 0\text{)} \tag{2.29b}$$

We first consider (2.29b) which defines one boundary for this region.

$$\begin{aligned} x_2^t &= \phi_2^t(0, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0,0,0) + \frac{a_{22}}{2}\lambda_2^t \\ &> \bar{x}_2^t \quad \text{(since } \phi_2^t \uparrow \lambda_2^t \text{ by Lemma B.2.3 and } a_{22} > 0\text{)} \end{aligned}$$

and $\lim_{\lambda_2^t \to 0} x_2^t = \phi_2^t(0,0,0) + \bar{x}_2^t - \phi_2^t(0,0,0) = \bar{x}_2^t$.

Furthermore, as $\phi_2^t \uparrow \lambda_2^t$ by Lemma B.2.3 and $a_{22} > 0$, x_2^t is strictly increasing with respect to λ_2^t in this region (equivalently, λ_2^t to be strictly increasing with respect to x_2^t), there is a one-to-one function defining λ_2^t in terms of x_2^t, that is $\lambda_2^t = \lambda_2^t(x_2^t)$.

The remaining boundaries are given by the inequalities in (2.29a). Since $\lambda_1^t = 0$, the constraints are not binding. By rearranging the terms, we get

$$\phi_1^t(0, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{12}}{2}\lambda_2^t - K_0 - K_1 < x_1^t < \phi_1^t(0, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{12}}{2}\lambda_2^t \tag{2.31}$$

Temporarily defining a function $\delta_1(\lambda_2^t) := \phi_1^t(0, \lambda_2^t, 0) - \phi_1^t(0,0,0) + \frac{a_{12}}{2}\lambda_2^t$, we can rewrite (2.31) as

$$\delta_1^t(\lambda_2^t) + \bar{x}_1^t - K_0 - K_1 < x_1^t < \delta_1^t(\lambda_2^t) + \bar{x}_1^t \tag{2.32}$$

Lemma B.2.3 and $a_{12} < 0$ yields $\delta_1^t(\lambda_2^t) < 0$ and that $\delta_1^t(\lambda_2^t)$ is strictly decreasing with respect to λ_2^t. If we now define $\gamma_1^t(x_2^t) := \bar{x}_1^t + \delta_1^t\left(\lambda_2^t(x_2^t)\right)$, we can write the boundaries for this region as

$$\gamma_1^t(x_2^t) - K_0 - K_1 < x_1^t < \gamma_1^t(x_2^t) \tag{2.33}$$

The fact that $\gamma_1^t(x_2^t)$ strictly decreasing with respect to x_2^t follows immediately from $\delta_1^t(\lambda_2^t)$ strictly decreasing with respect to λ_2^t and λ_2^t strictly increasing with respect to x_2^t.

Next, the region denoted by $A_{\{0,-1\}}$ corresponds to $\lambda_1^t = 0$ and $\lambda_2^t < 0$. In this region constraints (2.26a) - (2.26c) reduce to the following:

$$x_1^t < \phi_1^t(0, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{12}}{2}\lambda_2^t < x_1^t + K_1 \tag{2.34a}$$

$$x_2^t = \phi_2^t(0, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0,0,0) + \frac{a_{22}}{2}\lambda_2^t + K_0 + K_2 \tag{2.34b}$$

The analysis of this region is very similar to the analysis of Region $A_{\{0,1\}}$. The three differences are (i) the right-hand side of inequality (2.34a) includes the term K_1 instead of $K_0 + K_1$, (ii) the right-hand side of equation (2.34b) has the additional terms $K_0 + K_2$ and (iii) $\lambda_2^t < 0$ throughout the region. Similar steps in the proof for Region $A_{\{0,1\}}$ yields the boundaries $x_2^t < \bar{x}_2^t - K_0 - K_2$ and $\gamma_1^t(x_2^t) - K_1 < x_1^t < \gamma_1^t(x_2^t)$ and ensures $\gamma_1^t(x_2^t)$ strictly decreasing with respect to x_2^t.

We next consider region $A_{\{1,1\}}$ that corresponds to $\lambda_1^t > 0$ and $\lambda_2^t > 0$. Since, in this region, both $\lambda_1^t > 0$, and $\lambda_1^t > 0$, constraints (2.26a) - (2.26c) reduce to the following equalities.

$$x_1^t = \phi_1^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{11}}{2}\lambda_1^t + \frac{a_{12}}{2}\lambda_2^t \tag{2.35a}$$

$$x_2^t = \phi_2^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0,0,0) + \frac{a_{21}}{2}\lambda_1^t + \frac{a_{22}}{2}\lambda_2^t \tag{2.35b}$$

Substituting in the previously defined function $\gamma_1^t(x_2^t)$, equation (2.35a) defines the first boundary of this region as:

$$\begin{aligned} x_1^t &= \phi_1^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{11}}{2}\lambda_1^t + \frac{a_{12}}{2}\lambda_2^t \\ &= \phi_1^t(\lambda_1^t, \lambda_2^t, 0) - \phi_1^t(0, \lambda_2^t, 0) + \frac{a_{11}}{2}\lambda_1^t + \gamma_1^t(x_2^t) \\ &> \gamma_1^t(x_2^t) \qquad (\text{since } \phi_1^t \uparrow \lambda_1^t \text{ and } a_{11} > 0,\ \lambda_1^t > 0) \end{aligned} \tag{2.36}$$

and $\lim_{\lambda_2^t \to 0} x_1^t = \phi_1^t(\lambda_1^t, 0, 0) - \phi_1^t(0, \lambda_2^t, 0) + \gamma_1^t(x_2^t) = \gamma_1^t(x_2^t)$, the left-hand-side boundary for Region $A_{\{0,1\}}$. Likewise, the remaining boundary is given by $x_2^t > \gamma_2^t(x_1^t)$ where $\gamma_2^t(x_1^t)$ is defined similar to $\gamma_1^t(x_2^t)$ in the analysis for Region $A_{\{0,1\}}$.

Lastly, we examine the region denoted by $A_{\{1,-1\}}$ where $\lambda_1^t > 0$ and $\lambda_2^t < 0$. Constraints (2.26a) - (2.26c) give the following equalities.

$$x_1^t = \phi_1^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0, 0, 0) + \frac{a_{11}}{2}\lambda_1^t + \frac{a_{12}}{2}\lambda_2^t \tag{2.37a}$$

$$x_2^t = \phi_2^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0, 0, 0) + \frac{a_{21}}{2}\lambda_1^t + \frac{a_{22}}{2}\lambda_2^t - K_0 - K_2 \tag{2.37b}$$

As in (2.36), one boundary is given by $x_1^t \geq \gamma_1^t(x_2^t)$. The remaining boundary is given by

$$\begin{aligned} x_2^t &= \phi_2^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0, 0, 0) + \frac{a_{11}}{2}\lambda_1^t + \frac{a_{22}}{2}\lambda_2^t - K_0 - K_2 \\ &= \phi_2^t(\lambda_1^t, \lambda_2^t, 0) - \phi_2^t(\lambda_1^t, 0, 0) + \gamma_2^t(x_1^t) + \frac{a_{22}}{2}\lambda_2^t - K_0 - K_2 \\ &< \gamma_2^t(x_1^t) - K_0 - K_2 \qquad (\text{since } \lambda_2^t < 0,\ \phi_2^t \uparrow \lambda_2^t \text{ and } a_{22} > 0) \end{aligned}$$

and $\lim_{\lambda_2^t \to 0} x_2^t = \phi_2^t(\lambda_1^t, 0, 0) - \phi_2^t(\lambda_1^t, 0, 0) + \gamma_2^t(x_1^t) - K_0 - K_2 = \gamma_2^t(x_1^t) - K_0 - K_2$, the lower boundary for Region $A_{\{1,0\}}$. □

Lemma 2.5. *Together with $\gamma_1^t(x_2^t)$ and $\gamma_2^t(x_1^t)$, a monotone function $\alpha^t(x_1^t) : [-\infty, \bar{x}_1^t - K_1] \to [-\infty, \bar{x}_2^t - K_2]$ with $\alpha^t(\bar{x}_1^t - K_1) = \bar{x}_2^t - K_2$ and $\alpha^t(x_1^t)$ strictly increasing with respect to x_1^t divides Region B into the three subregions:*

- $B_{(0,-1)} := B'_{(0,-1)} \cup B''_{(0,-1)}$ *where*

$$B'_{(0,-1)} := \{(x_1^t, x_2^t) : \bar{x}_1^t - K_1 < x_1^t \leq \gamma_1^t(x_2^t) - K_1 \text{ and } x_2^t \leq \bar{x}_2^t - K_0 - K_2\}$$

$$B''_{(0,-1)} := \{(x_1^t, x_2^t) : x_1^t \leq \bar{x}_1^t - K_1 \text{ and } x_2^t \leq \alpha^t(x_1^t) - K_0\}$$

- $B_{(-1,0)} := B'_{(-1,0)} \cup B''_{(-1,0)}$ *where*

$$B'_{(-1,0)} := \{(x_1^t, x_2^t) : x_1^t \leq \bar{x}_1^t - K_0 - K_1 \text{ and } \bar{x}_2^t - K_2 \leq x_2^t \leq \gamma_2^t(x_1^t) - K_2\}$$

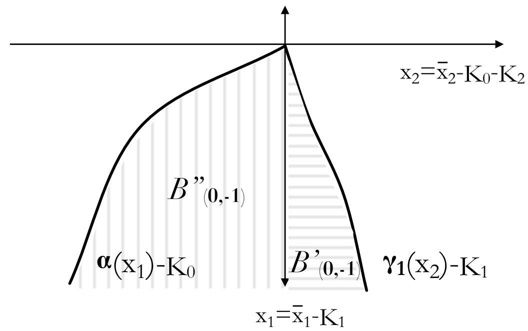


Figure 2.9: Subregions for the proof of Lemma 2.5

$$B''_{(-1,0)} := \{(x_1^t, x_2^t) : x_1^t \leq \bar{x}_1^t - K_0 - K_1 \ and \ \ \alpha^t(x_1^t + K_0) < x_2^t \leq \bar{x}_2^t - K_2\}$$

- $B_{(0,0)} := \Re^2 \backslash \{A \cup \big(B_{(0,-1)} \cup B_{(-1,0)}\big)\}$.

Proof: We only provide the proof for region $B_{\{0,-1\}}$ as the analysis of $B_{\{-1,0\}}$ is similar and region $B_{\{0,0\}}$ is defined by the remaining area in Region B. Region $B_{\{0,-1\}}$ corresponds to $\lambda_2^t < 0$ and $\mu^t > 0$ for which constraints (2.26a) - (2.26c) reduce to

$$x_1^t = \phi_1^t(0, \lambda_2^t, \mu^t) - \phi_1^t(0,0,0) + \bar{x}_1^t + \frac{a_{12}}{2}\lambda_2^t - \frac{(a_{11}+a_{12})}{2}\mu^t - K_1 \tag{2.38a}$$

$$x_2^t = \phi_2^t(0, \lambda_2^t, \mu^t) - \phi_2^t(0,0,0) + \bar{x}_2^t + \frac{a_{22}}{2}\lambda_2^t - \frac{(a_{21}+a_{22})}{2}\mu^t - K_0 - K_2 \tag{2.38b}$$

The analysis of this region is simpler if we consider the cases where $x_1^t > \bar{x}_1^t - K_1$ and $x_1^t \leq \bar{x}_1^t - K_1$ separately corresponding to subregions $B'_{(0,-1)}$ and $B''_{(0,-1)}$, respectively. For subregion $B'_{(0,-1)}$, we first find the feasible values for x_2^t and then show that $\gamma_1^t(x_2^t) - K_1$ defines the possible values for x_1^t. For subregion $B''_{(0,-1)}$, we find a function $\alpha^t(x_1^t)$ that is defined on the domain $x_1^t \leq \bar{x}_1^t - K_1$ which establishes the boundaries for the subregion. Figure 2.9 illustrates the subregions $B'_{(0,-1)}$ and $B''_{(0,-1)}$. We first show that in the subregion $B'_{\{0,-1\}}$, we have $x_2^t \leq \bar{x}_2^t - K_0 - K_2$. For arbitrary $\lambda_2^t < 0$ and $\mu^t > 0$, by (2.38b), we have,

$$
\begin{aligned}
x_2^t &= \phi_2^t(0,\lambda_2^t,\mu^t) - \phi_2^t(0,0,0) + \frac{a_{22}}{2}\lambda_2^t - \frac{(a_{21}+a_{22})}{2}\mu^t + \bar{x}_2^t - K_0 - K_2 \\
&< \phi_2^t(0,\lambda_2^t,\mu^t) - \phi_2^t(0,\lambda_2^t,0) + \bar{x}_2^t - K_0 - K_2 \\
&\qquad\qquad (\text{since } \lambda_2^t < 0,\ \mu^t > 0, \text{ and } a_{22} > 0,\ a_{11}+a_{12} > 0) \\
&< \bar{x}_2^t - K_0 - K_2 \qquad (\text{since } \phi_2^t \uparrow \lambda_2^t,\ \downarrow \mu^t \text{ and } \lambda_2^t < 0, \mu^t > 0)
\end{aligned}
$$

and $\lim_{\lambda_2^t,\mu^t\to 0} x_2^t = \phi_2^t(0,0,0) - \phi_2^t(0,0,0) + \bar{x}_2^t - K_0 - K_2 = \bar{x}_2^t - K_0 - K_2$.

Examining the expression for x_1^t given in (2.38a), we get

$$
\begin{aligned}
x_1^t &= \phi_1^t(0,\lambda_2^t,\mu^t) - \phi_1^t(0,0,0) + \bar{x}_1^t + \frac{a_{12}}{2}\lambda_2^t - \frac{(a_{11}+a_{12})}{2}\mu^t - K_1 \\
&= \phi_1^t(0,\lambda_2^t,\mu^t) - \phi_1^t(0,\lambda_2^t,0) - \frac{(a_{11}+a_{12})}{2}\mu^t + \gamma_1^t(x_2^t) - K_1 \\
&< \gamma_1^t(x_2^t) - K_1 \qquad (\text{since } \mu^t > 0,\ \phi_1^t \downarrow \mu^t \text{ and } a_{11}+a_{12} > 0)
\end{aligned}
$$

and $\lim_{\mu^t\to 0} x_1^t = \phi_1^t(0,\lambda_2^t,0) - \phi_1^t(0,0,0) + \bar{x}_1^t + \frac{a_{12}}{2}\lambda_2^t - K_1 = \gamma_1^t(x_2^t) - K_1$, the left-hand-side boundary for Region $A_{\{0,-1\}}$. We note that the increasing property of $\gamma_1^t(x_2^t)$ established in the proof of Lemma 2.4 ensures that $\bar{x}_1^t - K_1 \le \gamma_1^t(x_2^t) - K_1$. Thus the expressions $\bar{x}_1^t - K_1 \le x_1^t \le \gamma_1^t(x_2^t) - K_1$ and $x_2^t \le \bar{x}_2^t - K_0 - K_2$ defines the states corresponding to $B'_{\{0,-1\}}$.

For subregion $B''_{\{0,-1\}}$, we first note that $\lim_{\lambda_2^t\to 0} x_1^t$ defines the boundary between regions $B''_{\{0,-1\}}$ and $B_{\{0,0\}}$. Along this boundary, by (2.38a), we have $x_1^t = \phi_1^t(0,0,\mu^t) - \phi_1^t(0,0,0) + \bar{x}_1^t - \frac{(a_{11}+a_{12})}{2}\mu^t - K_1$. Using Lemma 2.3, we find that x_1^t is strictly decreasing with respect to μ^t. Hence x_1^t falls solely in this subregion $B''_{\{0,-1\}}$ ($x_1^t \le \bar{x}_1^t - K_1$). Further x_1^t strictly decreasing with respect to μ^t implies, there is a one-to-one function defining x_1^t and μ^t, i.e., $\mu^t(x_1^t)$ where μ^t is strictly decreasing with respect to x_1^t.

Considering (2.38b) along the boundary, we have $x_2^t = \phi_2^t(0,0,\mu^t) - \phi_2^t(0,0,0) + \bar{x}_2^t - \frac{(a_{21}+a_{22})}{2}\mu^t - K_0 - K_2$. Let us temporarily define $\sigma_2(\mu^t) = \phi_2^t(0,0,\mu^t) - \phi_2^t(0,0,0) +$

$\bar{x}_2^t - \frac{(a_{21}+a_{22})}{2}\mu^t$. Then, by Lemma 2.3, $\sigma_2(\mu^t)$ is strictly decreasing with respect to μ^t. Consequently, as $\mu^t(x_1^t)$ is strictly decreasing with respect to x_1^t, we have $\sigma_2\left(\mu^t(x_1^t)\right)$ strictly increasing with respect to x_1^t. We now introduce and define $\alpha(x_1^t) := \sigma_2\left(\mu^t(x_1^t)\right) - K_2$, thus the function $x_2^t = \alpha(x_1^t) - K_0$ forms the boundary between regions $B''_{\{0,-1\}}$ and $B_{\{0,0\}}$. Since $\sigma_2\left(\mu^t(x_1^t)\right) \uparrow x_1^t$, we have $\alpha(x_1^t) \uparrow x_1^t$. Lastly, by (2.38a), $x_1^t = \bar{x}_1^t - K_1$ implies $\phi_1^t(0,0,\mu^t) - \phi_1^t(0,0,0) - \frac{(a_{21}+a_{22})}{2}\mu^t = 0$ for which the only solution is $\mu^t = 0$. (Note $\phi_1^t \downarrow \mu^t$ and $\mu^t \geq 0$). Hence by (2.38b), we have $x_2^t = \bar{x}_2^t - K_0 - K_2$, which also yields $\alpha(\bar{x}_1^t - K_1) = \bar{x}_2^t - K_2$. $\square$

To complete the Proof of Theorem 2.1, we note that part 1(a) follows directly from the definitions of the monotone functions $\gamma_1^t(x_2^t)$ and $\gamma_2^t(x_1^t)$ in Lemma 2.4 and the complementary slackness conditions. For example, in region $A_{\{-1,1\}}$, the binding constraints yield $y_1^t = x_1^t + K_0 + K_1$ and $y_2^t = x_2^t$. In region $A_{\{0,1\}}$, we have $y_1^t = z_1^t + \bar{d}_1^t = \phi_1^t(0,\lambda_2^t,0) + b_1 - a_{11}p_1^t - a_{12}p_2^t = \phi_1^t(0,\lambda_2^t,0) - \phi_1^t(0,0,0) + \phi_1^t(0,0,0) + b_1 - a_{11}p_{1L}^t - a_{12}p_{2L}^t + \frac{a_{12}}{2}\lambda_2^t = \phi_1^t(0,\lambda_2^t,0) - \phi_1^t(0,0,0) + \bar{x}_1^t + \frac{a_{12}}{2}\lambda_2^t = \gamma_1^t(x_2^t)$.

For part 1(b), in Regions $A_{\{0,j\}}$, the optimal order-up-to level for product 1 is independent of its own starting inventory x_1^t and by Lemma 2.4 and part (a), it is non-increasing with x_2^t. Specifically, in Region $A_{\{0,0\}}$, it is independent of x_2^t and in regions $A_{\{0,1\}}$ and $A_{\{0,-1\}}$, it is strictly decreasing with the inventory position of x_2^t. In Regions $A_{\{1,j\}}$, by part (a), we have $y_1^t = x_1^t$, hence the order-up-to level of product 1 is increasing with x_1^t and independent of x_2^t. For regions $A_{\{-1,0\}}$ and $A_{\{-1,1\}}$, again by part (a), we have $y_1^t = x_1^t + K_0 + K_1$, thus the order-up-to level of product 1 is increasing with x_1^t and independent of x_2^t. Symmetric arguments hold for product 2.

The proofs of part 2 (a) and (b) are due to Lemma 2.5. Suppose $l^t(x_1^t, x_2^t)$ denotes the optimal amount of flexible capacity allocated to product 1. Since in Region B, the complementary slackness conditions imply full utilization of each resource,

$K_0 - l^t(x_1^t, x_2^t)$ will be the amount of flexible capacity allocated to product 2. After the allocation of the flexible resource and employing the dedicated resources, the optimal production policy brings inventories of products 1 and 2 to $x_1^t + K_1 + l^t(x_1^t, x_2^t)$ and $x_2^t + K_2 + K_0 - l^t(x_1^t, x_2^t)$, respectively. Specifically, in region $B_{\{-1,0\}}$, complementary slackness yields $y_1^t - x_1^t = K_0 + K_1$ and $y_2^t - x_2^t = K_2$, thus $l^t(x_1^t, x_2^t) = K_0$. Similarly, in region $B_{\{0,-1\}}$, complementary slackness conditions give $y_1^t - x_1^t = K_1$ and $y_2^t - x_2^t = K_0 + K_2$, hence $l^t(x_1^t, x_2^t) = 0$. For region $B_{\{0,0\}}$, the definition of $l^t(x_1^t, x_2^t)$ yields $y_2^t - x_2^t = K_0 + K_2 - l^t(x_1^t, x_2^t)$ which is equivalent to $\phi_2^t(0,0,\mu^t) - \phi_2^t(0,0,0) + \bar{x}_2^t - \frac{(a_{21}+a_{22})}{2}\mu^t - x_2^t = K_0 + K_2 - l^t(x_1^t, x_2^t) = \alpha^t(x_1^t + l^t(x_1^t, x_2^t)) = K_0 + K_2 - l^t(x_1^t, x_2^t)$. Therefore, the optimal production policy satisfies $x_2^t + K_0 - l^t(x_1^t, x_2^t) = \alpha^t(x_1^t + l^t(x_1^t, x_2^t))$. Furthermore, the complementary slackness condition (2.26c) yields μ^t to be a function of x_1^t and x_2^t only through their sum $x_1^t + x_2^t$. Thus, the optimal modified base stock levels for products 1 and 2 are identical for starting inventory positions for which the total inventory level, $x_1^t + x_2^t$, is identical.

For part 2(c), first let $\alpha'^t(x_1^t + l^t)$ denote the derivative of $\alpha^t(x_1^t + l^t)$ with respect to its argument. By Lemma 2.5, α^t is increasing, thus $\alpha'^t(x_1^t + l^t) > 0$. Next, differentiating both sides of $l^t(x_1^t, x_2^t) + \alpha^t(x_1^t + l^t(x_1^t, x_2^t)) = x_2^t + K_0$ with respect to x_1^t, we get

$$\frac{\partial l^t}{\partial x_1^t} = -\frac{\alpha'^t(x_1^t + l^t)}{1 + \alpha'^t(x_1^t + l^t)} < 0 \tag{2.39}$$

Thus, l^t is decreasing with respect to x_1^t. Similarly, differentiating both sides of $l^t(x_1^t, x_2^t) + \alpha^t(x_1^t + l^t(x_1^t, x_2^t)) = x_2^t + K_0$ with respect to x_2^t, we get

$$\frac{\partial l^t}{\partial x_2^t} = \frac{1}{1 + \alpha'^t(x_1^t + l^t)} > 0 \tag{2.40}$$

Hence, l^t is increasing with respect to x_2^t. Finally, for part 2(d), the order-up-to level for product 1 is $x_1^t + l^t(x_1^t, x_2^t) + K_1$. Differentiating it with respect to x_1^t and with respect to x_2^t and using (2.39) and (2.40) we get

$$\frac{\partial x_1^t + l^t(x_1^t, x_2^t) + K_1}{\partial x_1^t} = \frac{1}{1 + \alpha'^t(x_1^t + l^t)} > 0$$
$$\frac{\partial x_1^t + l^t(x_1^t, x_2^t) + K_1}{\partial x_2^t} = \frac{1}{1 + \alpha'^t(x_1^t + l^t)} > 0$$

The order-up-to level for product 2 is $x_2^t + K_0 - l^t(x_1^t, x_2^t) + K_2$. Again, differentiating it with respect to x_1^t and with respect to x_2^t, we get

$$\frac{\partial x_2^t + K_0 - l^t(x_1^t, x_2^t) + K_2}{\partial x_1^t} = \frac{\alpha'^t(x_1^t + l^t)}{1 + \alpha'^t(x_1^t + l^t)} > 0$$
$$\frac{\partial x_2^t + K_0 - l^t(x_1^t, x_2^t) + K_2}{\partial x_2^t} = \frac{\alpha'^t(x_1^t + l^t)}{1 + \alpha'^t(x_1^t + l^t)} > 0$$

Hence, the order-up-to level for both products is increasing with respect to either starting inventory position x_i^t. $\square$

Proof of Theorem 2.2 (Optimal Pricing Policy):

The proof of part (a) follows from the expression given in (2.22) and (2.23). In region A, $\mu^t = 0$, therefore combining (2.22) and (2.4) we have

$$m_1^t(x_1^t, x_2^t) = -\frac{1}{2}\lambda_1^t(x_1^t, x_2^t) \tag{2.41a}$$
$$m_2^t(x_1^t, x_2^t) = -\frac{1}{2}\lambda_2^t(x_1^t, x_2^t) \tag{2.41b}$$

Thus, we have $m_1^t(x_1^t, x_2^t) > 0$ for $\lambda_1^t < 0$, $m_1^t(x_1^t, x_2^t) = 0$ for $\lambda_1^t = 0$, and $m_1^t(x_1^t, x_2^t) < 0$ for $\lambda_1^t > 0$. Following the state space segmentation set forth in Lemma B.2.4, $\lambda_1^t < 0$, $\lambda_1^t = 0$, and $\lambda_1^t > 0$ correspond to item 1 being critically understocked, moderately understocked, and overstocked, respectively. Similar arguments yield the results corresponding to $m_2^t(x_1^t, x_2^t)$. The expressions for current period list prices are given by (2.23). In Region B, the expressions given in (2.22) yields

$$m_1^t(x_1^t, x_2^t) = -\frac{1}{2}\left(\lambda_1^t(x_1^t, x_2^t) - \mu^t(x_1^t, x_2^t)\right) \tag{2.42a}$$
$$m_2^t(x_1^t, x_2^t) = -\frac{1}{2}\left(\lambda_2^t(x_1^t, x_2^t) - \mu^t(x_1^t, x_2^t)\right) \tag{2.42b}$$

Since region B is defined as the states corresponding to $\mu^t > 0$ and non-positive λ_1^t and λ_2^t, we have $m_1^t(x_1^t, x_2^t) > 0$ and $m_2^t(x_1^t, x_2^t) > 0$. For part (b), in the states that correspond to $B_{\{0,0\}}$, (2.42a) and (2.42b) reduce to

$$m_1^t(x_1^t, x_2^t) = \frac{1}{2}\mu^t(x_1^t, x_2^t) \tag{2.43a}$$

$$m_2^t(x_1^t, x_2^t) = \frac{1}{2}\mu^t(x_1^t, x_2^t) \tag{2.43b}$$

Therefore, we have $m_1^t(x_1^t, x_2^t) = m_2^t(x_1^t, x_2^t)$. Further, (2.4) then yields

$$p_2^t(x_1^t, x_2^t) = p_1^t(x_1^t, x_2^t) + C^t \tag{2.44a}$$

where $C^t = p_{2L}^t - p_{1L}^t$. The fact that $m_i^t(x_1^t, x_2^t)$ is a function of x_1^t and x_2^t through their sum follows from (2.26c) which for this region implies that μ^t is a function of $x_1^t + x_2^t$. For part (c), we only show the proof for product 1, as similar arguments yield the desired monotonicity results for product 2. In regions $A_{\{0,0\}}$, $A_{\{0,1\}}$, and $A_{\{0,-1\}}$, we have $p_1^t = p_{1L}^t$ and hence p_1^t is independent of both x_1^t and x_2^t. In region $A_{\{1,0\}}$, we have $p_1^t = p_{1L}^t - \frac{1}{2}\lambda_1^t$. Based on Lemma 2.4, in this region λ_1^t increases with x_1^t and is independent of x_2^t, hence p_1^t decreases with x_1^t and is independent of x_2^t. With a similar analysis, we also find that p_1^t decreases with x_1^t and is independent of x_2^t in region $A_{\{-1,0\}}$ as well.

In region $A_{\{1,1\}}$, we have

$$x_1^t = \phi_1^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_1^t - \phi_1^t(0,0,0) + \frac{a_{11}}{2}\lambda_1^t + \frac{a_{12}}{2}\lambda_2^t \tag{2.45a}$$

$$x_2^t = \phi_2^t(\lambda_1^t, \lambda_2^t, 0) + \bar{x}_2^t - \phi_2^t(0,0,0) + \frac{a_{21}}{2}\lambda_1^t + \frac{a_{22}}{2}\lambda_2^t \tag{2.45b}$$

By differentiating both sides of (2.45a) and (2.45b) with respect to x_1^t, we find that both λ_1^t and λ_2^t are increasing with respect to x_1^t. Similarly, λ_1^t and λ_2^t are increasing with respect to x_2^t. Since, in this region p_1^t is given by $p_1^t = p_{1L}^t - \frac{1}{2}\lambda_1^t$, p_1^t decreases with respect to both x_1^t and x_2^t. Similar analysis yield λ_1^t to be independent of x_1^t and x_2^t in Regions $B_{\{0,-1\}}$ and $B_{\{0,0\}}$ and be increasing with respect to x_1^t and x_2^t in Region $B_{\{-1,0\}}$. Likewise, we find λ_2^t to be independent of x_1^t and x_2^t in Regions $B_{\{-1,0\}}$ and $B_{\{0,0\}}$ and be increasing with respect to x_1^t and x_2^t in Region $B_{\{0,-1\}}$. Lastly, we find μ^t to be increasing with respect to x_1^t and x_2^t in regions $B_{\{-1,0\}}, B_{\{0,-1\}}$ and $B_{\{0,0\}}$. Therefore, the desired monotonicity results follow immediately form the definitions of p_1^t in these regions. $\square$

2.9.3 Proofs of Sensitivity Results

Proof of Theorem 2.3

For sensitivity with respect to production cost, we only present the proof for instances where both items are critically understocked and share the flexible capacity or when one item is moderately understocked while the other is overstocked. The analysis of the other cases are similar and omitted for brevity. When both items are critically understocked and share the flexible capacity, we have $\lambda_1^{t*} = \lambda_2^{t*} = 0$ and $\mu^{t*} > 0$. The optimality conditions in this region yield to the following relationships that define the sensitivity of the optimal policy to a change in item 1's production cost, c_1.

$$
\begin{aligned}
&-2a_{11}\frac{\partial p_1^{t*}}{\partial c_1} - 2a_{12}\frac{\partial p_2^{t*}}{\partial c_1} + a_{11} + (a_{11}+a_{12})\frac{\partial \mu^{t*}}{\partial c_1} = 0\\
&-2a_{12}\frac{\partial p_1^{t*}}{\partial c_1} - 2a_{22}\frac{\partial p_2^{t*}}{\partial c_1} + a_{12} + (a_{12}+a_{22})\frac{\partial \mu^{t*}}{\partial c_1} = 0\\
&-1 + \beta V_{11}^{t-1}(z_1^{t*}-\epsilon_1^t, z_2^{t*}-\epsilon_2^t)\frac{\partial z_1^{t*}}{\partial c_1} + \beta V_{12}^{t-1}(z_1^{t*}-\epsilon_1^t, z_2^{t*}-\epsilon_2^t)\frac{\partial z_2^{t*}}{\partial c_1} - \frac{\partial \mu^{t*}}{\partial c_1} = 0\\
&\beta V_{12}^{t-1}(z_1^{t*}-\epsilon_1^t, z_2^{t*}-\epsilon_2^t)\frac{\partial z_1^{t*}}{\partial c_1} + \beta V_{22}^{t-1}(z_1^{t*}-\epsilon_1^t, z_2^{t*}-\epsilon_2^t)\frac{\partial z_2^{t*}}{\partial c_1} - \frac{\partial \mu^{t*}}{\partial c_1} = 0
\end{aligned}
$$

$$\frac{\partial z_1^{t*}}{\partial c_1} - \frac{\partial y_1^{t*}}{\partial c_1} - a_{11}\frac{\partial p_1^{t*}}{\partial c_1} - a_{12}\frac{\partial p_2^{t*}}{\partial c_1} = 0$$
$$\frac{\partial z_2^{t*}}{\partial c_1} - \frac{\partial y_1^{t*}}{\partial c_1} - a_{12}\frac{\partial p_1^{t*}}{\partial c_1} - a_{22}\frac{\partial p_2^{t*}}{\partial c_1} = 0$$
$$\frac{\partial y_1^{t*}}{\partial c_1} + \frac{\partial y_2^{t*}}{\partial c_1} = 0 \tag{2.46}$$

Solving the set of equations in (2.46) for $\frac{\partial p_1^{t*}}{\partial c_1}$, $\frac{\partial p_2^{t*}}{\partial c_1}$, $\frac{\partial y_1^{t*}}{\partial c_1}$ and $\frac{\partial y_2^{t*}}{\partial c_1}$,we find

$$\frac{\partial p_1^{t*}}{\partial c_1} = \frac{-2V_{11} + 2V_{12} - a_{12}\beta V_{12}^2 - a_{22}\beta V_{12}^2 + a_{12}\beta V_{11}V_{22} + a_{22}\beta V_{11}V_{22}}{-4V_{11} + 8V_{12} - 4V_{22} + 4a_{12}\beta V_{11}V_{22} + 2(a_{11} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.47a}$$
$$\frac{\partial p_2^{t*}}{\partial c_1} = \frac{-2V_{12} + 2V_{22} + a_{11}\beta V_{12}^2 + a_{12}\beta V_{12}^2 - a_{11}\beta V_{11}V_{22} - a_{12}\beta V_{11}V_{22}}{-4V_{11} + 8V_{12} - 4V_{22} + 4a_{12}\beta V_{11}V_{22} + 2(a_{11} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.47b}$$
$$\begin{aligned}\frac{\partial y_1^{t*}}{\partial c_1} &= -\left(\frac{\partial y_2^{t*}}{\partial c_1}\right)\\ &= -\frac{4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} - 2a_{22}\beta V_{22} + (a_{11}a_{22} - a_{12}^2)\beta^2(V_{11}V_{22} - V_{12}^2)}{\beta(-4V_{11} + 8V_{12} - 4V_{22} + 4a_{12}\beta V_{11}V_{22} + 2(a_{11} + a_{22})\beta(V_{11}V_{22} - V_{11}^2))}\end{aligned} \tag{2.47c}$$

The term in the denominators of (2.47a-c) is positive. (Recall $\Lambda' > 0$ in the proof of Lemma 2.2). The numerator of (2.47a) is nonnegative since $2V_{11} \leq 2V_{12}$, and $(a_{12}+a_{22})\beta(V_{11}V_{22}-V_{12}^2) \geq 0$ by $a_{12}+a_{22} > 0$ and $V_{11}V_{22}-V_{12}^2 \geq 0$. Hence, $\frac{\partial p_1^{t*}}{\partial c_1} \geq 0$. With a similar argument, we find $\frac{\partial p_2^{t*}}{\partial c_1} \leq 0$. The numerator of (2.47c) is also positive (with a similar reasoning in showing $\Lambda > 0$ in Case 2 of Lemma 2.2). Thus, we have $\frac{\partial y_1^{t*}}{\partial c_1} \leq 0$ and $\frac{\partial y_2^{t*}}{\partial c_1} \geq 0$.

Next, we consider the case where item 1 is moderately understocked and item 2 is overstocked. In this region, $\lambda_1^{t*} = 0$, $\lambda_2^{t*} > 0$, and $\mu = 0$, thus the modified base-stock level for product 2 is $y_2^{t*} = x_2^t$ and is independent of the production cost parameter. The optimality conditions result in the following relationships regarding the sensitivity with respect to c_1.

$$\frac{\partial p_1^{t*}}{\partial c_1} = \frac{1}{2} \tag{2.48a}$$

$$\frac{\partial p_2^{t*}}{\partial c_1} = -\frac{-2V_{12} + a_{12}\beta V_{12}^2 - a_{12}\beta V_{11}V_{22}}{2\left(2V_{11} + a_{22}\beta V_{12}^2 - a_{22}\beta V_{11}V_{22}\right)} \tag{2.48b}$$

$$\frac{\partial y_1^{t*}}{\partial c_1} = -\frac{\left(\begin{array}{l} 4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} + a_{12}^2\beta^2 V_{12}^2 - a_{11}a_{22}\beta^2 V_{12}^2 \\ \qquad\qquad -2a_{22}\beta V_{22} - a_{12}^2\beta^2 V_{11}V_{22} + a_{11}a_{22}\beta^2 V_{11}V_{22} \end{array}\right)}{2\beta(-2V_{11} - a_{22}\beta V_{12}^2 + a_{22}\beta V_{11}V_{22})} \tag{2.48c}$$

Hence $\frac{\partial p_1^{t*}}{\partial c_1} > 0$. It can be easily shown that the numerator of (2.48b) is nonnegative and the denominator of (2.48b) is negative. With the negative sign in front, we thus find $\frac{\partial p_2^{t*}}{\partial c_1} > 0$. Finally, the numerator of (2.48c) is positive (recall proof of case 2 in Lemma 2.2). The denominator is easily shown to be positive. Therefore, we have $\frac{\partial y_1^{t*}}{\partial c_1} < 0$.

The proof of part (b) regarding holding and backorder costs is similar and thus omitted for brevity. □

Proof of Theorem 2.4

We show the result for sensitivity regarding and increase in the dedicated capacity, for the two instances when item 1 is critically understocked, while item 2 is overstocked, and when both items are critically understocked sharing the flexible resource. The analysis of the remaining cases are similar.

When item 1 is critically understocked and item 2 is overstocked, we have, $\lambda_1^{t*} < 0$ while $\lambda_2^{t*} > 0$ indicating that $y_1^{t*} = x_1^t + K_0 + K_1$ and $y_2^{t*} = x_2^t$. Hence we immediately have $\frac{\partial y_1^{t*}}{\partial K_1} = 1 > 0$ and $\frac{\partial y_2^{t*}}{\partial K_1} = 0$. For sensitivity results for p_1^{t*} and p_2^{t*}, we again consider the optimality conditions to get

$$\frac{\partial p_1^{t*}}{\partial K_1} = \frac{2\beta V_{11} + a_{22}\beta^2(V_{12}^2 - V_{11}V_{22})}{\begin{pmatrix} 4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} + a_{12}^2\beta^2 V_{12}^2 - a_{11}a_{22}\beta^2 V_{12}^2 \\ -2a_{22}\beta V_{22} - a_{12}^2\beta^2 V_{11}V_{22} + a_{11}a_{22}\beta^2 V_{11}V_{22} \end{pmatrix}} \tag{2.49a}$$

$$\frac{\partial p_2^{t*}}{\partial K_1} = \frac{2\beta V_{12} - a_{12}\beta^2(V_{12}^2 - V_{11}V_{22})}{\begin{pmatrix} 4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} + a_{12}^2\beta^2 V_{12}^2 - a_{11}a_{22}\beta^2 V_{12}^2 \\ -2a_{22}\beta V_{22} - a_{12}^2\beta^2 V_{11}V_{22} + a_{11}a_{22}\beta^2 V_{11}V_{22} \end{pmatrix}} \tag{2.49b}$$

It is straightforward to check that the numerators in (2.49a) and (2.49b) are nonpositive and the denominators are positive. Thus, we find $\frac{\partial p_1^{t*}}{\partial K_1} \leq 0$ and $\frac{\partial p_2^{t*}}{\partial K_1} \leq 0$. When both items are critically understocked and receive a share of the flexible resource, the optimality conditions yield

$$\frac{\partial p_1^{t*}}{\partial K_1} = \frac{\beta(V_{12}^2 - V_{11}V_{22})}{-2V_{11} + 4V_{12} - 2V_{22} + (a_{11} + 2a_{12} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.50a}$$

$$\frac{\partial p_2^{t*}}{\partial K_1} = \frac{\beta(V_{12}^2 - V_{11}V_{22})}{-2V_{11} + 4V_{12} - 2V_{22} + (a_{11} + 2a_{12} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.50b}$$

$$\frac{\partial y_1^{t*}}{\partial K_1} = \frac{2(V_{12} - V_{22}) + (a_{11} + a_{12})\beta(V_{11}V_{22} - V_{12}^2)}{-2V_{11} + 4V_{12} - 2V_{22} + (a_{11} + 2a_{12} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.50c}$$

$$\frac{\partial y_2^{t*}}{\partial K_1} = \frac{2(V_{12} - V_{11}) + (a_{12} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)}{-2V_{11} + 4V_{12} - 2V_{22} + (a_{11} + 2a_{12} + a_{22})\beta(V_{11}V_{22} - V_{12}^2)} \tag{2.50d}$$

Similar to the previous results, we verify that $\frac{\partial p_1^{t*}}{\partial K_1} \leq 0$, $\frac{\partial p_2^{t*}}{\partial K_1} \leq 0$, $\frac{\partial y_1^{t*}}{\partial K_1} \geq 0$ and $\frac{\partial y_2^{t*}}{\partial K_1} \geq 0$.

The proof of part (b) is similar. $\square$

Proof of Theorem 2.5

We show the result for sensitivity regarding and increase in the demand intercept, for the two instances when both items are overstocked, and when both are critically understocked sharing the flexible resource. The analysis of the remaining cases are similar and therefore omitted for brevity. When both items are overstocked, we find,

$$\frac{\partial p_1^{t*}}{\partial b_1} = \frac{\left(\begin{array}{l} 2a_{22} + a_{12}^2 \beta V_{11} - 2a_{11}a_{22}\beta V_{11} - 2a_{12}a_{22}\beta V_{12} \\ \qquad -a_{22}^2 \beta V_{22} + a_{22}\beta^2 \left(V_{11}V_{22} - V_{12}^2\right)\left(a_{11}a_{22} - a_{12}^2\right) \end{array} \right)}{\left(\begin{array}{l} (a_{11}a_{22} - a_{12}^2)(4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} \\ \qquad -2a_{22}\beta V_{22} + (a_{11}a_{22} - a_{12}^2)\beta^2 (V_{11}V_{22} - V_{12}^2)) \end{array} \right)} \tag{2.51a}$$

$$\frac{\partial p_2^{t*}}{\partial b_1} = \frac{\left(\begin{array}{l} -2a_{12} + a_{11}a_{12}\beta V_{11} + 3a_{12}^2 \beta V_{12} - a_{11}a_{22}\beta V_{12} \\ \qquad +a_{12}a_{22}\beta V_{22} - a_{12}\beta^2 \left(V_{11}V_{22} - V_{12}^2\right)\left(a_{11}a_{22} - a_{12}^2\right) \end{array} \right)}{\left(\begin{array}{l} (a_{11}a_{22} - a_{12}^2)(4 - 2a_{11}\beta V_{11} - 4a_{12}\beta V_{12} \\ \qquad -2a_{22}\beta V_{22} + (a_{11}a_{22} - a_{12}^2)\beta^2 (V_{11}V_{22} - V_{12}^2)) \end{array} \right)} \tag{2.51b}$$

In equation (2.51a), both the numerator and denominator are positive. The denominator is positive since $a_{11}a_{22} - a_{12}^2 > 0$ and the remaining term in the paranthesis is also positive as discussed in the previous cases. The numerator is positive because $2a_{22} > 0$, and $a_{22}\beta^2 \left(V_{11}V_{22} - V_{12}^2\right)\left(a_{11}a_{22} - a_{12}^2\right) \geq 0$ and $a_{12}^2\beta V_{11} - 2a_{11}a_{22}\beta V_{11} - 2a_{12}a_{22}\beta V_{12} - a_{22}^2\beta V_{22} = -\left(a_{11}a_{22} - a_{12}^2\right)\beta V_{11} - a_{22}\beta\left(a_{11}V_{11} + 2a_{12}V_{12} + a_{22}V_{22}\right) \geq 0$ (nonpositivity of term in last paranthesis follows from similar arguments as in the proof of Lemma 2.2 case 2.) Therefore, $\frac{\partial p_1^{t*}}{\partial b_1} \geq 0$.

In equation (2.51b), both the numerator and denominator are also positive. The denominator is positive since $a_{11}a_{22} - a_{12}^2 > 0$ and the remaining term in the paranthesis is also positive as discussed previously. The numerator is positive because $-2a_{12} > 0$, and $-a_{12}\beta^2 \left(V_{11}V_{22} - V_{12}^2\right)\left(a_{11}a_{22} - a_{12}^2\right) \geq 0$ and $a_{11}a_{12}\beta V_{11} + 3a_{12}^2\beta V_{12} - a_{11}a_{22}\beta V_{12} + a_{12}a_{22}\beta V_{22} = -\left(a_{11}a_{22} - a_{12}^2\right)\beta V_{12} + a_{12}\beta\left(a_{11}V_{11} + 2a_{12}V_{12} + a_{22}V_{22}\right) \geq 0$ Therefore, we also have $\frac{\partial p_2^{t*}}{\partial b_1} \geq 0$.

When both items are critically understocked and receive a share of the flexible resource, the optimality conditions yield

$$\frac{\partial p_1^{t*}}{\partial b_1} = -\frac{\left(\begin{array}{l} 2a_{22}V_{11} - 4a_{22}V_{12} + 2a_{22}V_{22} \\ \qquad +(V_{11}V_{22} - V_{12}^2)(a_{12}^2 - a_{22}^2 - 2a_{22}(a_{11} + a_{12})) \end{array}\right)}{2(a_{11}a_{22} - a_{12}^2)\left(\begin{array}{l} (-2V_{11} + 4V_{12} - 2V_{22} \\ \qquad +\beta(a_{11} + 2a_{12} + a_{22})(V_{11}V_{22} - V_{12}^2)) \end{array}\right)} \tag{2.52a}$$

$$\begin{aligned} \frac{\partial p_2^{t*}}{\partial b_1} &= -\frac{a_{12}}{2(a_{11}a_{22} - a_{12}^2)} \\ &\quad + \frac{(-a_{12}(a_{11} + a_{12}) + a_{11}(a_{12} + a_{22}))\beta(V_{11}V_{22} - V_{12}^2)}{2(a_{11}a_{22} - a_{12}^2)\left(\begin{array}{l} -2V_{11} + 4V_{12} - 2V_{22} \\ \qquad +\beta(a_{11} + 2a_{12} + a_{22})(V_{11}V_{22} - V_{12}^2) \end{array}\right)} \end{aligned} \tag{2.52b}$$

$$\begin{aligned} \frac{\partial y_1^{t*}}{\partial b_1} &= -\left(\frac{\partial y_2^{t*}}{\partial b_1}\right) \\ &= -\frac{2V_{11} - 2V_{12} + a_{12}\beta V_{12}^2 + a_{22}\beta V_{12}^2 - a_{12}\beta V_{11}V_{22} - a_{22}\beta V_{11}V_{22}}{2(-2V_{11} + 4V_{12} - 2V_{22} + \beta(a_{11} + 2a_{12} + a_{22})(V_{11}V_{22} - V_{12}^2))} \end{aligned} \tag{2.52c}$$

Based on results from previous cases, the denomiator of (2.52a) is positive. The numerator is negative since $2a_{22}V_{11} - 4a_{22}V_{12} + 2a_{22}V_{22} < 0$ and $(V_{11}V_{22} - V_{12}^2)(a_{12}^2 - a_{22}^2 - 2a_{22}(a_{11} + a12))$ may be easily shown to be negative. Hence, with the negative sign in front, we find $\frac{\partial p_1^{t*}}{\partial b_1} \geq 0$. The first term in (2.52b) is positive. The denominator of the second term is positive due to earlier results. The numerator is also clearly positive. Hence, we find $\frac{\partial p_2^{t*}}{\partial b_1} \geq 0$. The denominator of (2.52c) is positive. The numerator may easily shown to be negative. With the negative sign in front, we find $\frac{\partial y_1^{t*}}{\partial b_1} \geq 0$ and $\frac{\partial y_2^{t*}}{\partial b_1} \leq 0$.

For part (b), we only show the result for the instances where both items are overstocked. The analysis for the other cases are similar. Through the optimality conditions we find

$$\frac{\partial p_1^{t*}}{\partial a_{11}} = \frac{A+B+C}{(a_{11}a_{22}-a_{12}^2)\left(\begin{array}{l}(4-2a_{11}\beta V_{11}-4a_{12}\beta V_{12}-2a_{22}\beta V_{22}\\ \qquad\qquad +(a_{11}a_{22}-a_{12}^2)\beta^2(V_{11}V_{22}-V_{12}^2))\end{array}\right)} \tag{2.53a}$$

where

$$A = 2a_{22}(c_1 - \lambda_1^{t*} - 2p_1^{t*}) \tag{2.54a}$$

$$B = (-a_{12}^2\beta V_{11} - 2a_{11}a_{22}\beta V_{12} - a_{22}^2\beta V_{22})(c_1 - \lambda_1^{t*} - 2p_1^{t*}) \tag{2.54b}$$

$$C = p_1^{t*}\left(2(a_{11}a_{22}-a_{12}^2)\beta V_{11} - a_{22}(a_{11}a_{22}-a_{12}^2)\beta^2(V_{11}V_{22}-V_{12}^2)\right) \tag{2.54c}$$

Substituting the expression for p_1^{t*} from (2.22) and (2.23a), we find that

$$c_1 - \lambda_1^{t*} - 2p_1^{t*} = \frac{a_{12}b_2 - a_{22}b_1}{a_{11}a_{22}-a_{12}^2} \leq 0$$

Furthermore, in A, $a_{22} > 0$ and in B, $-a_{12}^2\beta V_{11} - 2a_{11}a_{22}\beta V_{12} - a_{22}^2\beta V_{22} \geq -a_{12}^2\beta V_{12} - 2a_{11}a_{22}\beta V_{12} - a_{22}^2\beta V_{12} = -\beta V_{12}(a_{12}+a_{22})^2 \geq 0$. Thus, both terms A and B are nonpositive. In C, $p_1^{t*} \geq 0$, and both $2(a_{11}a_{22}-a_{12}^2)\beta V_{11}$ and $-a_{22}(a_{11}a_{22} - a_{12}^2)\beta^2(V_{11}V_{22}-V_{12}^2)$ are nonpositive. Thus C is nonpositive as well. Since the denominator of (2.53a) has been shown to be nonnegative previously, we find that $\frac{\partial p_1^{t*}}{\partial a_{11}} \leq 0$.

A similar analysis proves part (c). □

CHAPTER III

Optimal Control of an Assembly System with Demand for the End-Product and Intermediate Components

3.1 Overview

In this chapter, we consider the production and admission control of a two-stage manufacturing system where intermediate components are produced to stock in the first stage and an end-product is assembled from these components through a second stage assembly operation which may allow backorders. The manufacturing firm faces two types of demand. The one directed at the end-product is satisfied immediately if there are available products in inventory, and the firm has the option to accept the order for later delivery or to reject the order if no inventory is available. The second type of demand is for any of the intermediate components and the firm again has the option to accept the order or reject it to keep the components available for assembly purposes. We provide structural results for the demand admission, component production and product assembly decisions.

We also extend the model to take into account multiple customer classes based on revenue and a more general revenue collecting scheme where only an upfront partial payment for an item is received if a customer demand is accepted for future delivery with the remaining revenue received upon delivery. Since the optimal policy structure is rather complex and defined by switching surfaces in a multidimensional space, we also propose a heuristic policy that is easily implementable regardless of the problem size and test its performance under a variety of example problems.

3.2 Introduction

This chapter focuses on a manufacturing setting where a firm, having both component production and final assembly operations, faces demands for its end-product as well as the intermediate components.

Several business practices may lead a firm to operate within this setting. For example, consider a major appliance manufacturer such as Whirlpool, that produces various components and assembles them into a refrigerator. In addition to the demand for the refrigerator, Whirlpool also supplies individual components such as compressors in order to sustain its after-sales service operations. In many instances, efficient production control and demand management skills may be crucial for profitability when the product and component sales both have a significant contribution to the firm's revenues. The after-sales service is regarded as a high profit margin business and has become a comparable revenue generator throughout numerous industries. According to Cohen et al [17], in industries such as automobiles and white goods, the earlier units that companies have sold over the years have created aftermarkets four to five times larger than the original product markets. Consequently, businesses across many industries earn on average 45% of gross profits from the aftermarket.

As another example, consider TRW, which produces a range of automotive components such as braking, steering and suspension systems. TRW has a unit that makes engineered fastener components for its own products but the fastener unit also sells fasteners to other Tier 1 automotive suppliers which may sometimes even compete with TRW. Thus, at any point in time, the fastener unit has the option to accept or reject outside demand but also needs to coordinate its production policies to serve the demand arising from the assembly of its own products.

Our main objective in this chapter is to address several important decisions that a firm operating within this setting faces. Specifically, we will be focusing on the following questions: 1) How many of each type of intermediate components should the firm produce? 2) How should the firm decide whether to accept or reject an order for any of these intermediate components? 3) How does the firm determine whether to initiate the assembly of another end-product? 4) How does the firm regulate end-product admissions to its assembly queue?

The remainder of this chapter is organized as follows. In Section 3.3, we review the related literature. We provide the problem formulation in Section 3.4. In Sections 3.5 and 3.6, we characterize the structure and sensitivity of the optimal policy. In Section 3.7, we discuss extensions of the original model to multiple customer classes and partial revenue collecting schemes. In Section 3.8, we devise a heuristic solution approach and provide numerical results to evaluate the performance of the heuristics for a variety of problem instances. Finally, we conclude in Section 3.9. The proofs of all theoretical results are provided in Section 3.10.

3.3 Literature Review

This study interconnects the two research areas of assembly and admission control. There exists a rich literature on inventory control of assembly systems. An extensive literature survey has been provided by Song and Zipkin [53]. In one of the earliest works, Schmidt and Nahmias [50] study an assembly system with two components and a single final product that is assembled-to-stock. They assume a two stage manufacturing system where both the production and assembly stages have deterministic lead times. They identify the optimal assembly policy which states that there exists a target assemble-up-to point to reach as long as there are available components. They also identify the optimal production policies for the components which follows a modified base-stock policy due to differing replenishment lead times for the components. Rosling [48] extends the findings of Schmidt and Nahmias to multi-stage assembly systems by also assuming deterministic lead times.

In more recent works on pure assembly systems which are closest to our setting, Benjaafar and ElHafsi [4] consider the production and inventory control of a multi-component assembly system with several customer classes. Assuming instantaneous assembly, they show that a state dependent base stock policy is optimal for component production and there exists state dependent rationing levels for different demand classes. A subsequent work by Benjaafar et al [5] relaxes their earlier assumption of instantaneous assembly and incorporates multiple production and inventory stages. They again characterize the structure of the optimal production and rationing policies in the presence of multiple customer classes and show that production at each stage follows a state dependent base-stock policy which decreases with the inventory level of downstream items and increases with the inventory level of all other items. As in their previous work, demand admission for the product follows state dependent rationing levels.

In Benjaafar et al. [5], all customer classes require the same end-product but are willing to pay different amounts for the product. Therefore, rationing decisions are taken at the end product inventory level to prioritize several demand streams for the same product. In our setting, different customers demand different products, either the end-product or any of the intermediate components.

There is also a rich literature on admission control which falls outside the assembly systems classification. Stidham [54] presents a review of the literature on admission control for a single class make-to-order queue. Ha [31] considers a single item make-to-stock production system with several demand classes and lost sales. He shows that the optimal admission control policy is characterized by stock rationing levels for each demand class. He later extends the results to allow backorders in Ha [30] for a make-to-stock production system with two priority classes.

Specifically, controlled arrival to multiple nodes of queues in series has also attracted interest. Ghoneim and Stidham [26] study one such setting with two queues in series where customer arrivals to the first queue go through service in both queues whereas customer arrivals to the second queue only require service by that queue. They show that the optimal demand admission policy has a monotonic structure. Ku and Jordan [37] also study a similar system with finite queue sizes. They introduce randomness on whether a customer admitted to the first queue will actually stay in the system to get service from the second queue. They show that the optimal admission policy is defined by a monotonic threshold. In a subsequent work, Ku and Jordan [38], extend their results to systems with parallel first-stage queues. Duenyas and Tsai [20] study a two-stage production/inventory system where there is demand for the end product as well as the intermediate product with admission control on the demand for the latter. They derive the structure of the optimal policy for the centralized control problem and consider several pricing schemes for the decentralized case that achieves the profits of the centralized problem. Their formulation for the centralized control problem where there is only a single component and no admission control on the end-product is a special case of the problem considered in this chapter.

Our work is also related to the general assemble-to-order manufacturing systems literature involving multiple products assembled from a selection of intermediate components. As stated in Song and Zipkin [53], optimal policies regarding such general systems are still unknown. Several authors have focused on control policies assuming an independent base-stock order policy along with some allocation rule, such as committing inventories to the earliest backlog or simply following a first come first serve allocation. Examples of such work include those of Hausman et al [32], Song et al [51], and Akcay and Xu [2]. Our formulation may be regarded as a special case of the general system in the sense that it only allows demands - in addition to the demand for the end product - for one component at a time rather than an arbitrary selection of multiple components. Albeit limited compared to a general product portfolio architecture, our model enables us to fully characterize the optimal policies.

3.4 Problem Formulation

We consider a two-stage assembly system as shown in Figure 3.1. In the first stage, N intermediate components (also referred to as intermediate products) are produced-to-stock in exclusive subassembly lines and in the second stage, they are assembled into a single end-product. There are two types of demand sources in the system. The first type of demand is for the end-product, arriving based on a Poisson distribution with rate λ_0. The second type of demand is directly for the intermediate components. Customers may request a specific component i with a demand rate λ_i, $i = 1, 2, ...N$.

We assume that production of a unit of component i ($i = 1, 2, ...N$) takes an exponentially distributed amount of time with mean $\frac{1}{\mu_i}$. For each unit of component

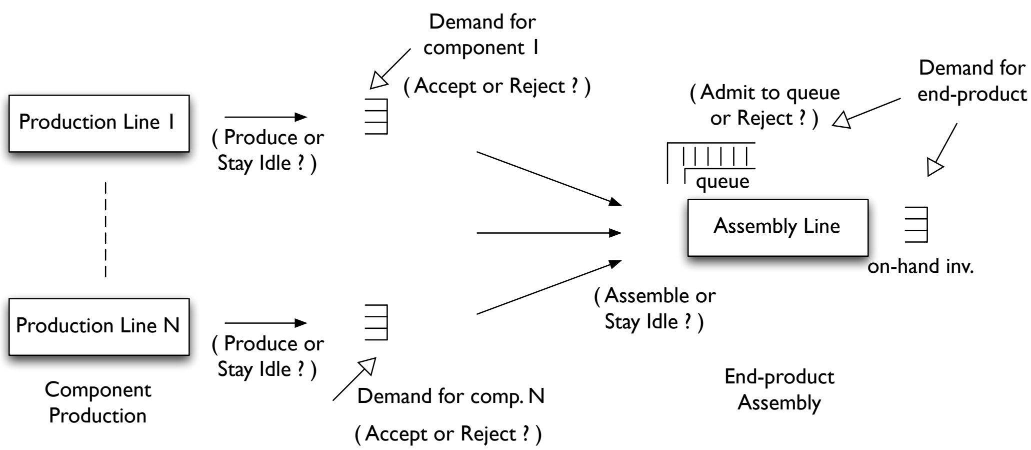


Figure 3.1: The assembly system demonstrating the demand for intermediate components as well as the end-product and the corresponding decisions.

i kept in stock, inventory holding costs are accrued at the rate of h_i per unit time.

The subassembly lines feed a single downstream assembly line, referred to as the second stage of the manufacturing system. During this assembly stage, one unit of each type of component is drawn from its inventory, and assembled into a single end-product. We let the assembly operation for the product also take an exponentially distributed amount of time with mean $\frac{1}{\mu_0}$. The firm incurs inventory holding and backorder costs at the rate of h_0 and b_0 per unit time for each product kept in stock and order kept in the assembly queue, respectively. We note that preemptions are allowed during both the assembly and production operations. In addition, we exclude the cost of subassembly and assembly operations from the model since, without loss of generality, we can define the revenues from individual component and end-product demands as marginal revenues.

The decision epochs considered in this model consists of all demand arrivals together with the production and assembly completions. At each decision epoch, a policy specifies whether a production server should stay idle or produce a unit of the corresponding component and whether the assembly line should stay idle or initiate the assembly of another end-product. At decision epochs corresponding to demand arrivals for the end-product or to individual components, the policy determines whether to accept or reject the orders.

The goal is to find a policy which maximizes the average profit per unit time through an infinite horizon. The profit is the revenue from accepted orders minus the inventory holding costs and the backorder costs due to orders waiting in the assembly queue.

The optimal production, assembly and admission control problem can be formulated as a Markov Decision Process. The state $(\mathbf{x}, y) \in S$ where $\mathbf{x} = (x_1, x_2, ..., x_N)$ and $x_i \geq 0$, $\forall\ i = 1, ..., N$ is defined such that $x_1, x_2, ..., x_N$ denote the amount of inventory of components 1 through N respectively, and y denotes the inventory position for the end-product. The states at which $y < 0$ correspond to customer orders waiting in the assembly queue whereas states at which $y > 0$ indicate that there is available inventory ready to satisfy end-product demand. We let $v(\mathbf{x}, y)$ denote the relative value function of being in state $(\mathbf{x}, y)$ and g be the average profit per transition, where transitions occur with rate $\Lambda = \sum_{i=0}^{N}(\lambda_i + \mu_i)$, resulting in an average profit per unit time of $g\Lambda$. We use uniformization as in Lippman [40] to write the average profit infinite horizon dynamic programming formulation. To keep the notation simple and to assist us in the analysis to follow, we first introduce a set of operators to represent the firm's decisions.

First, we will consider the end-product demand admission decision. When an order for the end-product arrives, if the end-product inventory is positive, the order is met immediately from inventory. However, if there is no available end-product inventory, the firm has the option to accept or reject the order. Each accepted order generates a revenue of R_0. If an order is rejected, it is considered as lost sales. If the end-product inventory is zero, and we accept the order, we incur a backorder cost. Operator T_0^1 in (3.1) represents this decision regarding product admission.

$$T_0^1 v(\mathbf{x}, y) = \begin{cases} v(\mathbf{x}, y-1) + R_0 & \text{if } y > 0 \\ \max[v(\mathbf{x}, y-1) + R_0, v(\mathbf{x}, y)] & \text{if } y \leq 0 \end{cases}$$

Similarly, there is an admission decision associated with each demand arrival for any of the intermediate components. Acceptance of a demand for a specific component i leads to a revenue of R_i whereas rejection of an order leads to lost sales. Since components are produced to stock, an order for a component may be accepted only if the inventory of the corresponding component is positive. Thus, component demand cannot be backordered. The operator T_i^1 defined below corresponds to the component demand admission decision where $I_{(\cdot)}$ denotes the indicator function and $\mathbf{e}_i$ is the i^{th} unit vector.

$$T_i^1 v(\mathbf{x}, y) = \max\left[(v(\mathbf{x} - \mathbf{e}_i, y) + R_i) \cdot I_{(x_i > 0)} + v(\mathbf{x}, y) \cdot I_{(x_i = 0)}, v(\mathbf{x}, y)\right]$$

Finally, operators T_0^2 and T_i^2 defined below correspond to the assembly initiation and component production decisions, respectively.

$$T_0^2 v(\mathbf{x}, y) = \max\left[(v(\mathbf{x} - \mathbf{1}, y+1)) \cdot I_{(x_i > 0\ \forall i)} + v(\mathbf{x}, y) \cdot I_{(\exists\, i\ |\ x_i = 0)}, v(\mathbf{x}, y)\right]$$

$$T_i^2 v(\mathbf{x}, y) = \max[v(\mathbf{x} + \mathbf{e}_i, y), v(\mathbf{x}, y)]$$

We now present the average profit infinite horizon dynamic programming formulation.

$$v(\mathbf{x}, y) + g = \frac{1}{\Lambda}\left(-\sum_{i=1}^{N}(h_i\, x_i) - h_0\, y^+ - b_0\, y^- + \sum_{i=0}^{N}\left(\lambda_i\, T_i^1\, v(\mathbf{x}, y) + \mu_i\, T_i^2\, v(\mathbf{x}, y)\right)\right) \tag{3.1}$$

where $y^+ := \max(y, 0)$, and $y^- := -\min(y, 0)$.

In (3.1), the terms $\frac{1}{\Lambda}(-\sum_{i=1}^{N} h_i x_i - h_0 y^+ - b_0 y^-)$ denote, respectively, the expected costs per decision epoch due to holding of component inventories, holding of end-product inventories, and backordering customer orders in the queue. The terms multiplied by $\lambda_i, (i = 0, 1, 2, ..., N)$ correspond to transitions and revenues generated with the arrival of a demand for the end-product when i=0 and the components when i=1,...,N. Finally, the terms multiplied by μ_i correspond to transitions and revenues generated by a product assembly when i=0 and by a component production completion opportunity when $i = 1, ...N$. We also define the operator T by letting $Tv(\mathbf{x}, y)$ to refer to the right hand side of (3.1).

3.5 Structure of Optimal Production, Assembly and Admission Policies

In this section, we characterize the optimal production, assembly and admission policies. The main questions of interest are the following: (1) Should the firm produce an additional unit of a component or not? (2) If a demand arrives for an individual component, should this demand be satisfied or rejected in order to keep the components available for the end-product assembly? (3) When there are available components, should another unit of an end-product be assembled? (4) When a demand arrives for the end-product while no inventory is available, should the firm admit the demand to the assembly queue or reject it?

First, we introduce the following difference operators that will facilitate the characterization of the optimal policy structure. For any real valued function v on the state space, we define:

$$D_i v(\mathbf{x}, y) = v(\mathbf{x} + \mathbf{e}_i, y) - v(\mathbf{x}, y) \ \ \forall \, i = 1, \ldots, N,$$

$$D_p v(\mathbf{x}, y) = v(\mathbf{x}, y + 1) - v(\mathbf{x}, y),$$

$$D_{\text{-}\mathbf{1},p} v(\mathbf{x}, y) = v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{1}, y).$$

D_i represents the additional value of an additional unit of component type-i inventory under value function v. D_p is the additional value of an additional unit of end-product inventory. Finally, $D_{\mathbf{-1},p}$ refers to the value of having an additional unit of an end-product relative to the value of keeping the components in component inventories.

Let V be the set of functions defined on the state space such that if $v \in V$, then $\forall\ i, j = 1, \ldots N$ where $j \neq i$;

(i) $D_i v(\mathbf{x}, y)\ \downarrow x_i, \uparrow x_j, \downarrow y\ \ \forall\ i = 1, \ldots, N$

(ii) $D_p v(\mathbf{x}, y)\ \downarrow x_i, \downarrow y$, and $\leq R_0$ for $y > 0$

(iii) $D_{\mathbf{-1},p} v(\mathbf{x}, y)\ \uparrow x_i, \downarrow y$

The above conditions are the sub- and super-modularity conditions on v and characterize the structure of the optimal component production and rationing policies. For example, $D_i \downarrow x_i$ means that the additional value gained by producing a unit of component type-i gets smaller with each additional unit of component type-i inventory. Hence, if it is optimal not to produce component type-i in state $(\mathbf{x}, y)$, it remains optimal not to produce it in state $(\mathbf{x} + \mathbf{e}_i, y)$. This in turn implies that if condition (i) holds, the component production policies follow state-dependent base-stock policies. Further, $D_i \uparrow x_j$ and $D_i \downarrow y$ mean, respectively, that the base-stock level for component type-i is nondecreasing with the inventory of other components and nonincreasing with the end-product inventory. Consequently, since backorders for the end-product imply a negative inventory position for this product, as the number of customers waiting in the assembly queue increases, the base-stock level for component type-i increases.

We further introduce secondary difference operators followed by a set of additional conditions that facilitate our derivation of the optimal policy structure.

$D_{\mathbf{1}}v(\mathbf{x}, y) = v(\mathbf{x} + \mathbf{1}, y) - v(\mathbf{x}, y)$

$D_{i,j}v(\mathbf{x}, y) = v(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j, y) - v(\mathbf{x}, y)\ \forall\ i, j = 1, \ldots, N$ where $j \neq i$

$D_{-i,p}v(\mathbf{x}, y) = v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{e}_i, y)$

$D_{i,-\mathbf{1},p}v(\mathbf{x}, y) = v(\mathbf{x} + \mathbf{e}_i, y + 1) - v(\mathbf{x} + \mathbf{1}, y)$

Lemma 3.1. *If a value function v satisfies conditions* (i)-(iii)*, then v also satisfies the following conditions $\forall\ i, j, k = 1, \ldots N$ where i, j, k are distinct:*

(iv) $D_{\mathbf{1}}v(\mathbf{x}, y) \downarrow x_i, \downarrow y$

(v) $D_{i,j}v(\mathbf{x}, y) \downarrow x_i, \downarrow x_j, \uparrow x_k, \downarrow y$

(vi) $D_{-i,p}v(\mathbf{x}, y) \uparrow x_i, \downarrow x_j, \downarrow y$

(vii) $D_{i,-\mathbf{1},p}v(\mathbf{x}, y) \downarrow x_i, \uparrow x_j, \downarrow y$

Proof: The proof of Lemma 3.1 is provided in Section 3.10.1.

As Lemma 3.1 reveals, the above relations are implied solely by conditions (i)-(iii). Although they do not have direct ramifications on the optimal policy structure, their frequent appearances in the analysis of the following lemma warrants their universal treatment. Lemma 3.2 shows that the conditions (i)-(iii) are preserved under the operator T.

Lemma 3.2. *If $v \in V$ then, $T_0^1 v$, $T_i^1 v$, $T_0^2 v$, $T_i^2 v$, and $Tv \in V \quad \forall i = 1, \ldots, N$.*

Proof: The proof of Lemma 3.2 is provided in Section 3.10.1.

We now present the main result. The policies defined below reflect the structure of the optimal policy for our model.

Definition 3.1. *Consider the N-dimensional integer valued vectors $(\boldsymbol{x}_{-i}, y)$ and $(\boldsymbol{x})$ where $\boldsymbol{x}_{-i}$ denotes the inventory level of all components except component type-i. Define the following (state-dependent) component rationing and product admission policies:*

(a) Rationing policy for component type-i: A demand for component type-i is satisfied if the amount of inventory for the component is higher than a rationing threshold $\alpha_i(\boldsymbol{x}_{-i}, y)$, *i.e. if* $x_i \geq \alpha_i(\boldsymbol{x}_{-i}, y)$. *If* $x_i < \alpha_i(\boldsymbol{x}_{-i}, y)$, *the demand for the component is rejected.*

(b) Admission policy for end-product: A demand for the end-product is admitted if the end-product inventory position is higher than an admission threshold $\beta(\boldsymbol{x})$. *If* $y < \beta(\boldsymbol{x})$, *the product demand is rejected.*

Definition 3.2. *For the integer valued vectors* $(\boldsymbol{x}_{-i}, y)$ *and* $(\boldsymbol{x})$, *define the following (state-dependent) base-stock policies for component production and product assembly:*

(a) Base-stock policy for component type-i: An additional unit of component type-i is produced if the inventory for component type-i is less than a production threshold $\gamma_i(\boldsymbol{x}_{-i}, y)$. *If* $x_i > \gamma_i(\boldsymbol{x}_{-i}, y)$, *the production resource for the component stays idle.*

(b) Base-stock policy for end-product: When all components are available, the assembly operation is initiated if the end-product inventory position is lower than an assembly threshold level $\delta(\boldsymbol{x})$. *Otherwise, the assembly resource for the product stays idle.*

The following theorem gives the characterization of the optimal policy structure.

Theorem 3.1. (a) *Demand admissions for each individual component type-i,* $i = 1, \ldots, N$, *follows a rationing policy characterized by the rationing threshold* $\alpha_i(\boldsymbol{x}_{-i}, y)$. *Furthermore,* $\alpha_i(\boldsymbol{x}_{-i}, y)$ *is non-decreasing with* x_j, $j = 1,\ldots,N$, $j \neq i$, *and non-increasing with* y.

(b) *Admissions for end-product demand follow an admission policy characterized by the threshold* $\beta(\boldsymbol{x})$ *which is non-increasing with* x_i, $\forall i = 1,\ldots, N$.

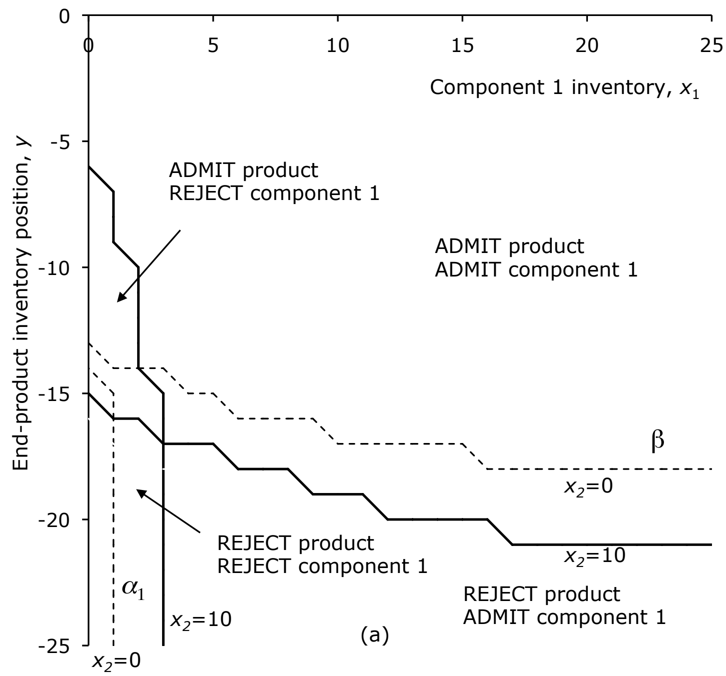


Figure 3.2: Optimal demand admission decisions for component type-1 and the end-product.

(c) *Production policy for each component type-i is defined by a base-stock policy with a production threshold* $\gamma_i(\boldsymbol{x}_{-i}, y)$. *Furthermore,* $\gamma_i(\boldsymbol{x}_{-i}, y)$ *is non-decreasing with* x_j, $j = 1,\ldots,N$, $j \neq i$, *and non-increasing with* y.

(d) *Assembly policy for the end-product follows a base-stock policy with an assembly threshold* $\delta(\boldsymbol{x})$ *which is non-decreasing with* x_i, $\forall i = 1,\ldots, N$.

Proof: The proof of Theorem 3.1 is provided in Section 3.10.1.

Figures 3.2 and 3.3 illustrate the structure of the optimal policies described in Theorem 3.1 for an example problem with two components with the following parameters: $\lambda_0 = 5$, $\lambda_1 = 3$, $\lambda_2 = 4$, $\mu_0 = 8$, $\mu_1 = \mu_2 = 10$, $R_0 = 40$, $R_1 = 20$, $R_2 = 10$, $b_0 = 4$, $h_0 = 2$, and $h_1 = h_2 = 1$.

The switching curves α_1 and β in Figure 3.2 depict the component rationing (for component type-1) and end-product admission threshold levels, respectively.

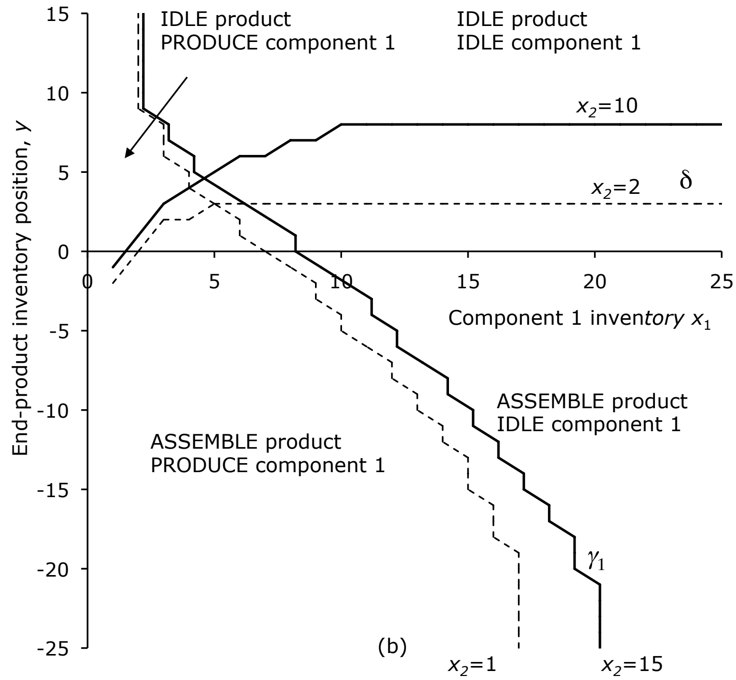


Figure 3.3: Optimal production and assembly decisions for component type-1 and the end-product.

As indicated in Theorem 3.1 part (a), we observe that the rationing threshold for component 1 increases as there are more units of component 2 in inventory and there are more end product customers backlogged in the system. Regarding the end product admission decision, we observe that the admission threshold decreases (more units are admitted) when there are more components of either type in inventory as stated in Theorem 3.1 part (b).

Figure 3.3 displays the structure of the optimal component production and end-product assembly policies represented by the switching curves γ_1 and δ, respectively. Similar switching curves exist for the type-2 component.

We observe that the production threshold for component 1 increases as there are more units of component 2 in inventory and as there are more end product customers backlogged in the system as depicted in Theorem 3.1 part (c). This is essentially a similar dynamics as the component demand rationing, i.e. when end-product queue is longer and when component type-2 is available, more units of component 1 are desired and used for the final product assembly. Lastly, as stated in Theorem 3.1 part (d), more units of the end product will be assembled if either type of component has additional inventories.

For this example problem consisting of two components, the optimal threshold values are defined by switching surfaces in three dimensions. The solid and the dotted curves in both figures are results of two-dimensional cuts on the switching surfaces at two separate values of the type-2 component inventory levels. For a general problem with N components, each threshold is defined by a switching surface embedded in an $N+1$ dimensional Euclidean space.

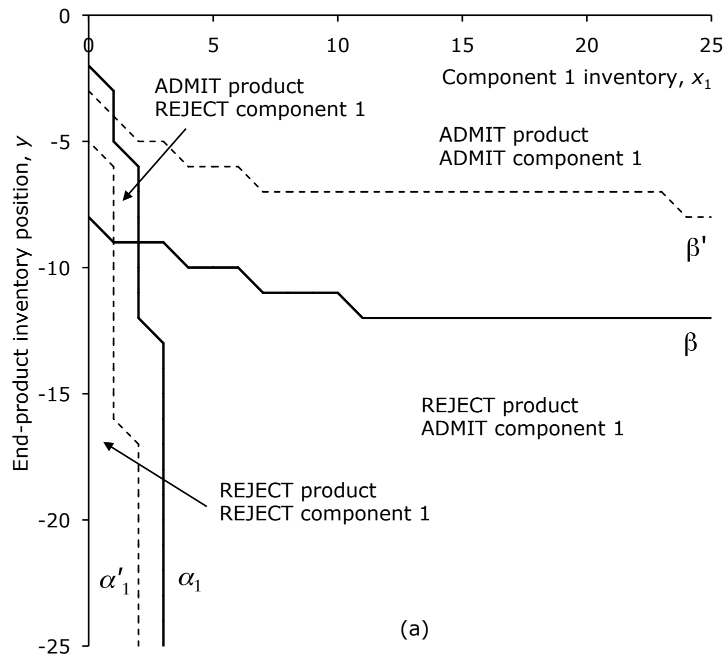
3.6 Sensitivity of the Optimal Policy

Next, we will examine how the optimal policies described in Theorem 3.1 change as the end-product revenue decreases. We will use the prime symbol (') while referring to the relative value function and parameters of the modified problem.

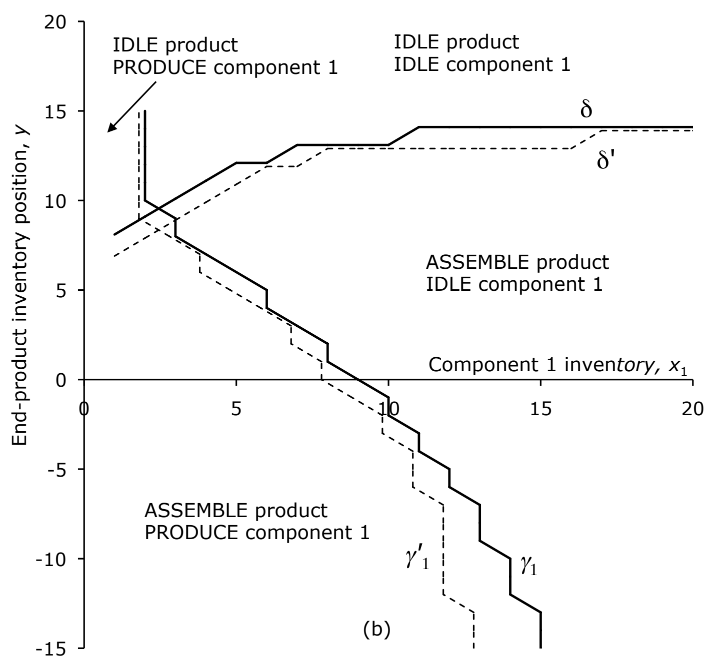
Theorem 3.2. *Suppose that* $\mu_i' = \mu_i$, $\lambda_i' = \lambda_i$ *and* $h_i' = h_i$ *for i=0,1,...,N;* $b_0' = b_0$, *and* $R_i' = R_i$ *for* $i = 1, ..., N$ *whereas* $R_0' < R_0$. *Then,* $\alpha_i'(\boldsymbol{x}_{-i}, y) \leq \alpha_i(\boldsymbol{x}_{-i}, y)$, $\beta'(\boldsymbol{x}) \geq \beta(\boldsymbol{x})$, $\gamma_i'(\boldsymbol{x}_{-i}, y) \leq \gamma_i(\boldsymbol{x}_{-i}, y)$, *and* $\delta'(\boldsymbol{x}) \leq \delta(\boldsymbol{x})$.

Proof: The proof of Theorem 3.2 is provided in Section 3.10.2.

Regarding the admission policies, Theorem 2 states that as the revenue from the end-product gets smaller, it may be optimal to switch from accepting a demand for the end-product to rejecting it, and from rejecting a demand for the intermediate product to accepting it. In terms of the production and assembly policies, it may be optimal to switch from assembling another end-product to staying idle, and from



(a) Sensitivity of the optimal demand admission decisions.



(b) Sensitivity of the optimal production and assembly decisions

Figure 3.4: Changes in optimal policies due to a decrease in the end-product revenue: $R_0' < R_0$

producing a component to staying idle. Figure 3.4 illustrates the changes in the optimal admission and production/assembly policies due to a product revenue change from $R_0 = 40$ to $R_0' = 25$ with remaining parameters set as $\lambda_0 = 6$, $\lambda_1 = 2$, $\lambda_2 = 3$, $\mu_0 = 8$, $\mu_1 = \mu_2 = 10$, $R_1 = 20$, $R_2 = 10$, $b_0 = 5$, $h_0 = 2$, and $h_1 = h_2 = 1$.

A decrease in the revenue of the end-product reduces the relative importance of satisfying an end-product demand compared to that of satisfying an individual component demand. Therefore, the system tends to admit less end-product demand in the assembly queue, and instead, it accepts more of the demand for the intermediate products. Consequently, this results in fewer end-products to be assembled. For component production, the decreased requirements due to fewer products assembled outweighs the increased requirement due to a higher number of individual component demands satisfied. As a result, fewer components of each type are produced.

Although the optimal policies are monotonic with respect to the end-product revenue, they don't necessarily have uniform monotonicities with respect to other problem parameters. We will show three such cases by way of counter examples. The solid lines in Figure 3.5 (a)-(c) display the threshold curves for assembly, component production, and component rationing, respectively for a two-component problem with parameters $\lambda_0 = 3, \lambda_1 = \lambda_2 = 2, \mu_0 = 4, \mu_1 = \mu_2 = 6, h_0 = 2, h_1 = h_2 = 1, b_0 = 2, R_0 = 30, R_1 = 10$, and $R_2 = 12$. In each of these figures, the dashed lines refer to the corresponding policies with one of the parameters modified as discussed below.

In Figure 3.5 (a), we observe the effects of lowering the revenue from a type-2 component on the optimal product assembly threshold. We observe that the threshold curves δ and δ' cross each other, hence the optimal policy does not possess monotonicity with respect to a change in R_2. A low value of R_2 shifts the priority from satisfying individual component type-2 demands towards producing end-products in-

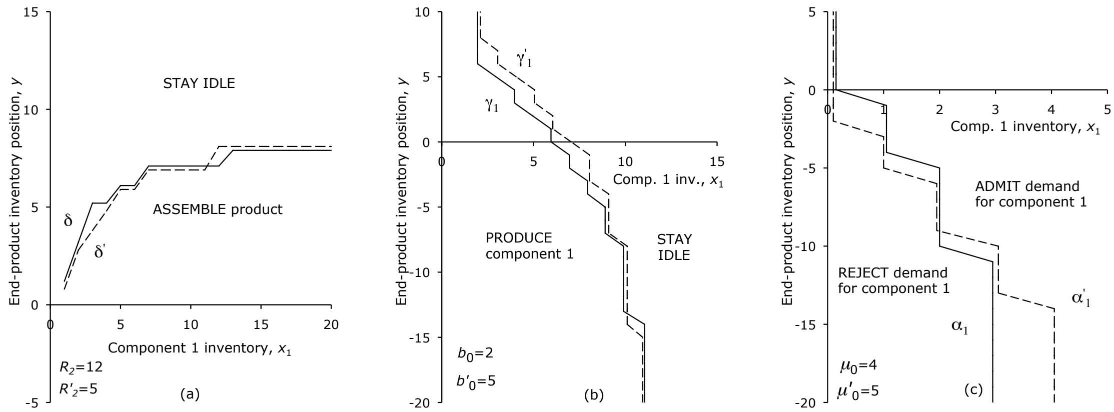


Figure 3.5: Counter examples for the optimal policy sensitivity on (a) component revenues, (b) backorder cost, and (c) product assembly rate

stead. This generates two sorts of dynamics. On the one hand, the priority shift towards the end-product enables the assembly line to have smoother access to component inventories thereby lowers the assembly threshold. On the other hand, as selling individual components separately brings in relatively lower revenues compared to selling them as an end-product, it is more favorable to turn the components into products, thus increasing the assembly threshold. Through numerical studies, we observe the former effect to have a higher influence when component inventories are low to moderate. This makes sense as the competition between the assembly operation and the individual component demands on a component is more critical when component inventories are scarce. At system states with high inventories for both components, the latter effect is more dominant resulting in a higher assembly threshold.

We analyze the effect of increasing the backorder cost on the optimal component production policies in Figure 3.5 (b). A higher backorder cost requires more of the product demand to be met from inventory and discourages demand admissions to the assembly queue if there are already a high number of backorders in the system.

Therefore, for states with on-hand inventories or moderate backorders, we tend to produce more components as we try to meet as much of the product demand as possible without further backordering. However, when there is already a high number of customers waiting in the queue, further product demand admission is prevented, hence the requirement for components decreases.

Finally, in Figure 3.5 (c) we observe the changes in the optimal component rationing policy based on an increase in the assembly process rate. We again observe two different dynamics. A faster assembly line can make up for a delayed availability of individual components. Therefore it enables more of the individual component demand to be satisfied resulting in lower component rationing thresholds. On the other hand, in order to utilize its fast pace, the assembly operation also requires high component availability allowing quick supplies. We observe that the first effect is stronger at system states where there is on-hand product inventory or only a few customers waiting in the assembly queue. However, when there is a high amount of backorders in the queue, the second effect is influential, saving the components for assembly purposes to quickly lower the number of backorders.

3.7 Extensions to the Original Model

3.7.1 Multiple Customer Classes

It is straightforward to extend our model to include multiple customer classes that are willing to pay different amounts for the same end-product as in Benjaafar et al [5] given that end-products supplied to any customer class require the same processing time and that they have identical backorder costs under which the original state space representation may be retained. In fact, we can also include multiple customer classes for the component demand. For example, let the demand for a type-i component arise by M_i customer classes with arrival rates λ_i^m where $m = 1, 2, ..., M_i$, generating a revenue of R_i^m. Without loss of generality, we rank R_i^m such that $R_i^1 \geq R_i^2 \geq ... \geq R_i^{M_i}$. In this modified problem, M_i operators of the form $T_i^{1,m} v(\mathbf{x}, y) = \max\left[(v(\mathbf{x} - \mathbf{e}_i, y) + R_i^m) \cdot I_{(x_i>0)} + v(\mathbf{x}, y) \cdot I_{(x_i=0)}, v(\mathbf{x}, y)\right]$ replace the original operator T_i^1. This results in the original problem given by (4.1) to have the terms $\sum_{m=1}^{M_i} \lambda_i^m \, T_i^{1,m} \, v(\mathbf{x}, y)$ instead of the term $\lambda_i \, T_i^1 \, v(\mathbf{x}, y)$.

The conditions set forth in the analysis of Lemma 3.2 and Theorem 3.1 suffices to show that the optimal policy has a similar structure as the one described in Theorem 3.1, except for a replacement of the component demand admission thresholds $\alpha_i(\mathbf{x}_{-i}, y)$ with multiple component admission threshold levels $\alpha_i^1(\mathbf{x}_{-i}, y) \leq \alpha_i^2(\mathbf{x}_{-i}, y) \leq ... \leq \alpha_i^{M_i}(\mathbf{x}_{-i}, y)$ for each demand class.

3.7.2 A Partial Revenue Collecting Scheme

In this section, we consider another extension to the basic model that takes into account a more general revenue collecting scheme. It is common practice in many businesses that if a customer is to be made to wait for and end-product, only an upfront partial payment for the item is collected rather than the item's full revenue. Consequently, the firm receives the remaining price of the item at the time of delivery. In such a setting, a discounted profit formulation is more valid since we would like to account for the time value of money. Interestingly, the policy structure described in Theorem 3.1 remains exactly the same with this more general revenue collection scheme.

Let r_0 denote the upfront payment amount received when a customer is admitted to the product assembly queue such that $R_0 \geq r_0 \geq 0$. Further, let the operators corresponding to the product demand admission and product assembly decisions be modified as follows:

$$T_0^1 v(\mathbf{x}, y) = \begin{cases} v(\mathbf{x}, y-1) + R_0 & \text{if } y > 0 \\ \max[v(\mathbf{x}, y-1) + r_0, v(\mathbf{x}, y)] & \text{if } y \leq 0 \end{cases}$$

$$T_0^2 v(\mathbf{x}, y) = \begin{cases} \max\left[(v(\mathbf{x}-\mathbf{1}, y+1)) \cdot I_{(x_i>0\ \forall i)} + v(\mathbf{x}, y) \cdot I_{(\exists\ i\ |\ x_i=0)}, v(\mathbf{x}, y)\right] \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } y \geq 0 \\ \max[(v(\mathbf{x}-\mathbf{1}, y+1) + R_0 - r_0) \cdot I_{(x_i>0\ \forall i)} \\ \qquad\qquad\qquad\qquad + v(\mathbf{x}, y) \cdot I_{(\exists\ i\ |\ x_i=0)}, v(\mathbf{x}, y)] \qquad \text{if } y < 0 \end{cases}$$

Then, using uniformization with transition rate $\Lambda = \phi + \sum_{i=0}^{N}(\lambda_i + \mu_i)$ where ϕ denotes the discount factor, we can rewrite the problem given in (3.1) as a discounted infinite horizon dynamic program as follows:

$$v(\mathbf{x}, y) = \frac{1}{\Lambda}\left(-\sum_{i=1}^{N}(h_i\ x_i) - h_0\ y^+ - b_0\ y^- + \sum_{n=0}^{N}\left(\lambda_j\ T_j^1\ v(\mathbf{x}, y) + \mu_n\ T_j^2\ v(\mathbf{x}, y)\right)\right) \tag{3.2}$$

The following theorem depicts the optimal policy structure.

Theorem 3.3. *For the problem given in* (3.2)*, the optimal demand admission, component production and product assembly policies follow the optimal policy structure described in Theorem 3.1. That is, demand admission for the end-product and the components are characterized by state-dependent admission thresholds and production and assembly decisions follow state-dependent base-stock policies with similar monotonicity properties as set forth in Theorem 3.1.*

Proof: The proof of Theorem 3.3 is provided in Section 3.10.3.

3.8 A Heuristic Policy

3.8.1 Construction of the Heuristic Policy

The optimal policy structure determined by Theorem 3.1 is fairly complex. In addition to the assembly and admission control decisions for the end-product, the firm also needs to make production and rationing decisions for each component. As shown in the previous section, all of these decisions are characterized by state dependent threshold levels. For a general problem with N components, each threshold is defined by a switching surface embedded in an $N+1$ dimensional space. Since the number of possible system states grows exponentially as the number of components gets larger, computing the optimal policy in such cases ceases to be a practical task.

For problems with a limited number of components, however, the optimal switching surfaces may be computed with relative ease as we have previously illustrated in Figures 3.2 and 3.3. Motivated by the ease of computation for a two-component problem and the prohibitive inefficiencies associated with problems of large sizes, we introduce the following heuristic solution approach. For each component type-i, we construct a two-component subproblem P_i that assumes its type-1 component as the component type-i of interest and aggregates all the remaining components into a type-2 component. In the original problem of N components, at each decision epoch corresponding to a demand arrival or production opportunity for component type-i, the heuristic policy maps the system state $(\mathbf{x}, y)$ to state $(x_i, \min\{x_j, j = 1, ..., N, j \neq i\}, y)$ in subproblem P_i and imitates the corresponding decision given at this state for component type-i.

The underlying assumption that leads us to construct two-component subproblems lies in the intuitive expectation that the production and rationing decisions for a certain component is influenced strongly by the component with the lowest inventory level as opposed to by others with higher inventory levels. In other words, it is the limiting component that mostly impacts the assembly capabilities and hence influences the amount of inventory to hold for others. Although exceptions to this conjecture may occur when there are discrepancies among component production rates, we will maintain this assumption as it allows us to develop a simpler heuristic.

Let the symbol ($\hat{}$) denote the parameters for the constructed two-component subproblem P_i. We define the revenues and cost parameters as $\hat{R_0} = R_0$, $\hat{R_1} = R_i$, $\hat{R_2} = \sum_{j \neq i} R_j$, $\hat{h_1} = h_i$, $\hat{h_2} = \sum_{j \neq i} h_j$, $\hat{h_0} = h_0$, and $\hat{b_0} = b_0$. The demand arrival, production and assembly rates for component type-1 and the end-product in problem P_i are set identically at their values in the original problem. Hence we define $\hat{\lambda_0} = \lambda_0$, $\hat{\lambda_1} = \lambda_i$, $\hat{\mu_0} = \mu_0$, and $\hat{\mu_1} = \mu_i$.

We handle the production and demand arrival rates corresponding to the type-2 component in subproblem P_i rather differently as this component reflects an aggregation of all the remaining ones. Since upon a demand arrival for component type-2, we receive the sum of the revenues from all remaining components, we adjust the demand arrival rate based on the time it takes for a demand to arrive for all components. This rate adjustment calls for a maximum of exponentially distributed random variables where this maximum itself no longer follows exponential distribution. The mean of the maximum of several random variables appears frequently in the reliability literature when calculating the reliability of a system consisting of several servers in parallel. Following equation (7.27) in Billinton and Allan [7], the mean of the maximum of n random variables $(Z_1, \ldots, Z_n)$ where each random variable Z_i is exponentially distributed with mean $1/\lambda_i$ is given by:

$$
\begin{aligned}
E[\max(Z_1,\ldots,Z_n)] = \;& \left(\tfrac{1}{\lambda_1}+\tfrac{1}{\lambda_2}+\cdots+\tfrac{1}{\lambda_n}\right)-\left(\tfrac{1}{\lambda_1+\lambda_2}+\tfrac{1}{\lambda_1+\lambda_3}+\cdots+\tfrac{1}{\lambda_i+\lambda_j}+\cdots\right)\\
&+\left(\tfrac{1}{\lambda_1+\lambda_2+\lambda_3}+\tfrac{1}{\lambda_1+\lambda_2+\lambda_4}+\cdots+\tfrac{1}{\lambda_i+\lambda_j+\lambda_k}+\cdots\right)\\
&-\cdots+(-1)^{n+1}\tfrac{1}{\sum_{i=1}^{n}\lambda_i}
\end{aligned}
$$

When calculating the demand arrival rate for the type-2 component, we treat the inverse of this mean time as the corresponding rate. For the production rate, we set $\hat{\mu_2}$ as the average of the production rate of all the remaining components, i.e., $\hat{\mu_2} = (\sum_{j\neq i} \mu_j)/(N-1)$.

For the component production and rationing decisions, we therefore construct a total of N such two-component subproblems P_i $(i = 1,\ldots,N)$. As for the end-product admission and assembly decisions, we follow an analogous argument by forming a single one-component subproblem P_0, where the parameters for the component are determined by aggregating all the components in a similar fashion. At each decision epoch corresponding to an end-product arrival or assembly opportunity, the heuristic policy maps the system state $(\mathbf{x}, y)$ to state $(\min\{x_i, i = 1, ..., N\}, y)$ in problem P_0 and imitates the corresponding decision given at this state.

3.8.2 Performance of the Heuristic Policy

Next, we evaluate the performance of the heuristic policy. Tables 3.1 and 3.2 compare the profits obtained by the optimal and the heuristic policies for 24 example problems each for systems consisting of three and four identical components, respectively. For each problem, we report the profits per unit time obtained by the optimal policy, our heuristic policy and a "strawman" heuristic policy we picked from the literature. The "strawman" heuristic policy is a base-stock/rationing policy for each product where a base-stock and a rationing level is set for each product which ignores the inventory levels for all other products (e.g. see Song et al. [51] and Benjaafar et al. [5] for such heuristic policies).

In Tables 3.1 and 3.2, the parameters varied include the revenues from the intermediate and end-products, the demand arrival rates for the intermediate and end-products, and the utilizations for the production and assembly lines.

The revenue parameters are selected to allow the testing of the heuristics for cases where the revenue from an end-product is higher, equal to, or lower than the sum of the revenues from intermediate components. Specifically, for both the three- and four-component problems, we evaluate the heuristics where the end-product revenue is $\frac{2}{3}, 1, \frac{4}{3}$ times the sum of the individual component revenues. For example, for the case of three intermediate components, the price pair $R_0 = 30$ and $R_i = 15$ corresponds to a revenue ratio of $\frac{2}{3}$ while the price pair $R_0 = 60$ and $R_i = 15$ corresponds to a revenue ratio of $\frac{4}{3}$. Generally, due to further processing, the end-product revenue may be at least as much as the sum of the revenues of its constituents. However, there may be examples where the reverse holds, such as after-sales parts that are sold at much higher prices. The revenue ratio of $\frac{2}{3}$ enables us to investigate the performance of the heuristics in such settings.

We also change the intermediate and end-product demand arrival rates between a high and a low ratio to observe the cases where individual sales are a significant part of the business and the cases where the focus is overwhelmingly on the end-product with occasional demands arising for intermediate products.

Finally, we vary both component production and assembly utilizations between low and high values by selecting μ_0 and μ_i such that $\frac{\lambda_0}{\mu_0} = \rho_0$ and $\frac{\lambda_i+\lambda_0}{\mu_i} = \rho_i \ \ \forall \ i = 1, ..., 4$. Throughout these example problems, we assume $b_0 = 0.2R_0$, $h_0 = 0.1R_0$, and $h_i = 0.1R_i \ \forall \ i = 1, \ldots, N$.

Table 3.1: Performance of the heuristics and independent base-stock/rationing policy for a system with three identical intermediate products

							Optimal	Heuristic		Indep. BS&R	
No	R_0	R_i	λ_0	λ_i	ρ_0	ρ_i	Profit	Profit	% diff.	Profit	% diff.
1	30	10	4	3	0.5	0.5	193.2	193.1	0.02	191.4	0.91
2	30	10	4	3	0.5	0.9	178.6	177.4	0.68	175.2	1.89
3	30	10	4	3	0.9	0.5	178.9	178.9	0.01	177.0	1.03
4	30	10	4	3	0.9	0.9	167.0	166.6	0.24	164.4	1.56
5	30	10	6	0.5	0.5	0.5	181.2	181.2	0.01	178.2	1.70
6	30	10	6	0.5	0.5	0.9	162.0	161.3	0.46	158.6	2.10
7	30	10	6	0.5	0.9	0.5	163.2	163.1	0.02	159.8	2.04
8	30	10	6	0.5	0.9	0.9	150.5	150.3	0.11	148.4	1.38
9	30	15	4	3	0.5	0.5	232.7	232.2	0.19	230.1	1.08
10	30	15	4	3	0.5	0.9	213.1	209.9	1.51	208.7	2.08
11	30	15	4	3	0.9	0.5	217.7	217.6	0.07	216.4	0.61
12	30	15	4	3	0.9	0.9	201.6	200.9	0.36	199.3	1.15
13	30	15	6	0.5	0.5	0.5	185.3	185.3	0.04	172.2	7.08
14	30	15	6	0.5	0.5	0.9	163.3	161.9	0.87	158.8	2.81
15	30	15	6	0.5	0.9	0.5	165.9	165.8	0.03	162.4	2.08
16	30	15	6	0.5	0.9	0.9	150.4	150.0	0.25	148.4	1.35
17	60	15	4	3	0.5	0.5	346.9	346.9	0.00	344.1	0.81
18	60	15	4	3	0.5	0.9	323.7	323.0	0.21	316.4	2.25
19	60	15	4	3	0.9	0.5	319.2	319.1	0.02	316.9	0.72
20	60	15	4	3	0.9	0.9	300.2	299.9	0.09	295.1	1.70
21	60	15	6	0.5	0.5	0.5	358.5	358.5	0.00	353.1	1.51
22	60	15	6	0.5	0.5	0.9	324.6	323.7	0.28	321.8	0.87
23	60	15	6	0.5	0.9	0.5	323.9	323.9	0.00	319.1	1.50
24	60	15	6	0.5	0.9	0.9	302.0	301.8	0.09	297.6	1.47

In Table 3.1, the average difference between the profits obtained by the optimal and the heuristic policy is 0.23% whereas the profit difference between the optimal policy and the independent base-stock/rationing policy is 1.74%. By making use of the inventory positions of the end-product and the limiting component, the heuristic policy performs better than the independent base-stock/rationing policy which only uses local inventory information. In fact, we observe that the heuristic policy outperforms the independent base-stock/rationing policy in all instances of the example problems.

Regarding product revenues, we find that the profit attained by the heuristic policy differs from that of the optimal policy for an average value of $0.41, 0.19$ and 0.09 corresponding to product revenue ratios of $\frac{2}{3}, 1$, and $\frac{4}{3}$, respectively, i.e. averages from problems No. 9-16, 1-8 and 17-24. Thus, we observe that the performance of the heuristic policy improves as the revenue from the end-product increases with respect to the sum of the revenues from intermediate components.

In terms of demand arrival rate ratios, we find that the heuristic policy performs slightly better at high arrival rates for the end-product and low arrival rates for intermediate products. These two properties suggest that the heuristic policy is also capable of controlling pure assembly systems with no exogenous demand for the intermediate components.

Lastly, we observe that the heuristic policy performs very well for problem instances where utilization for the production line is low with an average difference of 0.04% between the profits obtained by the optimal and the heuristic. This is somewhat expected since the heuristic plans the production of an item based on the inventory of the limiting product. A faster production rate allows withholding the processing of an item until the inventory position of the limiting item is restored. In settings where the production line utilization is high and the assembly line utilization is low, however, we find that the performance of the heuristic policy is degraded to an average difference of 0.67%. On the other hand, settings with low assembly utilizations also lead to even lower performance by the independent policy with an average difference of 2.09% and a maximum difference of up to 7.08% as observed in problem No. 13.

Table 3.2 is constructed in a similar fashion to Table 3.1 in order to evaluate the performance of the heuristic policy when applied to systems with a larger number of intermediate products.

We find that the average difference between the optimal and the heuristic profit is 0.31% while the average difference between the optimal profit and the profit obtained by the independent base-stock/rationing policy is 1.91%. Comparing the three- and four-component results, an important characteristic of the heuristic policy seems to be its retained robustness moving from a three-component problem to a four-component one.

In accordance with the results obtained by Table 3.1, a closer look into Table 3.2 also reveals that the performance of the heuristic policy is strongest in settings where the revenue from the end-product is higher compared to the sum of the revenues from intermediate components, when the demand rate for the end-product is higher compared to the demand rate for intermediate components, and when the production line utilizations are low. The heuristic is robust with respect to the assembly line utilization.

Finally, we are interested in how the heuristic policy performs in systems with asymmetric rate, revenue and cost parameters. In Table 3.3, we construct a base case labeled as problem No. 0, for which we set $R_0 = 50, R_1 = 20, R_2 = 15, R_3 = 10, R_4 =$

Table 3.2: Performance of the heuristics and an independent base-stock/rationing policy for a system with four identical intermediate products

							Optimal	Heuristic		Indep. BS&R	
No	R_0	R_i	λ_0	λ_i	ρ_0	ρ_i	Profit	Profit	% diff.	Profit	% diff.
1	40	10	4	3	0.5	0.5	257.1	257.0	0.07	254.8	0.91
2	40	10	4	3	0.5	0.9	236.9	234.7	0.90	232.2	1.97
3	40	10	4	3	0.9	0.5	238.0	237.9	0.04	235.6	1.02
4	40	10	4	3	0.9	0.9	221.5	220.7	0.35	216.3	2.34
5	40	10	6	0.5	0.5	0.5	240.9	240.8	0.05	235.7	2.17
6	40	10	6	0.5	0.5	0.9	213.1	211.7	0.67	208.6	2.13
7	40	10	6	0.5	0.9	0.5	216.8	216.7	0.00	212.4	2.04
8	40	10	6	0.5	0.9	0.9	198.2	198.0	0.07	195.3	1.45
9	40	15	4	3	0.5	0.5	309.6	309.2	0.14	305.0	1.49
10	40	15	4	3	0.5	0.9	282.3	277.1	1.82	272.4	3.50
11	40	15	4	3	0.9	0.5	289.5	289.4	0.05	287.2	0.79
12	40	15	4	3	0.9	0.9	267.0	266.1	0.36	264.1	1.11
13	40	15	6	0.5	0.5	0.5	246.2	245.9	0.11	227.1	7.76
14	40	15	6	0.5	0.5	0.9	214.0	211.2	1.33	208.1	2.76
15	40	15	6	0.5	0.9	0.5	220.0	219.9	0.06	215.7	1.96
16	40	15	6	0.5	0.9	0.9	197.4	196.8	0.33	194.1	1.69
17	80	15	4	3	0.5	0.5	461.9	461.7	0.05	457.9	0.86
18	80	15	4	3	0.5	0.9	429.7	428.4	0.32	417.8	2.79
19	80	15	4	3	0.9	0.5	424.8	424.6	0.05	421.5	0.76
20	80	15	4	3	0.9	0.9	398.6	398.0	0.15	393.7	1.22
21	80	15	6	0.5	0.5	0.5	476.9	476.8	0.02	471.8	1.05
22	80	15	6	0.5	0.5	0.9	428.0	426.2	0.41	425.1	0.67
23	80	15	6	0.5	0.9	0.5	430.7	430.6	0.01	422.7	1.85
24	80	15	6	0.5	0.9	0.9	399.1	398.4	0.16	393.1	1.49

Table 3.3: Performance of the heuristics and an independent base-stock/rationing policy for an asymmetric system with four intermediate products

No	R_0	R_1	λ_0	λ_1	ρ_0	ρ_1	b_0	h_0	h_1	Optimal Profit	Heuristic Profit	Heuristic % diff.	Indep. BS&R Profit	Indep. BS&R % diff.
0	50	20	10	1	0.7	0.8	10	5	2	552.1	551.0	0.20	540.8	2.05
1	100	20	10	1	0.7	0.8	10	5	2	1058.2	1057.7	0.05	1044.9	1.26
2	30	20	10	1	0.7	0.8	10	5	2	356.5	349.8	1.88	345.8	3.00
3	50	40	10	1	0.7	0.8	10	5	2	571.4	570.3	0.19	561.4	1.75
4	50	5	10	1	0.7	0.8	10	5	2	538.8	537.8	0.17	526.8	2.22
5	50	20	20	1	0.7	0.8	10	5	2	1046.6	1045.9	0.07	1035.2	1.09
6	50	20	5	1	0.7	0.8	10	5	2	307.0	306.0	0.32	299.0	2.62
7	50	20	10	5	0.7	0.8	10	5	2	630.5	629.6	0.13	612.4	2.87
8	50	20	10	0.2	0.7	0.8	10	5	2	536.3	535.3	0.17	527.5	1.64
9	50	20	10	1	0.9	0.8	10	5	2	528.3	527.7	0.15	522.2	1.15
10	50	20	10	1	0.5	0.8	10	5	2	560.6	560.0	0.11	532.1	5.09
11	50	20	10	1	0.7	0.9	10	5	2	545.5	543.0	0.47	537.0	1.56
12	50	20	10	1	0.7	0.4	10	5	2	559.7	559.2	0.09	548.9	1.93
13	50	20	10	1	0.7	0.8	20	5	2	544.5	542.8	0.31	535.7	1.61
14	50	20	10	1	0.7	0.8	5	5	2	560.3	559.7	0.11	552.1	1.47
15	50	20	10	1	0.7	0.8	10	10	2	545.1	544.5	0.11	539.6	1.01
16	50	20	10	1	0.7	0.8	10	2	2	562.9	561.5	0.26	554.8	1.45
17	50	20	10	1	0.7	0.8	10	5	4	545.0	543.6	0.26	526.5	3.39
18	50	20	10	1	0.7	0.8	10	5	1	557.2	556.4	0.14	550.3	1.24

5 for revenues, $\lambda_0 = 10, \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 4$ for demand arrival rates, $\rho_0 = 0.7, \rho_1 = 0.8, \rho_2 = 0.9, \rho_3 = 0.5, \rho_4 = 0.6$ for utilizations, and $b_0 = 0.2R_0, h_0 = 0.1R_0, R_i = 0.1 \forall i$ for backorder and holding cost parameters. In the 18 instances to follow, a parameter corresponding to the end-product and one of the intermediate products (product type-1) is either increased or decreased. For this experiment, we

keep the parameters for the remaining intermediate products unchanged and hence omit representing their values in Table 3.3. In example 1, the revenue from the end-product is much higher than the sum of the revenues from intermediate products. In example 2, on the other hand, the end-product revenue is lower. Similarly, examples 3 and 4 depict instances when the revenue from the intermediate product type-1 is high and low. Examples 5-8 correspond to high and low demand arrival rates for the end-product and the intermediate product. Examples 9-12 explore the effects of high and low assembly and production utilizations. Finally, in examples 13-15 we change the backorder and holding costs for the end-product and the intermediate product type-1.

In Table 3.3, we observe that the difference between the profits obtained by the optimal and the heuristic policy is 0.27% whereas the difference between the profits attained by the optimal and the independent base-stock/rationing policy is 2.02%. Hence, we find that the heuristic policy maintains its performance when there are differences in the rate, revenue, and cost parameters of various intermediate products. In addition, as was the case in Tables 3.1 and 3.2, the heursitic policy performs better than the independent base-stock/rationing policy in every problem instance. Through the experiment in Table 3.3, we again observe that the performance of the heuristic policy improves when the end-product revenue and demand rate is high and when production line utilization is low. The heuristic also performed better when assembly backorder cost was low, end-product holding cost was high, and intermediate product holding cost was low. The heuristic has been robust with respect to changes in the revenue and demand rate of the intermediate component as well as the assembly utilization.

We would like to end this section with a note on systems with larger number of components. The computational requirements for the heuristic policy grows linearly with the number of components (a total of $N+1$ subproblems need to be constructed and solved optimally for a system with N components). However, as discussed during the motivation for the development of a heuristic procedure, computing the optimal policies for large systems is computationally impractical since the number of states that the system as a whole may be in grows exponentially with N. Although we can determine the parameters of the heuristic policy for a large system with ease, computational limitations for evaluating the profits obtained by the optimal and heuristic policy prevent us from exploring its performance in such systems. However, as the results for Table 3.1 and 3.2 imply, we expect the heuristic to maintain a highly satisfactory level of performance for systems with a moderately large number of components. Therefore, due to its performance in the problems tested, ease of implementation, and requirement of only a manageable number of subproblems, we believe this heuristic policy would be very effective and beneficial for the control of such systems in practice.

3.9 Conclusions

In this chapter, we studied an assembly system where there is demand for both the end-product and intermediate products. For a general system composed of an arbitrary number of components, we showed that demand admission for the product and for any of the intermediate products are characterized by state-dependent rationing and admission threshold levels while both component production and product assembly follow state-dependent base-stock levels. We explored the sensitivity properties of the optimal policy to various problem parameters.

In addition, we provided two extensions for the basic model, one concerning with multiple customer classes based on revenue in addition to the classes based on the type of item they request, and the other, investigating the effects of a partial payment scheme on the optimal policy structure. Since the optimal policies were rather complex and defined by switching surfaces in a multidimensional space, we also introduced a heuristic policy that performed well under a variety of example problems.

Extensions that allow customer demands for a selection of components may constitute interesting and challenging problems for future research. The scope of this paper may be regarded as a special case of the general assemble-to-order problem for which assembly is done from a selection of components chosen by a customer. The optimal policies have not been fully characterized for such systems and extensions to this research may provide valuable additional results and insights applicable to these problems.

3.10 Appendix

3.10.1 Proofs of Optimal Policy Structure

Proof of Lemma 3.1

(**iv**) $D_1 v(\mathbf{x}, y) \downarrow x_i$: $D_{\mathbf{1}} v(\mathbf{x}+\mathbf{e}_i, y) - D_{\mathbf{1}} v(\mathbf{x}, y) = v(\mathbf{x}+\mathbf{e}_i+\mathbf{1}, y) - v(\mathbf{x}+\mathbf{e}_i, y) - v(\mathbf{x}+\mathbf{1}, y) + v(\mathbf{x}, y) \leq v(\mathbf{x}+\mathbf{e}_i+\mathbf{1}, y) - v(\mathbf{x}+\mathbf{e}_i, y+1) - v(\mathbf{x}+\mathbf{1}, y) + v(\mathbf{x}, y+1) = -D_{\mathbf{-1},p} v(\mathbf{x}+\mathbf{e}_i, y) + D_{\mathbf{-1},p} v(\mathbf{x}, y) \leq 0$ where the first and second inequalities follow from (i) and (iii), respectively.

$D_1 v(\mathbf{x}, y) \downarrow y$: By expanding and regrouping the terms we get $D_{\mathbf{1}} v(\mathbf{x}, y+1) - D_{\mathbf{1}} v(\mathbf{x}, y) = D_p v(\mathbf{x}+\mathbf{1}, y) - D_p v(\mathbf{x}, y) \leq 0$ where the inequality follows from successive applications of $D_p \downarrow x_i$ in (ii).

(**v**) $D_{i,j}v(\mathbf{x}, y) \downarrow x_i$ ($\downarrow x_j$ similar): By (iv), we have $D_{\mathbf{1}}v(\mathbf{x}, y) \downarrow x_i$, i.e., $v(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2+ ... + \mathbf{e}_N, y) - v(\mathbf{x}, y) \downarrow x_i$. Adding and subtracting the term $v(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j, y)$, we get $[v(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2+...+\mathbf{e}_N, y)-v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y)]+[v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y)-v(\mathbf{x}, y)] \downarrow x_i$. Since, the term in the first brackets is a combination of difference operators for having one more unit of a component k, where $k = 1, ..., N$, $k \neq i, j$, this term is increasing with respect to x_i by (i). Therefore, the term in the second brackets must be decreasing with respect to x_i, i.e., $D_{i,j}v(\mathbf{x}, y) \downarrow x_i$.

$D_{i,j}v(\mathbf{x}, y) \uparrow x_k$ and $\downarrow y$: Follows immediately from (i).

(**vi**): $D_{-i,p}v(\mathbf{x}, y) \uparrow x_i$: By (iii), we have $D_{\mathbf{-1},p}v(\mathbf{x}, y) \uparrow x_i$, i.e. $v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2 + ... + \mathbf{e}_N, y) \uparrow x_i$. Adding and subtracting the term $v(\mathbf{x} + \mathbf{e}_i, y)$, we get $[v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{e}_i, y)] + [v(\mathbf{x} + \mathbf{e}_i, y) - v(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2 + ... + \mathbf{e}_N, y)] \uparrow x_i$. The term in the second bracket is decreasing with respect to x_i due to (i). Therefore, $D_{-i,p}v(\mathbf{x}, y) \ \uparrow x_i$.

$D_{-i,p}v(\mathbf{x}, y) \downarrow x_j$: $D_{-i,p}v(\mathbf{x}, y) = v(\mathbf{x}, y+1)-v(\mathbf{x}+\mathbf{e}_i, y)$. Adding and subtracting the term $v(\mathbf{x}, y)$, we get $D_pv(\mathbf{x}, y) - D_iv(\mathbf{x}, y)$. The first term is $\downarrow x_j$ due to (ii) and the second term, excluding the negative sign, is $\uparrow x_j$ due to (i), thus $D_{-i,p}v(\mathbf{x}, y) \downarrow x_j$.

$D_{-i,p}v(\mathbf{x}, y) \downarrow y$: By (iii), we have, $D_{-1,p}v(\mathbf{x}, y) \downarrow y$, that is $v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2 + ... + \mathbf{e}_N, y) \downarrow y$. Adding and subtracting the term $v(\mathbf{x} + \mathbf{e}_i, y)$, we get $[v(\mathbf{x}, y + 1) - v(\mathbf{x} + \mathbf{e}_i, y)] + [v(\mathbf{x} + \mathbf{e}_i, y) - v(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2 + ... + \mathbf{e}_N, y)]$ By (i), the term in the second bracket is increasing with respect to y. Therefore, the term in the second bracket must be decreasing with respect to y. Hence, $D_{-i,p}v(\mathbf{x}, y) \ \downarrow y$.

(**vii**): $D_{i,-1,P}v(\mathbf{x}, y) \downarrow x_i$: $D_{i,-1,P}v(\mathbf{x}, y) = D_{i,-1}v(\mathbf{x}, y + 1) + D_pv(\mathbf{x}, y)$. Since the second term is $\downarrow x_i$ by (ii), we only need to show that the first term is $\downarrow x_i$. $D_{i,-1}v(\mathbf{x}+\mathbf{e}_i, y + 1) - D_{i,-1}v(\mathbf{x}, y + 1) = D_iv(\mathbf{x} + \mathbf{e}_i - \mathbf{1}, y + 1) - D_iv(\mathbf{x}, y + 1) \leq D_iv(\mathbf{x} + \mathbf{e}_i +$

$\mathbf{e}_j - \mathbf{1}, y+1) - D_i v(\mathbf{x}, y+1) \leq D_i v(\mathbf{x} + \mathbf{1} - \mathbf{1}, y+1) - D_i v(\mathbf{x}, y+1) = 0$ where the inequalities are due to successive applications of $D_i \uparrow x_j$ in (i).

$D_{i,-1,p}v(\mathbf{x}, y) \uparrow x_j$ and $\downarrow y$: $D_{i,-1,p}v(\mathbf{x}, y) = D_i v(\mathbf{x} - \mathbf{1}, y+1) + D_{-1,p}v(\mathbf{x}, y)$. The terms are $\uparrow x_j$ and $\downarrow y$ by (i) and (iii), respectively. □

Proof of Lemma 3.2

For controlled queueing systems, Koole [34] develops a framework on the preservation of the properties of dynamic programming value functions and gives examples for a rich set of properties and operators that are frequently encountered in the analysis of production systems including single server and tandem settings (see for example, Definition 5.2 and Theorem 7.2 in [34]). In our model, the arbitrary number of production lines feeding a shared assembly operation and admission decisions on both the component and the assembly stage prevents us from using his results directly in most cases.

Further, the recent work by Benjaafar et al [5] also establishes preservation results for an assembly system with multiple stages. Although our model also requires similar sub- and supermodularity conditions, it also necessitates us to show that they are preserved under different operators. Our analysis follows the framework given by Ha [31] and Carr and Duenyas [9].

For brevity, we only present the proof for a supermodularity condition given in (i). The proofs that the other conditions are also preserved under the operators are similar and therefore omitted. Specifically, we will show that if v satisfies $D_i v(\mathbf{x}, y) \uparrow x_j$, then $T_0^1 v$, $T_i^1 v$, $T_j^1 v$, $T_k^1 v$, $T_0^2 v$, $T_i^2 v$, $T_j^2 v$, $T_k^2 v$, and Tv (where i, j, k are distinct) all satisfy the same condition.

We start by showing that $D_i T_0^1 v(\mathbf{x}, y) \uparrow x_j$. For $y > 0$, as we satisfy the demand from available inventory, $D_iT_0^1v(\mathbf{x}+\mathbf{e}_j,y) - D_iT_0^1v(\mathbf{x},y) = D_iv(\mathbf{x}+\mathbf{e}_j,y-1) - D_iv(\mathbf{x},y-1) \geq 0$ by $D_iv \uparrow x_j$.

For $y \leq 0$, we have

$$D_iT_0^1v(\mathbf{x}+\mathbf{e}_j,y) = \max[v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y-1), v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y)] - \max[v(\mathbf{x}+\mathbf{e}_j,y-1), v(\mathbf{x}+\mathbf{e}_j,y)] \tag{3.3}$$

$$D_iT_0^1v(\mathbf{x},y) = \max[v(\mathbf{x}+\mathbf{e}_i,y-1), v(\mathbf{x}+\mathbf{e}_i,y)] - \max[v(\mathbf{x},y-1), v(\mathbf{x},y)] \tag{3.4}$$

We need to show that (3.3) minus (3.4) ≥ 0.

The outcome $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y) - v(\mathbf{x}+\mathbf{e}_j,y-1)$ is not feasible for (3.3) due to $D_p \downarrow x_i$ in (ii). The three remaining feasible outcomes are $D_iv(\mathbf{x}+\mathbf{e}_j,y-1)$, $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y-1) - v(\mathbf{x}+\mathbf{e}_j,y)$, and $D_iv(\mathbf{x}+\mathbf{e}_j,y)$.

(a) Suppose $D_iv(\mathbf{x}+\mathbf{e}_j,y-1)$ is the outcome for (3.3). We have three possibilities for the outcome of (3.4). If $D_iv(\mathbf{x},y-1)$ is the result of (3.4), then (3.3) minus (3.4) yields $D_iv(\mathbf{x}+\mathbf{e}_j,y-1) - D_iv(\mathbf{x},y-1)$ which is ≥ 0 by $D_iv \uparrow x_j$. If, on the other hand, the outcome of (3.4) is $v(\mathbf{x}+\mathbf{e}_i,y-1) - v(\mathbf{x},y)$, then (3.3) minus (3.4) becomes $D_iv(\mathbf{x}+\mathbf{e}_j,y-1) - v(\mathbf{x}+\mathbf{e}_i,y-1) + v(\mathbf{x},y) \geq D_iv(\mathbf{x}+\mathbf{e}_j,y-1) - D_iv(\mathbf{x},y-1) \geq 0$ where the first inequality follows from the case requirement that $v(\mathbf{x},y) \geq v(\mathbf{x},y-1)$. Finally, if $D_iv(\mathbf{x}+\mathbf{e}_j,y)$ is the outcome for (3.4), then (3.3) minus (3.4) results in $D_{-j,P}v((\mathbf{x},y) - D_{-j,P}v((\mathbf{x}+\mathbf{e}_i,y)$ and that is ≥ 0 since $D_{-j,P} \downarrow x_i$ by (vi).

(b) Suppose $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y-1) - v(\mathbf{x}+\mathbf{e}_j,y)$ is the outcome for (3.3). Then (3.3) minus (3.4) either yields $D_{-i,P}v((\mathbf{x},y) - D_{-i,P}v((\mathbf{x}+\mathbf{e}_j,y)$ which is ≥ 0 by (vi), or yields $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y-1) - v(\mathbf{x}+\mathbf{e}_j,y) - D_iv(\mathbf{x},y) \geq D_iv(\mathbf{x}+\mathbf{e}_j,y) - D_iv(\mathbf{x},y) \geq 0$ where the first inequality is due the underlying case assumption for the outcomes and the second inequality follows from $D_iv \uparrow x_j$ in (i).

(c) Finally, if the outcome of (3.3) is $D_i v(\mathbf{x}+\mathbf{e}_j, y)$, then (3.3) minus (3.4) reduces to $D_i v(\mathbf{x}+\mathbf{e}_j, y) - D_i v(\mathbf{x}, y) \geq 0$.

Next, we show $D_i T_i^1 v(\mathbf{x}, y) \uparrow x_j$ and first consider the states away from the boundary, i.e., $x_i > 0$.

$$D_i T_1^1 v(\mathbf{x}+\mathbf{e}_j, y) = \max[v(\mathbf{x}+\mathbf{e}_j, y) + R_i, v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y)] - \max[v(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j, y) + R_i, v(\mathbf{x}+\mathbf{e}_j, y)] \tag{3.5}$$

$$D_i T_1^1 v(\mathbf{x}, y) = \max[v(\mathbf{x}, y) + R_i, v(\mathbf{x}+\mathbf{e}_i, y)] - \max[v(\mathbf{x}-\mathbf{e}_i, y) + R_i, v(\mathbf{x}, y)] \tag{3.6}$$

We again need to show that (3.5) minus (3.6) ≥ 0. Eliminating the infeasible outcomes due to conditions (i) and (v), we have the following possible cases:

(a) If the outcome for (3.5) is $D_i v(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j, y)$, the only feasible outcome for (3.6) results in (3.5) minus (3.6) to equal $D_i v(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j, y) - D_i v(\mathbf{x}-\mathbf{e}_i, y)$ and that is ≥ 0 by (i).

(b) Suppose the outcome for (3.5) is $v(\mathbf{x}+\mathbf{e}_j, y) + R_i - v(\mathbf{x}+\mathbf{e}_j, y)$. Then, the result for (3.5) minus (3.6) is either zero or $R_i - v(\mathbf{x}, y) + v(\mathbf{x}-\mathbf{e}_i, y) \geq 0$ by the case assumption.

(c) Suppose now that the outcome for (3.5) is $D_i v(\mathbf{x}+\mathbf{e}_j, y)$. Then, (3.5) minus (3.6) either becomes $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y) - v(\mathbf{x}+\mathbf{e}_j, y) - R_i \geq 0$ by the case assumption or becomes $D_i v(\mathbf{x}\mathbf{e}_j, y) - D_i v(\mathbf{x}, y)$ and that is ≥ 0 by (i).

At the boundary states where $x_i = 0$, the outcome analyzed in part (a) also becomes infeasible and the resulting cases are identical to the ones in parts (b) and (c) with x_i set to zero. The analysis for operators $D_i T_j^1$ and $D_i T_k^1$ where $k \neq i, j$ are very similar and thus omitted for space.

Next, we consider the operator T_0^2 and show that $D_i T_0^2 v(\mathbf{x}, y) \uparrow x_j$. For $x_i > 0 \; \forall i$, we have

$$D_iT_0^2v(\mathbf{x}+\mathbf{e}_j,y)=\max[v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1},y+1),v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j,y)] \\ -\max[v(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y+1),v(\mathbf{x}+\mathbf{e}_j,y)] \tag{3.7}$$

$$D_iT_1^1v(\mathbf{x},y)=\max[v(\mathbf{x}+\mathbf{e}_i-\mathbf{1},y+1),v(\mathbf{x}+\mathbf{e}_i,y)]-\max[v(\mathbf{x}-\mathbf{1},y+1),v(\mathbf{x},y)] \tag{3.8}$$

Eliminating the infeasible outcomes due to (iii), we show that for each of the remaining cases we have (3.7) minus (3.8) ≥ 0.

(a) If (3.7) results in $D_iv(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y+1)$, then (3.7) minus (3.8) may result in three possible expressions. $D_iv(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y+1)-D_iv(\mathbf{x}-\mathbf{1},y+1)\geq 0$ by (i), $D_iv(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y+1)-v(\mathbf{x}+\mathbf{e}_i-\mathbf{1},y+1)+v(\mathbf{x},y)\geq D_iv(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y+1)-D_iv(\mathbf{x}-\mathbf{1},y+1)\geq 0$ by (i), and $D_{j,-1,P}v(\mathbf{x}+\mathbf{e}_i-\mathbf{1},y)-D_{j,-1,P}v(\mathbf{x}-\mathbf{1},y)\geq 0$ by (vii).

(b) If the outcome of (3.7) is $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1},y+1)-v(\mathbf{x}+\mathbf{e}_j,y)$, then (3.7) minus (3.8) is either $D_{i,-1,P}v(\mathbf{x}+\mathbf{e}_j-\mathbf{1},y)-D_{i,-1,P}v(\mathbf{x}-\mathbf{1},y)\geq 0$ by (vii) or $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1},y+1)-v(\mathbf{x}+\mathbf{e}_j,y)-D_iv(\mathbf{x},y)\geq D_iv(\mathbf{x}+\mathbf{e}_j,y)-D_iv(\mathbf{x},y)\geq 0$ where the first inequality is due to the case assumption and the second inequality follows from (i).

(c) Finally, if (3.7) results in $D_iv(\mathbf{x}+\mathbf{e}_j,y+1)$, the only possible outcome for (3.7) minus (3.8) is $D_iv(\mathbf{x}+\mathbf{e}_j,y)-D_iv(\mathbf{x},y)\geq 0$ by (i).

For boundary states where $x_i=0$ or $x_j=0$, the outcome discussed in part (a) becomes infeasible and only the cases in parts (b) and (c) apply with identical reasoning. For boundary states with $x_k=0$, cases analyzed in parts (a) and (b) become infeasible and only part (c) applies.

The analysis for operators $D_iT_i^2$, $D_iT_j^2$ and $D_iT_k^2$ are very similar to the ones for operators $D_iT_i^1$, $D_iT_j^1$ and $D_iT_k^1$. Hence, they are omitted for brevity.

Finally, we consider the operator T. By definition, T is formed by (a) the addition and multiplication of positive constants with the functions $T_n^1 v$ and $T_n^2 v$ for $n =$ 0,...,N that are shown to be $\uparrow x_j$ and (b) linear inventory holding and assembly queue backorder costs. Therefore, $D_i T v \uparrow x_j$ as well. $\square$

Proof of Theorem 3.1

Consider a value iteration algorithm to solve the optimal policy for the problem given in (4.1) where initial values $v_0(\mathbf{x}, y) = 0$ are used for every state $(\mathbf{x}, y)$. Conditions (i)-(iii) are trivially satisfied by $v_0(\mathbf{x}, y)$, hence $v_0(\mathbf{x}, y) \in V$. We apply $v_{k+1}(\mathbf{x}, y) = T v_k(\mathbf{x}, y)$ for $k = 0, 1, 2, ...$ to determine the relative value functions for successive iterations. Suppose now that the value functions in iteration k satisfy (i)-(iii), i.e. $v_k(\mathbf{x}, y) \in V$. Then, Lemma 3.1 shows that $v_{k+1}(\mathbf{x}, y)$ also satisfy (i)-(iii). Therefore $v_{k+1}(\mathbf{x}, y) \in V$.

We note that, without loss of optimality, we can add the following constraints to the original problem that we cannot admit a product demand when $R_0 < b_0\ y^-/\Lambda$, we cannot produce a component type-i when $h_i\ x_i/\Lambda > \max(R_i, (b_0 + \sum_{i=1}^{N} h_i/\Lambda))$, and we cannot assemble another end product when $h_0\ y^+/\Lambda > \max(R_0, ((\sum_{i=1}^{N} h_i) - h_0)/\Lambda)$. For example, if $b_0\ y^-/\Lambda > R_0$, this suggests that the amount of backorder cost incurred until the next transition is greater than any potential revenue of R_0 that would be received if the next event were a product demand arrival. (As another example, if $h_i\ x_i/\Lambda > \max(R_i, (b_0 + \sum_{i=1}^{N} h_i/\Lambda))$, this indicates that the amount of holding cost due to a type-i component incurred during a transition epoch is greater than the potential benefits of (a) selling that component for a revenue of R_i were the next event a demand arrival for component i, and (b) assembling another unit of a backordered product that would save backorder and holding costs for a transition epoch if the next event were a product assembly opportunity.) Thus, the original problem can be converted to a finite state, finite action set problem. The underlying Markov chain is also unichain. Thus, Theorem 8.4.5 of Puterman [46] ensures the existence of a long-run average profit and the validity of the value iteration algorithm to determine it.

To complete the proof of Theorem 3.1, we note that conditions (i)-(iii) are sufficient to demonstrate the structural properties of the optimal policy. Due to (i), if it is optimal not to produce component i in state $(\mathbf{x}, y)$, it remains optimal not to do so in state $(\mathbf{x} + \mathbf{e}_i, y)$, implying a base-stock production policy. Further, the sub- and super-modularity conditions imply that the base-stock level is nondecreasing with the inventory of other components and nonincreasing with the end-product inventory. Condition (i) also implies that if it is optimal to accept a demand for component i in state $v(\mathbf{x}, y)$, i.e., $R_i \geq v(\mathbf{x}, y) - v(\mathbf{x} - \mathbf{e}_i, y)$, it is also optimal to accept a demand for component i in state $v(\mathbf{x}+\mathbf{e}_i, y)$, i.e. $R_i \geq v(\mathbf{x}+\mathbf{e}_i, y) - v(\mathbf{x}, y)$. Thus component demand admission follows a rationing policy. Similarly, the sub- and super-modularity conditions imply that the rationing level for a component is nondecreasing with the inventory of other products and nonincreasing with the end-product inventory position.

Condition (ii) indicates that if $v(\mathbf{x}, y-1) + R_0 \geq v(\mathbf{x}, y)$, then $v(\mathbf{x}, y) + R_0 \geq v(\mathbf{x}, y+1)$, thus implying an admission threshold for the end-product. Moreover, $D_{-P} \uparrow x_i$ implies that the state dependent admission threshold is nondecreasing with the amount of component inventories. Finally, condition (iii) implies the structure of the optimal product assembly policy. $D_{\mathbf{-1},p} \downarrow y$ means that the additional value gained by assembling an end-product gets smaller with each additional unit of end-product in the inventory, implying that product assembly follows a state dependent threshold structure. $D_{\mathbf{-1},p} \uparrow x_i$ suggests that the assembly threshold level is non-decreasing with the amount of component inventories. □

3.10.2 Proofs of Sensitivity Results

Proof of Theorem 3.2

We construct two systems that are identical in all problem parameters except the end-product revenues which are chosen such that $R_0' < R_0$. We refer to the original problem where the end product revenue is R_0 as problem A, and the modified setting with R_0' as problem B. We initialize problems A and B with $v_0(\mathbf{x}, y) = 0$ and $v_0'(\mathbf{x}, y) = 0$, respectively. We then apply $v_{k+1}(\mathbf{x}, y) = Tv_k(\mathbf{x}, y)$ and $v_{k+1}'(\mathbf{x}, y) = Tv_k'(\mathbf{x}, y)$. By Lemma 2, $v_k(\mathbf{x}, y)$ and $v_k'(\mathbf{x}, y)$ satisfy conditions (i)-(iii) $\forall$ k. Hence both v_k and $v_k' \in V$. We first prove the following lemma.

Lemma 3.3. *Let v_k and $v_k' \in V$ for $\forall$ k. For each state $(\boldsymbol{x}, y)$ and for every $k =$ $0, 1, 2, \ldots,$ the following conditions jointly hold:*

(a) $D_i v_k'(\mathbf{x}, y) - D_i v_k(\mathbf{x}, y) \leq 0$

(b) $D_p v_k'(\mathbf{x}, y) - D_p v_k(\mathbf{x}, y) + R_0 - R_0' \geq 0$

(c) $D_{\mathbf{-1},p} v_k'(\mathbf{x}, y) - D_{\mathbf{-1},p} v_k(\mathbf{x}, y) \leq 0$

Proof: The proof of Lemma 3.3 is by induction. Conditions (a)-(c) hold trivially for $v_0(\mathbf{x}, y)$ and $v_0'(\mathbf{x}, y)$. We assume the conditions hold for iteration k and show that they are preserved in iteration $k + 1$. For brevity, we only present the proof that $D_p v_k'(\mathbf{x}, y) - D_p v_k(\mathbf{x}, y) + R_0 - R_0' \geq 0$. The proofs of other conditions are similar and hence omitted. Since $v_{k+1}(\mathbf{x}, y) = Tv_k(\mathbf{x}, y)$ and $v_{k+1}'(\mathbf{x}, y) = Tv_k'(\mathbf{x}, y)$, we have

$$D_p v_{k+1}'(\mathbf{x}, y) - D_p v_{k+1}(\mathbf{x}, y) = D_p T v_k'(\mathbf{x}, y) - D_p T v_k(\mathbf{x}, y)$$

As in the proof of Lemma 3.2, we start by showing $D_pT_0^1v_k'(\mathbf{x},y) - D_pT_0^1v_k(\mathbf{x},y) + R_0 - R_0' \geq 0$ and proceed to show that the condition holds for the remaining operators T_i^1, T_0^2, T_i^2, and T.

$$\begin{aligned} &D_pT_0^1v_k'(\mathbf{x},y) - D_pT_0^1v_k(\mathbf{x},y) + R_0 - R_0' \\ &\qquad = \max[v_k'(\mathbf{x},y) + R_0', v_k'(\mathbf{x},y+1)] - \max[v_k'(\mathbf{x},y-1) + R_0', v_k'(\mathbf{x},y)] \\ &\qquad\quad - \max[v_k(\mathbf{x},y) + R_0, v_k(\mathbf{x},y+1)] + \max[v_k(\mathbf{x},y-1) + R_0, v_k(\mathbf{x},y)] \\ &\qquad\quad + R_0 - R_0' \end{aligned} \tag{3.9}$$

In order to simplify the analysis, we adapt a similar notation used in Carr and Duenyas [9] and introduce two functions w and w' on $\{0,1\}$ x S. Let w be defined as

$$w(u,\mathbf{x},y) = \begin{cases} v_k(\mathbf{x},y-1) + R_0 & \text{if } u = 1 \\ v_k(\mathbf{x},y) & \text{if } u = 0 \end{cases}$$

and w' be defined similarly with the corresponding value function v_k' and product revenue R_0'. Therefore $T_0^1v_k(\mathbf{x},y) = \max_{u\in 0,1} w(u,\mathbf{x},y)$ and $T_0^1v_k'(\mathbf{x},y)$=$\max_{u\in 0,1} w'(u,\mathbf{x},y)$. We let $u_{(\mathbf{x},y)} = \text{argmax}_u w(u,\mathbf{x},y)$ and $u'_{(\mathbf{x},y)} = \text{argmax}_u w'(u,\mathbf{x},y)$.

By condition (ii) ($D_p \downarrow y$), we have $u'_{(\mathbf{x},y+1)} \geq u'_{(\mathbf{x},y)}$ and $u_{(\mathbf{x},y+1)} \geq u_{(\mathbf{x},y)}$. By condition (b), we further have $u_{(\mathbf{x},y)} \geq u'_{(\mathbf{x},y)}$ and $u_{(\mathbf{x},y+1)} \geq u'_{(\mathbf{x},y+1)}$. Hence the vector ($u'_{(\mathbf{x},y+1)}$, $u'_{(\mathbf{x},y)}$, $u_{(\mathbf{x},y+1)}$, $u_{(\mathbf{x},y)}$) has the following six possible values: $(0,0,0,0)$, $(0,0,1,0)$, $(1,0,1,0)$, $(0,0,1,1)$, $(1,1,0,1)$, and $(1,1,1,1)$. We show that (3.9) analyzed for each of these six cases is ≥ 0.

(0,0,0,0): (3.9) equals $D_pv_k'(\mathbf{x},y) - D_pv_k(\mathbf{x},y) + R_0 - R_0'$ and that is ≥ 0 by condition (b).

(0,0,1,0): yields $D_pv_k'(\mathbf{x},y) - R_0'$ which is ≥ 0 since $u'_{(\mathbf{x},y+1)} = 0$ implies $v_k'(\mathbf{x},y) + R_0' \leq v_k'(\mathbf{x},y+1)$.

(1,0,1,0): (3.9) equals 0.

(0,0,1,1): results in $D_p v'_k(\mathbf{x}, y) - D_p v_k(\mathbf{x}, y-1) + R_0 - R'_0 \geq 0$ by the case assumptions that $u'_{(\mathbf{x},y+1)} = 0$ and $u_{(\mathbf{x},y)} = 1$.

(1,1,0,1): (3.9) $\geq D_p v'_k(\mathbf{x}, y-1) - D_p v_k(\mathbf{x}, y) + R_0 - R'_0 \geq D_p v'_k(\mathbf{x}, y) - D_p v_k(\mathbf{x}, y) + R_0 - R'_0 \geq 0$ where the first inequality follows due to the case assumption $u_{(\mathbf{x},y)} = 1$ implying $v_k(\mathbf{x}, y-1) + R_0 > v_k(\mathbf{x}, y)$ and the second inequality follows from (b).

(1,1,1,1): yields $D_p v'_k(\mathbf{x}, y-1) - D_p v_k(\mathbf{x}, y-1) + R_0 - R'_0$ and that is ≥ 0 by condition (b).

Next, we show the result for operator T_i^1 and first consider the states away from boundary, i.e., $x_i > 0$.

$$
\begin{aligned}
D_p T_i^1 v'_k(\mathbf{x}, y) &- D_p T_i^1 v'_k(\mathbf{x}, y) + R_0 - R'_0 \\
&= \max[v'_k(\mathbf{x} - \mathbf{e}_i, y+1) + R_i, v'_k(\mathbf{x}, y+1)] - \max[v'_k(\mathbf{x} - \mathbf{e}_i, y) + R_i, v'_k(\mathbf{x}, y)] \\
&\quad - \max[v_k(\mathbf{x} - \mathbf{e}_i, y+1) + R_i, v_k(\mathbf{x}, y+1)] \\
&\quad + \max[v_k(\mathbf{x} - \mathbf{e}_i, y) + R_i, v_k(\mathbf{x}, y)] + R_0 - R'_0 \qquad (3.10)
\end{aligned}
$$

We redefine w and w' as

$$
w(u, \mathbf{x}, y) = \begin{cases} v_k(\mathbf{x} - \mathbf{e}_i, y) + R_i, & \text{if } u = 1 \\ v_k(\mathbf{x}, y) & \text{if } u = 0 \end{cases}
$$

with w' again defined similar to w, using its corresponding value function v'_k. Using the definitions of u and u' with the new functions w and w', we have $T_i^1 v_k(\mathbf{x}, y) = \max_{u \in 0,1} w(u, \mathbf{x}, y)$ and $T_i^1 v'_k(\mathbf{x}, y) = \max_{u \in 0,1} w'(u, \mathbf{x}, y)$.

Condition (i) ($D_i \downarrow y$), implies $u_{(\mathbf{x},y+1)} \geq u_{(\mathbf{x},y)}$ and $u'_{(\mathbf{x},y+1)} \geq u'_{(\mathbf{x},y)}$. Further, by condition (a), we have $u'_{(\mathbf{x},y)} \geq u_{(\mathbf{x},y)}$ and $u'_{(\mathbf{x},y+1)} \geq u_{(\mathbf{x},y+1)}$.

Hence the vector $(u'_{(\mathbf{x},y+1)}, u'_{(\mathbf{x},y)}, u_{(\mathbf{x},y+1)}, u_{(\mathbf{x},y)})$ now has the following six possible values: $(0,0,0,0)$, $(1,0,0,0)$, $(1,1,0,0)$, $(1,0,1,0)$, $(1,1,1,0)$, and $(1,1,1,1)$. We analyze (3.10) for each of these cases.

(0,0,0,0): yields $D_p v'_k(\mathbf{x},y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0$ and that is ≥ 0 by condition (b).

(1,0,0,0): (3.10) $\geq D_p v'_k(\mathbf{x},y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0$ and that is ≥ 0 by condition (b) where the first inequality follows from $v'_k(\mathbf{x}-\mathbf{e}_i,y) + R_i > v'_k(\mathbf{x},y)$ implied by $u'_{(\mathbf{x},y+1)} = 0$.

(1,1,0,0): results in $D_p v'_k(\mathbf{x}-\mathbf{e}_i,y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0 \geq D_p v'_k(\mathbf{x},y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0 \geq 0$ where the inequalities follow from conditions (b) and (i), respectively.

(1,0,1,0): gives (by adding and subtracting the terms $v'_k(\mathbf{x},y+1)$ and $v_k(\mathbf{x},y+1)$) $[-D_i v'_k(\mathbf{x}-\mathbf{e}_i,y) + D_i v_k(\mathbf{x}-\mathbf{e}_i,y) + [D_p v'_k(\mathbf{x},y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0]$ where the terms in the first and second brackets are ≥ 0 by conditions (a) and (b), respectively.

(1,1,1,0): (3.10) $\geq D_p v'_k(\mathbf{x}-\mathbf{e}_i,y) - D_p v_k(\mathbf{x}-\mathbf{e}_i,y) + R_0 - R'_0$ and that is ≥ 0 by condition (b) where the first inequality follows from $v_k(\mathbf{x},y) > v_k(\mathbf{x}-\mathbf{e}_i,y) + R_i$ implied by $u_{(\mathbf{x},y)} = 0$.

(1,1,1,1): yields $\geq D_p v'_k(\mathbf{x}-\mathbf{e}_i,y) - D_p v_k(\mathbf{x}-\mathbf{e}_i,y) + R_0 - R'_0$ which is ≥ 0 by (ii).

For the boundary states with $x_i = 0$, only case (0,0,0,0) applies and (3.10) results in $D_p v'_k(\mathbf{x},y) - D_p v_k(\mathbf{x},y) + R_0 - R'_0$ which is ≥ 0 by condition (b).

For Operator T_0^2, we first consider the states $(\mathbf{x}, y)$ for which $x_i > 0\ \forall i$.

$$\begin{aligned}
&D_p T_i^1 v_k'(\mathbf{x}, y) - D_p T_i^1 v_k'(\mathbf{x}, y) + R_0 - R_0' \\
&\qquad = \max[v_k'(\mathbf{x} - \mathbf{1}, y+2), v_k'(\mathbf{x}, y+1)] - \max[v_k'(\mathbf{x} - \mathbf{1}, y), v_k'(\mathbf{x}, y)] \\
&\qquad\quad - \max[v_k(\mathbf{x} - \mathbf{1}, y+2), v_k(\mathbf{x}, y+1)] + \max[v_k(\mathbf{x} - \mathbf{1}, y), v_k(\mathbf{x}, y)] \\
&\qquad\quad + R_0 - R_0' \qquad (3.11)
\end{aligned}$$

We redefine w and w' as

$$w(u, \mathbf{x}, y) = \begin{cases} v_k(\mathbf{x} - \mathbf{1}, y+1) & \text{if } u = 1 \\ v_k(\mathbf{x}, y) & \text{if } u = 0 \end{cases}$$

with w' defined similar to w. We let $T_0^2 v_k(\mathbf{x}, y) = \max_{u \in 0,1} w(u, \mathbf{x}, y)$ and $T_0^2 v_k'(\mathbf{x}, y) = \max_{u \in 0,1} w'(u, \mathbf{x}, y)$. By condition (iii) ($D_{-1,p} \downarrow y$) we have $u'_{(\mathbf{x},y)} \geq u'_{(\mathbf{x},y+1)}$ and $u_{(\mathbf{x},y)} \geq u_{(\mathbf{x},y+1)}$. Further, by (c), we also have $u_{(\mathbf{x},y)} \geq u'_{(\mathbf{x},y)}$ and $u_{(\mathbf{x},y+1)} \geq u'_{(\mathbf{x},y+1)}$. The vector $(u'_{(\mathbf{x},y+1)}, u'_{(\mathbf{x},y)}, u_{(\mathbf{x},y+1)}, u_{(\mathbf{x},y)})$ therefore has the six possible values: $(0,0,0,0)$, $(0,0,0,1)$, $(0,0,1,1)$, $(0,1,0,1)$, $(0,1,1,1)$, and $(1,1,1,1)$. We analyze (3.11) for each of these six cases. All cases except for $(0,1,0,1)$ are very similar to the ones analyzed previously, hence we only show that case $(0,1,0,1)$ yields $D_1 v_k'(\mathbf{x} - \mathbf{1}, y + 1) - D_1 v_k(\mathbf{x} - \mathbf{1}, y+1) + R_0 - R_0' \geq 0$ which is implied by jointly by conditions (b) and (c). For the states where $x_i = 0$ for some component i, the only feasible case $(0,0,0,0)$ yields $D_p v_k'(\mathbf{x}, y) - D_p v_k(\mathbf{x}, y) + R_0 - R_0' \geq 0$ by (b).

The analysis for operator T_i^2 is very similar to the one for operator T_i^1 and is thus omitted. The results hold for operator T as this operator is formed by addition and multiplication of positive constants with the functions $T_n^1 v$ and $T_n^2 v$ for $n = 0,\ldots,N$ and linear inventory holding and assembly queue backorder costs. $\square$

To complete the proof of Theorem 2, we note that conditions (a)-(c) are sufficient for the sensitivity results to hold. For example, condition (a) implies that if $v_k'(\mathbf{x} +$

$\mathbf{e}_i, y) - v'_k(\mathbf{x}, y) > 0$, then $v_k(\mathbf{x} + \mathbf{e}_i, y) - v_k(\mathbf{x}, y) > 0$. Hence, if it is optimal to produce an additional unit of component i at state $(\mathbf{x}, y)$ in problem B, then it is also optimal to produce an additional unit of component i in state $(\mathbf{x}, y)$ in problem A. Therefore $\gamma'_i(\mathbf{x}_{-i}, y) \leq \gamma_i(\mathbf{x}_{-i}, y)$. Through a similar argument, (a) also implies the shift in the component admission threshold $\alpha_i(\mathbf{x}, y)$.

In terms of the product demand admission policy, Condition (ii) implies that if $v'_k(\mathbf{x}, y-1) + R'_0 \geq v_k(\mathbf{x}, y)$, then $v_k(\mathbf{x}, y-1) + R_0 \geq v_k(\mathbf{x}, y)$. Hence, if it is optimal to admit a demand for the end product in problem B, it remains optimal to admit a product demand at that state in problem A resulting in $\beta'(\mathbf{x}) \geq \beta(\mathbf{x})$. As for the optimal assembly policies, analogous arguments yield $\delta'(\mathbf{x}) \leq \delta(\mathbf{x})$. □

3.10.3 Proofs of Extension Results

Proof of Theorem 3.3

The proof of Theorem 3.3 closely follows the steps in the proof of Theorem 3.1. Therefore, we only provide the sufficient conditions implying the structure of the optimal policy and show that they are preserved across transitions. We note that, as described in the proof of Theorem 3.1, without loss of optimality, the problem may be converted to a finite state and finite action space problem. Therefore, a stationary policy for the discounted profit infinite horizon problem exists by Theorem 6.2.10 of Puterman [46].

Let V be the set of functions defined on the state space such that if $v \in V$, then $\forall\ i, j = 1, \ldots N$ where $j \neq i$;

(i') $D_i v(\mathbf{x}, y) \ \downarrow x_i, \uparrow x_j, \downarrow y \ \ \forall\, i = 1, \ldots, N$

(ii') $D_p v(\mathbf{x}, y) \ \downarrow x_i, \downarrow y$, and $\leq R_0$ for $y > 0$

$D_p v(\mathbf{x}, 0) - D_p v(\mathbf{x}, -1) \leq R_0 - r_0$

(iii') $D_{\mathbf{-1},p} v(\mathbf{x}, y) \ \uparrow x_i, \downarrow y$

$D_{\mathbf{-1},p} v(\mathbf{x}, 0) - D_{\mathbf{-1},p} v(\mathbf{x}, -1) \leq R_0 - r_0$

Conditions (i')-(iii') are similar to the ones (i)-(iii) associated with the original problem except for the conditions given in (ii') and (iii') which enable the optimal policy to hold at the boundary states as well. We first prove the following lemma.

Lemma 3.4. *If $v \in V$ then, $T_n^1 v$, $T_n^2 v$, and $Tv \in V \quad \forall n = 0, \ldots, N$.*

Proof: For brevity, once again we only present the proof for the supermodularity condition given in (i'). We will show that if v satisfies $D_i v(\mathbf{x}, y) \uparrow x_j$, then $T_0^1 v$, $T_i^1 v$, $T_j^1 v$, $T_k^1 v$, $T_0^2 v$, $T_i^2 v$, $T_j^2 v$, $T_k^2 v$, and Tv (where i, j, k are distinct) all satisfy the same condition. As the operators T_0^1 and T_0^2 are the ones that have been modified and the analysis for the remaining operators are similar to the ones in the proof of Lemma 3.1, we restrict our illustration of the proof for these two operators.

For operator T_0^1 and for $y > 0$, we have $D_i T_0^1 v(\mathbf{x}+\mathbf{e}_j, y) - D_i T_0^1 v(\mathbf{x}, y) = D_i v(\mathbf{x}+\mathbf{e}_j, y-1) - D_i v(\mathbf{x}, y-1) \geq 0$ by $D_i v \uparrow x_j$. For $y \leq 0$, we have

$$\begin{aligned} D_i T_0^1 v(\mathbf{x}+\mathbf{e}_j, y) &= \max[v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y-1) + r_0, v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y)] \\ &\quad - \max[v(\mathbf{x}+\mathbf{e}_j, y-1) + r_0, v(\mathbf{x}+\mathbf{e}_j, y)] \end{aligned} \tag{3.12}$$

$$\begin{aligned} D_i T_0^1 v(\mathbf{x}, y) &= \max[v(\mathbf{x}+\mathbf{e}_i, y-1) + r_0, v(\mathbf{x}+\mathbf{e}_i, y)] \\ &\quad - \max[v(\mathbf{x}, y-1) + r_0, v(\mathbf{x}, y)] \end{aligned} \tag{3.13}$$

We need to show that (3.12) minus (3.13) ≥ 0. The outcome $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y) - v(\mathbf{x}+\mathbf{e}_j, y-1) - r_0$ is not feasible for (3.12) due to $D_p \downarrow x_i$ in (ii'). The three remaining feasible outcomes are $D_iv(\mathbf{x}+\mathbf{e}_j, y-1)$, $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y-1) - v(\mathbf{x}+\mathbf{e}_j, y)$, and $D_iv(\mathbf{x}+\mathbf{e}_j, y)$.

(a) Suppose $D_iv(\mathbf{x}+\mathbf{e}_j, y-1)$ is the outcome for (3.12). We have three possibilities for the outcome of (3.13). If $D_iv(\mathbf{x}, y-1)$ is the result of (3.13), then (3.12) minus (3.13) yields $D_iv(\mathbf{x}+\mathbf{e}_j, y-1) - D_iv(\mathbf{x}, y-1)$ which is ≥ 0 by $D_iv \uparrow x_j$. If, on the other hand, the outcome of (3.13) is $v(\mathbf{x}+\mathbf{e}_i, y-1) + r_0 - v(\mathbf{x}, y)$, then (3.12) minus (3.13) becomes $D_iv(\mathbf{x}+\mathbf{e}_j, y-1) - v(\mathbf{x}+\mathbf{e}_i, y-1) - r_0 + v(\mathbf{x}, y) \geq D_iv(\mathbf{x}+\mathbf{e}_j, y-1) - D_iv(\mathbf{x}, y-1) \geq 0$ where the first inequality follows from the case requirement that $v(\mathbf{x}, y) \geq v(\mathbf{x}, y-1) + r_0$. Finally, if $D_iv(\mathbf{x}+\mathbf{e}_j, y)$ is the outcome for (3.13), then (3.12) minus (3.13) results in $D_{-j,P}v((\mathbf{x}, y) - D_{-j,P}v((\mathbf{x}+\mathbf{e}_i, y)$ and that is ≥ 0 (see Lemma 3.1 for the derivation of a similar condition (iv).)

(b) Suppose $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y-1) + r_0 - v(\mathbf{x}+\mathbf{e}_j, y)$ is the outcome for (3.12). Then (3.12) minus (3.13) either yields $D_{-i,P}v((\mathbf{x}, y) - D_{-i,P}v((\mathbf{x}+\mathbf{e}_j, y)$ which is ≥ 0 (see Lemma 3.1 for the derivation of a similar condition (vi).), or yields $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y-1) + r_0 - v(\mathbf{x}+\mathbf{e}_j, y) - D_iv(\mathbf{x}, y) \geq D_iv(\mathbf{x}+\mathbf{e}_j, y) - D_iv(\mathbf{x}, y) \geq 0$ where the first inequality is due the underlying case assumption for the outcomes and the second inequality follows from $D_iv \uparrow x_j$ in (i').

(c) Finally, if the outcome of (3.12) is $D_iv(\mathbf{x}+\mathbf{e}_j, y)$, then (3.12) minus (3.13) reduces to $D_iv(\mathbf{x}+\mathbf{e}_j, y) - D_iv(\mathbf{x}, y) \geq 0$.

For operator T_0^2 we only show the result for $y < 0$ as the cases for $y \geq 0$ are identical to the ones analyzed in Lemma 3.2. For $x_i > 0 \; \forall i$, we have

$$\begin{aligned} D_iT_0^2v(\mathbf{x}+\mathbf{e}_j, y) &= \max[v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1}, y+1) + R_0 - r_0, v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j, y)] \\ &\quad - \max[v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y+1) + R_0 - r_0, v(\mathbf{x}+\mathbf{e}_j, y)] \end{aligned} \tag{3.14}$$

$$D_i T_1^1 v(\mathbf{x}, y) = \max[v(\mathbf{x}+\mathbf{e}_i-\mathbf{1}, y+1)+R_0-r_0, v(\mathbf{x}+\mathbf{e}_i, y)]$$
$$- \max[v(\mathbf{x}-\mathbf{1}, y+1)+R_0-r_0, v(\mathbf{x}, y)] \tag{3.15}$$

Eliminating the infeasible outcomes due to (iii'), we show that for each of the remaining cases we have (3.14) minus (3.15) ≥ 0.

(a) If (3.14) results in $D_i v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y+1)$, then (3.14) minus (3.15) may result in three possible expressions. $D_i v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y+1) - D_i v(\mathbf{x}-\mathbf{1}, y+1) \geq 0$ by (i'), $D_i v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y+1) - v(\mathbf{x}+\mathbf{e}_i-\mathbf{1}, y+1) - R_0 + r_0 + v(\mathbf{x}, y) \geq D_i v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y+1) - D_i v(\mathbf{x}-\mathbf{1}, y+1) \geq 0$ by (i'), and $D_{j,-1,P} v(\mathbf{x}+\mathbf{e}_i-\mathbf{1}, y) - D_{j,-1,P} v(\mathbf{x}-\mathbf{1}, y) \geq 0$ (see a similar justification outlined in Lemma 3.1 for (vii).)

(b) If the outcome of (3.14) is $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1}, y+1)+R_0-r_0-v(\mathbf{x}+\mathbf{e}_j, y)$, then (3.14) minus (3.15) is either $D_{i,-1,P} v(\mathbf{x}+\mathbf{e}_j-\mathbf{1}, y) - D_{i,-1,P} v(\mathbf{x}-\mathbf{1}, y) \geq 0$ or $v(\mathbf{x}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{1}, y+1)+R_0-r_0-v(\mathbf{x}+\mathbf{e}_j, y)-D_i v(\mathbf{x}, y) \geq D_i v(\mathbf{x}+\mathbf{e}_j, y)-D_i v(\mathbf{x}, y) \geq 0$ where the first inequality is due to the case assumption and the second inequality follows from (i').

(c) Finally, if (3.14) results in $D_i v(\mathbf{x}+\mathbf{e}_j, y+1)$, the only possible outcome for (3.14) minus (3.15) is $D_i v(\mathbf{x}+\mathbf{e}_j, y) - D_i v(\mathbf{x}, y) \geq 0$ by (i').

For boundary states where $x_i = 0$ or $x_j = 0$, the outcome discussed in part (a) becomes infeasible and only the cases in parts (b)-(c) apply with identical reasoning. For boundary states with $x_k = 0$, cases analyzed in parts (a)-(b) become infeasible and only part (c) applies. $\square$

CHAPTER IV

Joint Production and Admission Control in a Two-Stage Assemble-to-Order Manufacturing System

4.1 Overview

Business models such as make-to-stock, which may usually be preferred if the number of products offered is limited, lead to very significant inventory costs for a high variety of end products especially under both production and demand uncertainties. On the other hand, a make-to-order system keeps inventory only at the component level and products are assembled after a customer order is received. As many firms increasingly implement a make-to-order strategy, we are motivated by the challenges faced by firms in this setting to effectively coordinate the production of components and to allocate the assembly line capacity shared across many different products.

In this chapter, we study a two-stage assemble-to-order (ATO) system where component production lines feed a downstream shared assembly line capacity. The main difference between the setting of this chapter and the one studied in the preceding one is that, we now allow a customer to choose which components will be assembled into the end-product. Hence, the setting discussed in this chapter may be viewed as a basic mass customization environment for which an ATO strategy is implemented.

We are interested in how a firm operating under this setting should decide on how many components of each type to produce and how it should set demand admission rules to prioritize orders for various products that compete for the shared assembly capacity. Our main contributions in this chapter are to partially characterize the optimal policy within a certain region of the state space and develop a heuristic algorithm that is effective and robust with respect to the number of product alternatives offered.

4.2 Introduction

Out of many challenges that manufacturing companies face in the recent years, one of them is to produce high-mix, low-volume customizable products targeted to satisfy a vast variety of customer demands with minimum possible inventories. Business models such as produce-to-stock, which may usually be preferred if the number of products offered is limited, may lead to very significant inventory costs for a high variety of end products especially under both production and demand uncertainties. On the other hand, strategies like produce-to-order, although reducing inventory costs, may lead to high lead times under capacity uncertainties that exceed customers' willingness to wait and result in revenue losses.

In order to utilize the advantages and reduce the shortcomings of these strategies, many firms choose to operate in an ATO manufacturing setting, where the components used in the assembly process are produced to stock while the demand is still uncertain and a specific selection of those components demanded by a customer is assembled only after the demand is realized. An ATO manufacturing system can achieve high levels of responsiveness to customer demands with fairly lower amounts of inventory because the inventory is kept only at the component level. Among some example industries that use ATO manufacturing principles as a business model are the computer, furniture and to some extent, the automotive industries where mass customization is popular. Many practitioners and researches believe that the effective use of ATO manufacturing systems was one of the major contributors to Dell Computer Corporation's success [36].

Exploiting the advantages of the ATO manufacturing systems however brings great challenges to control the component production and assembly admission processes of a manufacturing firm. A firm has to make production or ordering decisions for components that will be required during the assembly stage while the demand is still uncertain. Yet, there are also uncertainties in the production process itself, so these decisions should be made taking into account possible machine outages and capacity disruptions.

When the firm observes demand for a variety of products, it also needs to determine admission rules regarding which type of product orders will be given priority. A thorough understanding of the characteristics of these decisions may help many companies focusing on low volume, high mix manufacturing to use their resources more efficiently, to lower overall inventories, and to achieve higher responsiveness to customer demands. As many researches recently identify, optimal policies in general ATO settings under production and demand uncertainty is still not known [53, 45].

In this chapter, we aim to provide insights to these problems by considering a firm in an ATO manufacturing setting that faces uncertainties in demand and due to the production and assembly capacities. Specifically, our focus will be on a manufacturing firm having a two-stage manufacturing system. The first stage consists of several dedicated production lines where each line produces a single type of component that will be required in the assembly stage. The components are produced to stock. These production lines feed a shared downstream assembly line which we refer to as the second stage of the manufacturing system. When a demand for a certain type of product arrives, the order may be accepted by picking up the corresponding component form the inventory and placing the order in the assembly line queue where the remaining assembly tasks will be performed. The firm also has the option to decline a demand for a certain type of product to ration its assembly line capacity to a more profitable product.

The problem can be summarized as joint component production and assembly admission control for an ATO manufacturing system. The component production lines and the product assembly line are modeled as M/M/1 queues in order to identify optimal policies regarding the firm's decision on when to accept or reject an incoming order for a certain product and how many components of each type is required. One of our contributions in this chapter is a partial characterization of the optimal production and demand admission decisions over a certain region of the state space. By performing extensive computational tests, we observed that the policy described for a certain region of the state space extends to the entire state space. Thus, we conjecture that the optimal production and admission policies described in this chapter hold throughout all possible states. Motivated by the insights we gain by the optimal policy structure, we develop a heuristic algorithm and test its performance against the optimal policy. Our computational studies based on numerous problem instances with varying problem parameters indicate that the algorithm is very effective and robust even when a higher number of end products is offered.

The remainder of this chapter is organized as follows. In Section 4.3, we review the related literature. We provide the problem formulation in Section 4.4 and analyze the structure of the component production and assembly admission policies in Section 4.5. In Section 4.6, we propose a heuristic solution approach and test its performance through numerical studies. We conclude in Section 4.7 and provide the proofs of theoretical results in Section 4.8.

4.3 Literature Review

Earlier works on assemble-to-order manufacturing systems have been conventionally focused on the characterization of optimal policies for the two special cases: (*i*) the distribution system where there is a single component and many products with the main issue being the allocation of the component and (*ii*) the assembly system where there are many components and a single product where the central concern is the coordination of the components. However, more recently, there has been growing interest in the general ATO manufacturing settings and an extensive literature survey has been provided by Song and Zipkin [53].

One of the earliest works is by Topkis [55], where he analyzes a distribution system. He uses a periodic review model where ordering decisions can be given only at specific periods. It is shown that a base-stock policy for ordering the components and a rationing policy for allocation of these components is optimal. Schmidt and Nahmias [50] study an assembly system with two components and one final product. They assume a two stage manufacturing system where both the production and assembly stages have deterministic lead times. They assume the end product is also produced before the stochastic demand is realized. Hence, there is inventory holding costs associated with the end product as well as the components. They develop the optimal assembly policy which states that there exists a target assemble-up-to point to reach as long as there are available components. They also identify the optimal production policies for the components which follows a modified produce up to policy due to differing replenishment lead times for the components. Rosling [48] extends the findings of Schmidt and Nahmias to multi-stage assembly systems with deterministic lead times and shows that under mild conditions on initial inventory levels, a balanced base-stock policy is optimal.

Gerchak and Henig [25] consider a periodic review problem of finding optimal production and allocation policies for a general ATO manufacturing system with stochastic demands and zero lead times under lost sales assumption. They show that a base stock production policy and a myopic allocation policy are optimal. Hausman et al [32] take a different approach by maximizing demand fulfillment probability, i.e. order fill rate, within a time window instead of profits. They also consider an ATO system with deterministic lead times. They assume that the production of each component is managed by an independent produce up to point which is known to be suboptimal but simpler to analyze. They allow backorders for unfulfilled demands and assume a first come first serve (FCFS) rule for satisfying demands from different periods. They determine optimal produce up to levels subject to an overall budget constraint.

Akcay and Xu [2] also consider a periodic review problem where the allocation of components across different periods is based on a FCFS basis. The system quotes a pre-specified response time window for each product and revenues are earned if the customer demand is satisfied within this time period. The production quantities and lead times are assumed to be deterministic and optimal base stock levels are found within a certain budget in order to achieve the maximum reward. Another work based on performance measurement is by Agrawal and Cohen [1]. Similar to Akcay and Xu, they assume an ATO system where backorders are permitted and there is a FCFS rule between different periods. They assume a consignment policy for component allocation meaning that a component is assigned to a product even if the assembly of that product will be delayed due to limitations on other components. They also propose a fair share allocation rule that allocates a fraction of the available inventory to different products where this fraction is determined by the ratio of the realized demand for a component due to a product to the total realized demand for that component. They use a service level constraint, order completion rate, as a constraint to determine the optimal base-stock levels for component production.

Song et al [51] performs an exact analysis on several performance measures regarding ATO systems. They assume independent production facilities for component production governed by independent base-stock levels. Their model assumes production uncertainty by using exponentially distributed processing times and hence is similar to the model discussed in the next section of this paper. They allow backlogging with a certain capacity of outstanding orders. Demands are satisfied based on a FCFS rule. For a given base-stock policy and a backlog capacity, they derive expressions for the performance measures such as order fill rate, service level, and waiting time distributions. In another work utilizing continuous time modeling, Lu and Song [41] compare the ATO model and single item newsvendor type models to derive upper bounds on base stock levels which could be used as starting points in a greedy search algorithm for the optimal base stock levels. They further study the effects of demand correlation to optimal base stock levels.

Plambeck and Ward [45] also consider a general ATO system with multiple components and multiple products. They assume stochastic component replenishment lead times with fixed ordering costs. Demand is also assumed to be random and backlogs are permitted. In addition, they allow expediting of components for a higher cost that enables the immediate availability of any limited component. They use discounted cost over a planning horizon as the objective to determine optimal decisions regarding sequencing the assembly of products, component production, and component expediting. Due to expediting, they show that the problem can be separated into single item inventory control problems, hence expediting is demonstrated as a means to simplify the analysis of ATO systems. Benjaafar and ElHafsi [4] focus on a special case of ATO systems with multiple demand classes based on demand rate and revenues but with all products having the same architecture and requiring one unit of each component. They assume independent production equipment for each component with exponentially distributed processing times. A queuing model is used to characterize the optimal production control and rationing policies.

The above selection of papers on assemble-to-order manufacturing systems shows that although the common issues in almost all of them are determining production and allocation decisions, there are quite distinctive approaches. While some researchers prefer to take profits and costs as objective functions to determine optimal production levels, others prefer service level performance measures and inventory budget constraints. Only a few of these papers aim to develop truly optimal production and assembly policies whereas many assume certain allocation rules during the assembly stage and independent base-stock levels for the production stage for practical purposes. Although, papers using continuous time models generally allow stochastic production lead times, their focus have rather been on certain special cases of ATO systems such as a single end product or allowing expediting to decouple the problem. Others considering a general ATO structure have been interested in developing expressions for performance measurements rather than control of the ATO system. In the vast majority of these papers, assembly is assumed to be instante-

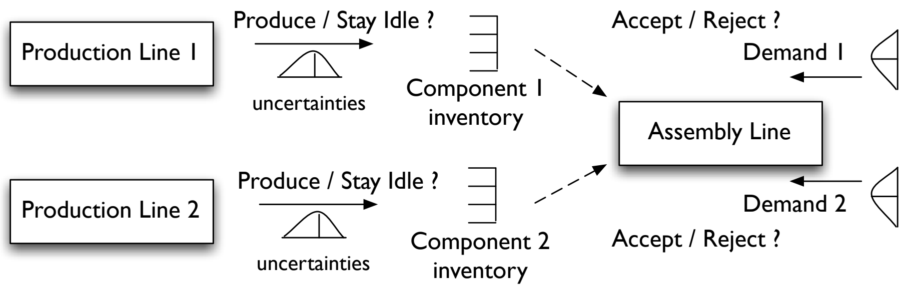


Figure 4.1: An assemble-to-order system for two products and demand types.

nous. We relax this assumption by allowing a finite production rate. We also relax the other assumption prevelant in the previous works that the products are assembled on a first come first serve basis. We consider the component production control problem for an assemble-to-order manufacturing systems jointly with the assembly admission problem.

4.4 Problem Formulation

We consider an assemble-to-order manufacturing system producing two distinct products. Demand for product type i ($i = 1, 2$) arrives based on a Poisson process with rate λ_i. Each product is composed of a unique, product-specific component and possibly some shared parts required by both products at the assembly stage.

The manufacturing system has two stages which is illustrated in Fig. 4.1. The first stage consists of two dedicated production lines where the product-specific components are produced to stock before demand is realized. It is assumed that production of a unit of component j ($j = 1, 2$) takes an exponentially distributed amount of time with mean $\frac{1}{\mu_j}$. Inventory costs are incurred at the rate of h_j per unit time for each unit of component j kept in stock. The production lines feed a shared assembly line. During this assembly stage, a product-specific component is turned into a final product through further operations which may include the joining of the component with other common parts shared by both products. We relax the assumption prevalent in most previous works that assembly is instantaneous and let the assembly operations also take an exponentially distributed amount of time with mean $\frac{1}{\mu_0}$, identical for both products. We note that identical assembly time distributions may seem restrictive at first, but it is a reasonable assumption in an assemble-to-order setting. Consider a computer manufacturer that assembles customized notebook computers. Customers may choose among different sized hard disks and computer memories as well as processors at different speeds. Although the production times of the variants within each category of these components may vary, once the customer's choice is placed in an assembly kit, the time to assemble the components are similar. This is due to the fact that even though each computer may have parts with different quality levels, the number of component categories required to assemble a computer do not vary significantly (i.e., all computers have hard disks, processors, etc.). Lastly, since all the products are assembled-to-order, no inventory is kept for finished products.

When an order for a certain type of product arrives, the firm has the option to accept the order and admit it into the assembly queue, or reject the order. Each accepted order for a product of type-i leads to a revenue of R_i. To account for the preference of customers' willingness to wait until delivery, the firm accrues a cost at a rate of b per unit time for each order in the assembly queue. If the firm decides to reject a demand by not admitting it into the assembly queue, the unsatisfied demand will be considered as lost sales.

The decision epochs considered in this model consists of all demand arrivals together with the production and assembly completions. At each decision epoch, a policy specifies whether a production server should stay idle or produce a unit of the corresponding component. In addition, at decision epochs corresponding to demand arrivals, the policy determines whether to accept or reject the assembly of that order.

Our objective is to find a control policy which maximizes the average profit per unit time over a long term. The profit is the revenue from accepted (assembled) orders minus the inventory holding costs for the components and the costs due to orders waiting in the assembly queue.

The overall problem consists of the control of (*i*) the production of the two product-specific components and (*ii*) the admission for assembly of the two types of product demands. The optimal production and admission control problem can be formulated as a Markov decision process. Let $S \in \mathbb{N}^3$ denote the state space and $(n_0, n_1, n_2) \in S$ be defined such that n_0 denotes the number of customer orders waiting in the assembly queue, n_1 and n_2 denote the amount of inventory for product-specific components 1 and 2 respectively. Note that our assumptions regarding the costs associated with customers waiting in the queue and the assembly times being identical for both products enables us to reduce the dimension of the state space due to the need of observing only the total number of customers waiting in the assembly queue.

When the demand for a certain type of product is admitted, its corresponding component is picked up form inventory to be assigned to the order and the order is placed in the assembly line queue where the remaining assembly tasks will be performed. Hence, if a product demand is accepted, the assembly queue length is increased by one unit and the corresponding component's inventory level is reduced by one unit. Letting $v(n_0, n_1, n_2)$ denote the relative value function of being in state (n_0, n_1, n_2) and g be the average profit per transition, we can present the average profit infinite horison dynamic programming problem as follows:

$$v(n_0, n_1, n_2) + g = \frac{1}{\Lambda} \left\{ \begin{aligned} & -bn_0 - h_1 n_1 - h_2 n_2 \\ & +\lambda_1 \max[(v(n_0+1, n_1-1, n_2) + R_1) \cdot I_{(n_1>0)} \\ & \qquad\qquad +v(n_0, n_1, n_2) \cdot I_{(n_1=0)}, v(n_0, n_1, n_2)] \\ & +\lambda_2 \max[(v(n_0+1, n_1, n_2-1) + R_2) \cdot I_{(n_2>0)} \\ & \qquad\qquad +v(n_0, n_1, n_2) \cdot I_{(n_2=0)}, v(n_0, n_1, n_2)] \\ & +\mu_0 \max[v(n_0-1, n_1, n_2) \cdot I_{(n_0>0)} \\ & \qquad\qquad +v(n_0, n_1, n_2) \cdot I_{(n_0=0)}, v(n_0, n_1, n_2)] \\ & +\mu_1 \max[v(n_0, n_1+1, n_2), v(n_0, n_1, n_2)] \\ & +\mu_2 \max[v(n_0, n_1, n_2+1), v(n_0, n_1, n_2)] \end{aligned} \right\} \tag{4.1}$$

where $\Lambda = \lambda_1 + \lambda_2 + \mu_0 + \mu_1 + \mu_2$, and $I_{(\cdot)}$ denotes the indicator function.

In (4.1), the terms $\frac{1}{\Lambda}\{-bn_0 - h_1 n_1 - h_2 n_2\}$ denote the expected costs per decision epoch due to customers' waiting in the queue and the holding of component inventories; the terms multiplied by λ_i $(i = 1, 2)$ correspond to transitions and revenues generated with the arrival of a demand for product type-i; and the terms multiplied by μ_j $(j = 0, 1, 2)$ correspond to transitions generated by either a product assembly or a component production completion opportunity. Since the transitions occur with rate Λ, the profit per unit time can be represented as $g\Lambda$.

4.5 A Partial Characterization of the Optimal Production and Admission Policy

The optimal component production policy states whether a production line should produce another component to meet anticipated demand or stay idle to avoid excessive inventory. The optimal assembly admission policy, on the other hand, determines whether it is more profitable to admit a certain type of product demand into the assembly queue or reject it to preserve the assembly capacity to the other product or to prevent higher costs associated with longer assembly queues. These two policies jointly affect the long run profitability of the manufacturing firm.

In this we section, we provide a characterization of the optimal policy structure for a certain region of the state space. To that end, we first define the following additional notation for any real valued function v:

$D_0 v(n_0, n_1, n_2) = v(n_0 + 1, n_1, n_2) - v(n_0, n_1, n_2)$,

$D_1 v(n_0, n_1, n_2) = v(n_0, n_1 + 1, n_2) - v(n_0, n_1, n_2)$,

$D_2 v(n_0, n_1, n_2) = v(n_0, n_1, n_2 + 1) - v(n_0, n_1, n_2)$,

and the combinations such as:

$D_{0,-1} v(n_0, n_1, n_2) = v(n_0 + 1, n_1 - 1, n_2) - v(n_0, n_1, n_2)$, and

$D_{0,-2} v(n_0, n_1, n_2) = v(n_0 + 1, n_1, n_2 - 1) - v(n_0, n_1, n_2)$

For any state (n_0, n_1, n_2), D_0 represents the additional value of having an additional order waiting in the assembly queue. D_1 and D_2 represent the additional value of having an additional unit of component 1 and component 2 inventory, respectively. $D_{0,-1}$ and $D_{0,-2}$ are the additional values of having accepted a type 1 and a type 2 demand, respectively.

The following sub- and super-modularity conditions are sufficient to prove the structure of the optimal policy where the symbols $\uparrow$ and $\downarrow$ are used in the weak sense and refer to non-decreasing and non-increasing, respectively.

$$
\begin{array}{rl}
\text{(i)} & D_0 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2, \text{ and } \leq 0 \\
\text{(ii)} & D_1 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2, \text{ and } \leq R_1 \\
\text{(iii)} & D_2 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2, \text{ and } \leq R_2 \\
\text{(iv)} & D_{0,-1} v(n_0, n_1, n_2) \downarrow n_0, \uparrow n_1, \text{ and } \downarrow n_2 \\
\text{(v)} & D_{0,-2} v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \text{ and } \uparrow n_2
\end{array}
\tag{4.2}
$$

In (4.2), condition (i), $D_0 \downarrow n_0, n_1$ and n_2 means the additional value gained by completing the assembly of an order in the queue (i.e., $v(n_0-1, n_1, n_2) - v(n_0, n_1, n_2)$ or equivalently, $-D_0 v(n_0 - 1, n_1, n_2)$) gets larger with each additional order in the queue and component in inventory. The fact that $D_0 \downarrow n_0$ and ≤ 0 implies that it is not optimal for the assembly line to stay idle if there are any orders waiting in the queue.

Conditions (ii) and (iii) directly imply the characteristics of the optimal production policies. For example, $D_1 \downarrow n_1$ means that the additional value gained by producing a unit of component 1 gets smaller with each additional unit of component 1 in inventory. Hence, if it is optimal not to produce component 1 in a state (n_0, n_1, n_2), (i.e., $D_1 v(n_0, n_1, n_2) < 0$) it remains optimal not to produce component 1 in state $(n_0, n_1 + 1, n_2)$. $D_1 \downarrow n_0$ and $D_1 \downarrow n_2$ imply that if it is optimal not to produce component 1 in state (n_0, n_1, n_2), then it is also optimal not to produce component 1 in states $(n_0 + 1, n_1, n_2)$ or $(n_0, n_1, n_2 + 1)$.

In a similar fashion, conditions (iv) and (v) imply the structure of the assembly admission policies. For example, $D_{0,-1} \downarrow n_0$, and $\downarrow n_2$ implies that if it is optimal not to admit a type 1 demand in state (n_0, n_1, n_2), (i.e., $D_{0,-1} v(n_0, n_1, n_2) < -R_1$) then it remains optimal not to admit a type 1 demand in states with a longer assembly queue, $(n_0 + 1, n_1, n_2)$, or with higher component 2 inventories, $(n_0, n_1, n_2 + 1)$. $D_{0,-1} \uparrow n_1$ suggests that if it is optimal to admit a type 1 demand in state (n_0, n_1, n_2), (i.e., $D_{0,-1} v(n_0, n_1, n_2) > -R_1$) then, it is also optimal to accept a type 1 demand when component 1 inventory is higher, $(n_0, n_1 + 1, n_2)$.

Conditions (i)-(v), if hold, are sufficient to the characterize the optimal policy structure. In our model, unfortunately, the conditions $D_{0,-1} v(n_0, n_1, n_2) \downarrow n_2$ and $D_{0,-2} v(n_0, n_1, n_2) \downarrow n_1$ in (iv) and (v) do not necessarily hold over the entire state space. However, through extensive numerical analysis, we find that they hold within a certain region of the state space away from the boundary where $n_0 = 0$. If we let a subregion of the state space $P \subset S$ to be defined as $P := \{(n_0, n_1, n_2) \mid D_{0,-1}v(n_0, n_1, n_2) \downarrow n_2$ and $D_{0,-2}v(n_0, n_1, n_2) \downarrow n_1\}$, then we can define the structure of the optimal production and admission policies within this particular region by the following Theorem.

Theorem 4.1. *For $(n_0, n_1, n_2) \in P$, the structure of the optimal policy is as follows:*
(a) The optimal assembly admission policy for each product is defined by a switching surface. For products of type i $(i = 1, 2)$, if $n_0 \leq \gamma_i(n_1, n_2)$, the optimal policy accepts to assemble the order for product i, otherwise it rejects the order. The switching surface for product i, $\gamma_i(n_1, n_2)$, is non-decreasing in n_i and non-increasing in $n_{(3-i)}$.

(b) The optimal component production policy for each component j $(j = 1, 2)$ is also defined by a switching surface $\beta_j(n_0, n_{(3-j)})$ such that the optimal policy is to produce an additional unit of component j if $n_j \leq \beta_j(n_0, n_{(3-j)})$, and to stay idle otherwise. Furthermore, $\beta_j(n_0, n_{(3-j)})$ is non-increasing in n_0 and in $n_{(3-j)}$.

(c) It is not optimal for the assembly line to stay idle if there are orders waiting in the assembly queue.

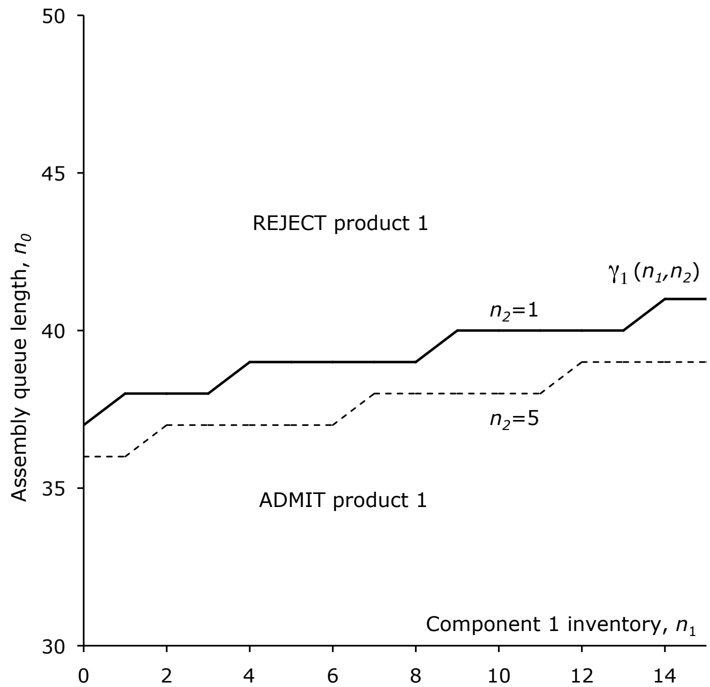
Proof: The proof of Theorem 4.1 is provided in Section 4.8.

In part (a) of the above theorem, $n_0 \leq \gamma_i(n_1, n_2)$ implies that there is a state-dependent assembly queue length threshold, i.e., a cutoff point, for a product to be assembled. The product will be assembled only if the assembly queue length is shorter than this threshold. This threshold level depends on the inventory levels of both components. Moreover, for example, $\gamma_1(n_1, n_2)$, being non-decreasing in n_1 implies that while it is not optimal to admit a product type-1 demand for assembly for a specific queue length and inventory levels, it may be optimal to admit a product type-1 demand for assembly when the inventory level of component 1 is higher. In a similar fashion, $\gamma_1(n_1, n_2)$, being non-increasing in n_2 implies that while it is optimal to admit a product type-1 demand for assembly for a certain instance, it may be optimal to reject a product type-1 demand when the inventory level of component 2 is higher.

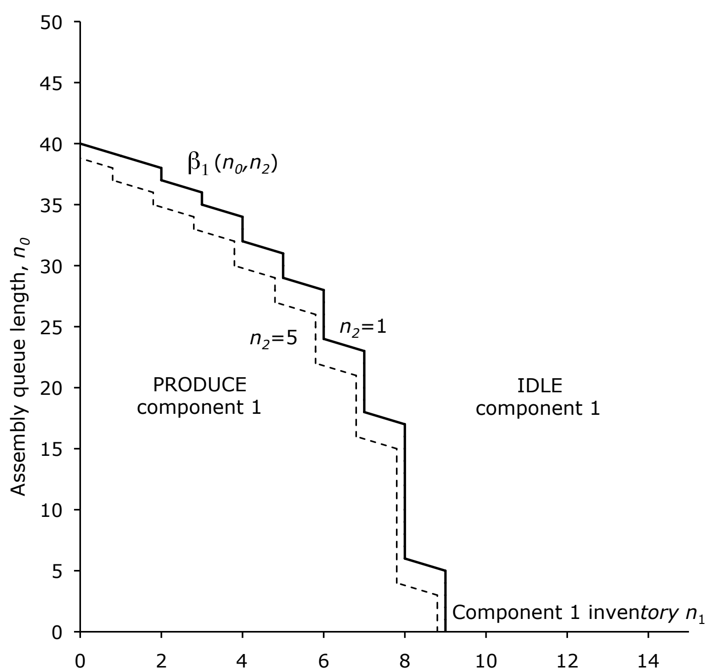
Similarly, in part (b) of the theorem, $n_j \leq \beta_j(n_0, n_{(3-j)})$ implies that the decision regarding whether another unit of a component needs to be produced or not is guided by a state-dependent base-stock policy. A certain type of component will be produced only if its inventory is lower than this base-stock level. This base-stock level depends on the assembly queue length as well as the inventory level of the other component. As an example, $\beta_2(n_0, n_1)$ being non-decreasing in n_0 and in n_1 implies that while it is optimal to produce another unit of component 2 for a specific queue length and inventory levels, it may be optimal not to produce another unit of component 2 if the assembly queue length was longer and/or the inventory level of component 1 was higher.

Even though we partially characterize the structure of the optimal policy in a specific subregion of the state space, throughout a large number of problems we have tested, we have observed that the general structure defined by Theorem 4.1 held over the entire state space. Therefore, we conjecture that the structure given in Theorem 4.1 is optimal for the entire state space, i.e. a state dependent produce-up-to policy is optimal for the production of components and there exists state dependent assembly queue length thresholds for each type of product, beyond which it is optimal to reject a demand for the corresponding product.

We will now illustrate the structure of the demand admission and component production policies for an example problem with two components with the following



(a) Structure of demand admission policies



(b) Structure of component production policies

Figure 4.2: Structure of production and demand admission policies

parameters: $\lambda_1 = \lambda_2 = 4$, $\mu_0 = 10$, $\mu_1 = \mu_2 = 5$,, $R_1 = 75$, $R_2 = 100$, $h_1 = h_2 = 2$, and $b = 5$. In Figure 4.2, the product admission and component production policies for product type-1 are displayed. For this example problem consisting of two components, the optimal threshold values are defined by switching surfaces in three dimensions. The solid and the dotted curves in both figures are results of two-dimensional cuts on the switching surfaces at two separate values of the type-2 component inventory levels.

As illustrated in Figure 4.2 (a), a type-1 product demand will be admitted for assembly if the current state of the system falls below this switching curve. In other words, any component inventory pair (n_1, n_2) corresponds to a certain admission threshold value for a type-1 product. A demand for this product will be accepted only if the assembly queue length, n_0, is less than this threshold value. For queue lengths exceeding this amount, it is optimal not to admit the product for assembly. It can be observed that the admission switching surface for product 1 is non-decreasing in the amount of component 1 inventory. The dashed line demonstrates how this threshold decreases when component 2 inventory is higher.

The production switching curve for the type-1 component is shown in Figure 4.2 (b). An assembly queue length and component 2 inventory pair (n_0, n_2) corresponds to a certain base-stock level for component 1. If component 1 inventory level, n_1, is lower than this base-stock level, then it is optimal to produce an additional unit of component 1. This base-stock level decreases with both the assembly queue length and the component 2 inventory.

4.6 A Heuristic Algorithm

In this section, we use the insights gained from the preceding analysis to develop a heuristic algorithm to determine component production and assembly admission decisions. The heuristic solution is obtained in two stages: First, we identify the appropriate admission threshold level for each type of product. Taking into consideration the results of this first stage, we then find corresponding produce-up-to levels for each component.

Stage 1: Determining admission thresholds:

Our goal is to determine multiple admission thresholds, one for each type of product. We first present an optimal admission problem for a single product, formulated as a Markov decision process. Define x as the number of items in the assembly queue and let $v(x)$ be the relative value function of being in state x and g denote the average profit per transition. In addition, let λ and μ denote the demand arrival and the assembly rates respectively, b denote the cost per unit time for each item kept in the assembly queue and R denote the revenue generated by an accepted demand. We can then write,

$$v(x) + g = \frac{1}{\lambda+\mu} \left\{ \begin{array}{l} -bx + \mu(v(x-1) \cdot I_{(x>0)} + v(x) \cdot I_{(x=0)}) \\ +\lambda \max[v(x+1) + R, v(x)] \end{array} \right\} \tag{4.3}$$

The optimal assembly admission policy has a threshold form [43, 40]. For some $l \geq 0$, the optimal policy will be to admit for assembly if $x \leq l$, and to reject the demand if $x > l$. This structure enables us to use a simple one dimensional search to find l, the cutoff point for assembly admission that maximizes the profit. For any choice of l, the number of items in the assembly queue, X, behaves as an $M/M/1/l$ queue.

In order to find the assembly admission cutoff points for two products, we use the following approach: We first find the admission threshold for the lower revenue

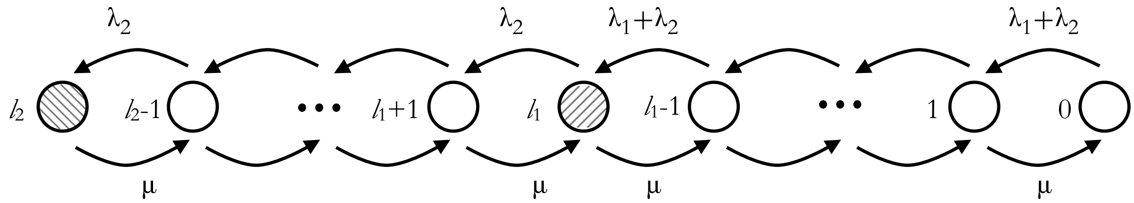


Figure 4.3: Transition rate diagram

item assuming that whenever a lower revenue item is admitted, the higher revenue item would also have been admitted. Then, we search for an admission threshold for the higher revenue item in a similar fashion. Without loss of generality, we label the products such that $R_1 \leq R_2$. To determine l_1, we let the total arrival into the assembly queue be given by $\lambda_1 + \lambda_2$ and substitute R_1 for R in (4.3). We opt not to use a weighted revenue such as $R = (\lambda_1 R_1 + \lambda_2 R_2)/(\lambda_1 + \lambda_2)$ as this would lead to an unfair amount of item 1 admissions if $R_2 >> R_1$. Next we determine l_2 by setting the arrival rate to the queue beyond l_1 be limited to λ_2 generating a revenue of R_2.

Stage 2: Determining base-stock levels:

To determine the base stock levels, we first adjust the demands for both products taking into consideration the orders that were rejected. For a system with admission threshold levels $l_2 \geq l_1$, standard results in queueing theory [47] allows us to easily compute the blocking probabilities for this $M/M/1/l_2$ queue (with transitions shown in Figure 4.3) by solving the global balance equations. Letting $p(l_1)$ and $p(l_2)$ denote the blocking probabilities for product 1 and product 2 arrivals, respectively, we adjust the demand for each product by $\lambda_1' := (1 - p(l_1))\lambda_1$ and $\lambda_2' := (1 - p(l_2))\lambda_2$.

Next, we consider the optimal production control problem for a single product. We use much of the previous notation except we now define x as the number of items in the component inventory. Further, let μ denote the component production rate and h denote the cost per unit time for each item kept in the component inventory. We can then write,

$$v(x) + g = \frac{1}{\lambda + \mu} \left\{ \begin{array}{l} -hx + \mu \max[(v(x+1), v(x)] \\ +\lambda((v(x-1) + R) \cdot I_{(x>0)} + v(x) \cdot I_{(x=0)}) \end{array} \right\} \tag{4.4}$$

The optimal policy is again of a threshold form. For some $s \geq 0$, the optimal policy will be to produce if $x \leq s$ and not to produce if $x > s$. For any choice of s, $s - X$ behaves like an $M/M/1/s$ queue where s implies the queue size. For example, when the physical component inventory, X, is zero, a demand arrival will be lost. This corresponds to a queue length of s in the $M/M/1/s$ queue, i.e., a totally full queue, hence a demand arrival will be lost. Using a one-dimensional search and substituting the adjusted demand levels obtained previously, we find the base stock levels s_1 and s_2 corresponding to products 1 and 2, respectively.

The following two tables provide results of numerical studies comparing the heuristic policy and the optimal policy. Table 4.1 displays a full numerical experiment including 32 problem instances for a two product case where the revenues form the products are $R_1 = 75$ and $R_2 = 100$. Two different demand arrival scenarios, one with the lower priced item having a higher demand rate, and the other with both products having the same demand rate, were tested. For each demand scenario, the corresponding production rates and the assembly rate were independently varied between utilization values of 80% and 90%. Finally, component holding and assembly queue waiting costs were alternated between a high and a low set of values. l_1 and l_2 correspond to the assembly admission threshold values found by the heuristics for product type 1 and 2, respectively. Similarly, s_1 and s_2 are the base-stock levels derived for the components. The average profit per unit time for the optimal and the heuristic policy are recorded together with the percent difference of the heuristics

Table 4.1: Performance of the heuristics for two products.

No	λ_1	λ_2	ρ_0	ρ_1	ρ_2	h_i	b	l_1	l_2	s_1	s_2	Opt.	Heur.	% Diff
1	4	2	0.8	0.8	0.8	1	3	28	37	13	11	464.5	463.9	0.13
2	4	2	0.9	0.8	0.8	1	3	18	31	12	11	452.9	450.9	0.44
3	4	2	0.8	0.9	0.8	1	3	28	37	17	11	460.4	460.0	0.08
4	4	2	0.9	0.9	0.8	1	3	18	31	16	11	449.6	447.8	0.39
5	4	2	0.8	0.8	0.9	1	3	28	37	11	15	461.3	461.0	0.05
6	4	2	0.9	0.8	0.9	1	3	18	31	10	14	450.2	449.1	0.24
7	4	2	0.8	0.9	0.9	1	3	28	37	17	15	457.1	456.7	0.09
8	4	2	0.9	0.9	0.9	1	3	18	31	16	14	446.8	445.2	0.35
9	4	2	0.8	0.8	0.8	2	5	18	32	10	9	443.8	442.3	0.33
10	4	2	0.9	0.8	0.8	2	5	13	26	9	9	429.4	425.8	0.84
11	4	2	0.8	0.9	0.8	2	5	18	32	13	9	438.8	437.4	0.31
12	4	2	0.9	0.9	0.8	2	5	13	26	12	9	425.4	421.7	0.88
13	4	2	0.8	0.8	0.9	2	5	18	32	10	11	440.0	438.8	0.28
14	4	2	0.9	0.8	0.9	2	5	13	26	9	11	426.2	422.9	0.77
15	4	2	0.8	0.9	0.9	2	5	18	32	13	11	435.0	433.9	0.26
16	4	2	0.9	0.9	0.9	2	5	13	26	12	11	422.2	418.8	0.80
17	3	3	0.8	0.8	0.8	1	3	22	36	11	13	489.1	488.5	0.12
18	3	3	0.9	0.8	0.8	1	3	15	35	11	13	477.5	474.4	0.65
19	3	3	0.8	0.9	0.8	1	3	22	36	15	13	485.7	485.1	0.11
20	3	3	0.9	0.9	0.8	1	3	15	35	14	13	474.7	472.0	0.57
21	3	3	0.8	0.8	0.9	1	3	22	36	11	17	484.9	484.5	0.07
22	3	3	0.9	0.8	0.9	1	3	15	35	11	17	473.8	471.0	0.60
23	3	3	0.8	0.9	0.9	1	3	22	36	15	17	481.4	481.1	0.07
24	3	3	0.9	0.9	0.9	1	3	15	35	14	17	471.0	468.6	0.52
25	3	3	0.8	0.8	0.8	2	5	15	32	9	10	467.9	466.5	0.30
26	3	3	0.9	0.8	0.8	2	5	10	32	8	10	453.5	448.5	1.10
27	3	3	0.8	0.9	0.8	2	5	15	32	11	10	463.9	462.8	0.23
28	3	3	0.9	0.9	0.8	2	5	10	32	10	10	450.4	445.7	1.05
29	3	3	0.8	0.8	0.9	2	5	15	32	9	13	462.8	461.5	0.30
30	3	3	0.9	0.8	0.9	2	5	10	32	8	13	449.2	444.2	1.11
31	3	3	0.8	0.9	0.9	2	5	15	32	11	13	458.8	457.7	0.23
32	3	3	0.9	0.9	0.9	2	5	10	32	10	13	445.9	441.3	1.04

Table 4.2: Performance of the heuristics for three products.

No	λ_1	λ_2	λ_3	μ_0	μ_1	μ_2	μ_3	h	b	Opt.	Heur.	% Diff
1	3	4	2	0.8	0.8	0.8	0.8	1	3	604.0	602.8	0.20
2	3	4	2	0.9	0.9	0.8	0.8	1	3	590.5	585.8	0.81
3	3	4	2	0.8	0.9	0.9	0.8	1	3	597.2	596.3	0.15
4	3	4	2	0.9	0.8	0.9	0.8	1	3	588.9	583.9	0.84
5	3	4	2	0.8	0.9	0.8	0.9	1	3	598.2	597.0	0.19
6	3	4	2	0.9	0.8	0.8	0.9	1	3	589.6	584.6	0.86
7	3	4	2	0.8	0.8	0.9	0.9	1	3	596.5	595.5	0.17
8	3	4	2	0.9	0.9	0.9	0.9	1	3	583.7	579.4	0.74
9	3	4	2	0.8	0.9	0.8	0.8	2	5	574.4	572.1	0.41
10	3	4	2	0.9	0.8	0.8	0.8	2	5	563.2	555.7	1.34
11	3	4	2	0.8	0.8	0.9	0.8	2	5	572.2	569.1	0.53
12	3	4	2	0.9	0.9	0.9	0.8	2	5	556.4	549.6	1.22
13	3	4	2	0.8	0.8	0.8	0.9	2	5	573.5	570.6	0.50
14	3	4	2	0.9	0.9	0.8	0.9	2	5	557.4	550.8	1.18
15	3	4	2	0.8	0.9	0.9	0.9	2	5	565.3	563.2	0.36
16	3	4	2	0.9	0.8	0.9	0.9	2	5	555.1	547.9	1.30
17	3	3	3	0.8	0.9	0.8	0.8	1	3	626.0	624.8	0.19
18	3	3	3	0.9	0.8	0.8	0.8	1	3	617.2	612.2	0.81
19	3	3	3	0.8	0.8	0.9	0.8	1	3	625.1	624.0	0.17
20	3	3	3	0.9	0.9	0.9	0.8	1	3	612.0	607.3	0.76
21	3	3	3	0.8	0.8	0.8	0.9	1	3	624.3	623.2	0.18
22	3	3	3	0.9	0.9	0.8	0.9	1	3	611.1	606.8	0.71
23	3	3	3	0.8	0.9	0.9	0.9	1	3	618.2	617.4	0.13
24	3	3	3	0.9	0.8	0.9	0.9	1	3	610.1	605.3	0.79
25	3	3	3	0.8	0.8	0.8	0.8	2	5	601.6	598.7	0.48
26	3	3	3	0.9	0.9	0.8	0.8	2	5	585.1	578.6	1.12
27	3	3	3	0.8	0.9	0.9	0.8	2	5	594.5	592.6	0.32
28	3	3	3	0.9	0.8	0.9	0.8	2	5	583.8	576.9	1.18
29	3	3	3	0.8	0.9	0.8	0.9	2	5	593.1	591.3	0.30
30	3	3	3	0.9	0.8	0.8	0.9	2	5	582.6	575.4	1.23
31	3	3	3	0.8	0.8	0.9	0.9	2	5	592.1	589.6	0.42
32	3	3	3	0.9	0.9	0.9	0.9	2	5	576.6	570.5	1.06

from the optimal.

The average difference between the heuristic and the optimal policy in Table 4.1 was approximately 0.45%. The heuristic algorithm was effective for all cases studied, where the maximum difference from the optimal was 1.11%. In general, when the products had different arrival rates (No 1-16), the heuristics performed slightly better. The assembly line rate was found as a significant factor on the performance of the heuristics, where the cases corresponding to lower assembly utlizations (odd numbered cases) performed better than the ones corresponding to higher utilization rates. However, the component production rates, i.e., the production line utilizations, did not have a significant effect on the performance. Finally, when the waiting cost in the assembly queue was higher compared to the component inventory holding costs (No 9-16, 25-32), the performance of the heuristics slightly deteriorated.

Table 4.2 shows numerical results for an extension to three products with revenues set as $R_1 = 50, R_2 = 75$, and $R_3 = 100$. A partial experiment was carried on to test the performance of the heuristics. The average difference between the heuristic and the optimal policy were found to be 0.65% with a maximum difference of 1.34%, indicating that the heuristic policy maintains its effectiveness at this higher number of end-products.

The heuristic algorithm that we have developed takes into account that demands are prioritized by their revenues and assigns separate thresholds for each product. Moreover, the base-stock levels for each product are determined by considering the effects of this admission threshold policy. Hence, compared to a first-come first-serve admission rule, the heuristic algorithm draws on several important properties of the optimal policy structure. In addition, the algorithm may easily be implemented as it does not require the current state information.

4.7 Conclusions

In this chapter, we considered a basic mass customization setting by focusing on a manufacturing firm that offers a customer to choose among several alternatives of a product. We were particularly interested in the manufacturer's decisions on how to decide the right amount of component inventory to hold and how to dynamically allocate its shared assembly capacity among customer orders generating different revenues.

We investigated and partially characterized the structures of the optimal component production and assembly admission policies. We showed that the optimal production decisions can be defined by a state-dependent base-stock policy where the base-stock level for a component decreases with the inventory level of other components and the assembly queue length. In addition, the optimal demand admission decision for each type of product is described by a state-dependent admission threshold level. Fewer demands for a certain product are admitted as more customer orders are waiting in the assembly queue or when the inventory level of the other components are higher. Through an extensive numerical analysis, we observed that the policy described for a certain region of the state space extends to the entire state space. Thus, we conjectured that the optimal production and admission policies described in this chapter hold throughout all possible states.

Finally, we devised a heuristic solution algorithm that is motivated by the insights we gained from the optimal policy structure. We tested the performance of the algorithm for a two- and a three-product setting and observed that the profits obtained by the heuristic policy remained within a narrow gap of the profits attained by the optimal policy.

4.8 Appendix

Proof of Theorem 4.1:

The proof of Theorem 4.1 follows the same framework as the proof of Theorem 3.1 in Section 3.5. In order to simplify the representation of the analysis, we first introduce some additional notation. For any value function v, we define the following operators:

$$\begin{aligned}
T_1 v(n_0, n_1, n_2) &= \max[(v(n_0+1, n_1-1, n_2) + R_1) \cdot I_{(n_1>0)} \\
&\qquad + v(n_0, n_1, n_2) \cdot I_{(n_1=0)}, v(n_0, n_1, n_2)], \\
T_2 v(n_0, n_1, n_2) &= \max[(v(n_0+1, n_1, n_2-1) + R_2) \cdot I_{(n_2>0)} \\
&\qquad + v(n_0, n_1, n_2) \cdot I_{(n_2=0)}, v(n_0, n_1, n_2)], \\
T_3 v(n_0, n_1, n_2) &= \max[v(n_0-1, n_1, n_2) \cdot I_{(n_0>0)} + v(n_0, n_1, n_2) \cdot I_{(n_0=0)}, v(n_0, n_1, n_2)], \\
T_4 v(n_0, n_1, n_2) &= \max[v(n_0, n_1+1, n_2), v(n_0, n_1, n_2)], \\
T_5 v(n_0, n_1, n_2) &= \max[v(n_0, n_1, n_2+1), v(n_0, n_1, n_2)], \\
T v(n_0, n_1, n_2) &= \tfrac{1}{\Lambda}[-bn_0 - h_1 n_1 - h_2 n_2 + \lambda_1 T_1 v(n_0, n_1, n_2) + \lambda_2 T_2 v(n_0, n_1, n_2) \\
&\qquad + \mu_0 T_3 v(n_0, n_1, n_2) + \mu_1 T_4 v(n_0, n_1, n_2) + \mu_2 T_5 v(n_0, n_1, n_2)]
\end{aligned}$$

We restate the sufficient conditions (i)-(v) given in (4.2) that directly imply the optimal policy structure with the addition of four other technical conditions that assists us in the proof.

Let V be the set of functions on P such that if $v \in V$, then:

(i) $D_0 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2$, and ≤ 0

(ii) $D_1 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2$, and $\leq R_1$

(iii) $D_2 v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1, \downarrow n_2$, and $\leq R_2$

(iv) $D_{0,-1} v(n_0, n_1, n_2) \downarrow n_0, \uparrow n_1$, and $\downarrow n_2$

(v) $D_{0,-2} v(n_0, n_1, n_2) \downarrow n_0, \downarrow n_1$, and $\uparrow n_2$

(vi) $D_{0,-1}v(n_0, n_1, n_2) \downarrow$ in the direction (1,0,-1)

(vii) $D_{0,-2}v(n_0, n_1, n_2) \downarrow$ in the direction (1,-1,0)

(viii) $D_{-1,2}v(n_0, n_1, n_2) \downarrow n_1$

(ix) $D_{1,-2}v(n_0, n_1, n_2) \downarrow n_2$

In the following lemma, we show that these conditions are preserved through recursions under the operator T.

Lemma 4.1. *If $v \in V$ then $T_1 v$, $T_2 v$, $T_3 v$, $T_4 v$, $T_5 v$, $Tv \in V$.*

Proof: For brevity, we present only the proof that D_1 is non-increasing in n_0. The proofs showing that other conditions are also preserved under the operators is similar and therefore omitted. We first show that $D_1 T_1 v(n_0, n_1, n_2) \downarrow n_0$. For any (n_0, n_1, n_2) such that $n_1 > 0$,

$$\begin{aligned} D_1 T_1 v(n_0, n_1, n_2) = {} & \max[v(n_0+1, n_1, n_2) + R_1, v(n_0, n_1+1, n_2)] \\ & - \max[v(n_0+1, n_1-1, n_2) + R_1, v(n_0, n_1, n_2)] \end{aligned} \tag{4.5}$$

$$\begin{aligned} D_1 T_1 v(n_0+1, n_1, n_2) = {} & \max[v(n_0+2, n_1, n_2) + R_1, v(n_0+1, n_1+1, n_2)] \\ & - \max[v(n_0+2, n_1-1, n_2) + R_1, v(n_0+1, n_1, n_2)] \end{aligned} \tag{4.6}$$

We need to show that (4.6) minus (4.5) ≤ 0. There are four possible outcomes for the two max functions in (4.6). One of these outcomes is $v(n_0+1, n_1+1, n_2) - v(n_0+2, n_1-1, n_2) - R_1$ which requires $v(n_0+2, n_1, n_2) - v(n_0+1, n_1+1, n_2) \leq -R_1$ and $v(n_0+2, n_1-1, n_2) - v(n_0+1, n_1, n_2) \geq -R_1$. However, this violates $D_{0,-1}v \uparrow n_1$ (recall $v \in V$) and therefore is not feasible. Three possible cases remain.

Case(1): If the outcome of (4.6) is $v(n_0+2, n_1, n_2) - v(n_0+2, n_1-1, n_2)$, the only possible outcome for (4.5) is $v(n_0+1, n_1, n_2) - v(n_0+1, n_1-1, n_2)$ due to $D_{0,-1}v \downarrow n_0$.

(4.6) minus (4.5) yields $D_1 v(n_0+2, n_1-1, n_2) - D_1 v(n_0+1, n_1-1, n_2)$, and this is ≤ 0 by $D_1 v \downarrow n_0$.

Case(2): If the outcome of (4.6) is $v(n_0+2, n_1, n_2) + R_1 - v(n_0+1, n_1, n_2)$, then $v(n_0+1, n_1, n_2) + R_1$ must be the outcome of the first max function in (4.5) (due to $D_{0,-1} v \downarrow n_0$). There are two possible outcomes for the second max function in (4.5). Case(2a): If the outcome of (4.5) is $v(n_0+1, n_1, n_2) + R_1 - v(n_0, n_1, n_2)$, then (4.6) minus (4.5) yields $v(n_0+2, n_1, n_2) - v(n_0+1, n_1, n_2) - v(n_0+1, n_1, n_2) + v(n_0, n_1, n_2)$, or equivalently $D_0 v(n_0+1, n_1, n_2) - D_0 v(n_0+1, n_1, n_2)$, and that is ≤ 0 by $D_0 v \downarrow n_0$. Case(2b): If the outcome of (4.5) is $v(n_0+1, n_1, n_2) - v(n_0+1, n_1-1, n_2)$, then (4.6) minus (4.5) yields $v(n_0+2, n_1, n_2) + R_1 - v(n_0+1, n_1, n_2) - v(n_0+1, n_1, n_2) + v(n_0+1, n_1-1, n_2) \le v(n_0+2, n_1, n_2) - v(n_0+1, n_1, n_2) - v(n_0+2, n_1-1, n_2) + v(n_0+1, n_1-1, n_2) = D_0 v(n_0+1, n_1, n_2) - D_0 v(n_0+1, n_1-1, n_2)$, and this is ≤ 0 by $D_0 v \downarrow n_1$. (The first inequality follows from $v(n_0+2, n_1-1, n_2) + R_1 \le v(n_0+1, n_1, n_2)$ which must hold based on the outcome of (4.6) for this specific case.)

Case(3): When the outcome of (4.6) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2)$, there are three possible outcomes for the two max functions in (4.5). Case(3a): If $v(n_0+1, n_1, n_2) - v(n_0+1, n_1-1, n_2)$ comes out of (4.5), then (4.6) minus (4.5) is equal to $D_1 v(n_0+1, n_1, n_2) - D_1 v(n_0+1, n_1-1, n_2)$, and that is ≤ 0 by $D_1 v \downarrow n_1$. Case(3b): If $v(n_0+1, n_1, n_2) + R_1 - v(n_0, n_1, n_2)$ comes out of (4.5), then (4.6) minus (4.5) gives $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2) - v(n_0+1, n_1, n_2) - R_1 + v(n_0, n_1, n_2) \le v(n_0+1, n_1+1, n_2) - v(n_0, n_1+1, n_2) - v(n_0+1, n_1, n_2) + v(n_0, n_1, n_2) = D_0 v(n_0, n_1+1, n_2) - D_0 v(n_0, n_1, n_2)$, and this is ≤ 0 by $D_0 v \downarrow n_1$. (The first inequality follows from $v(n_0+1, n_1, n_2) + R_1 \ge v(n_0, n_1+1, n_2)$ which holds due to the outcome of (4.5) for this specific case.) Case(3c): The only remaining possible outcome for (4.5) is $v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2)$ in which case (4.6) minus (4.5) results in $D_1v(n_0+1,n_1,n_2) - D_1v(n_0,n_1,n_2)$, and this is ≤ 0 by $D_1v \downarrow n_0$.

Finally, we consider the boundary states $(n_0, 0, n_2)$:

$$D_1T_1v(n_0,0,n_2) = \max[v(n_0+1,0,n_2)+R_1, v(n_0,1,n_2)] - v(n_0,0,n_2) \qquad (4.7)$$

$$D_1T_1v(n_0+1,0,n_2) = \max[v(n_0+2,0,n_2)+R_1, v(n_0+1,1,n_2)] - v(n_0+1,0,n_2) \quad (4.8)$$

We need to show that (4.8)-(4.7) ≤ 0. There are two possibilities for the outcome of the max function in (4.8). If (4.8) is $v(n_0+2,0,n_2)+R_1-v(n_0+1,0,n_2)$, then the only possible outcome for (4.7) is $v(n_0+1,0,n_2)+R_1-v(n_0,0,n_2)$ (since $D_{0,-1}v \downarrow n_0$), and the fact that (4.8)-(4.7) ≤ 0 can be shown by following the arguments in Case(2a). On the other hand, if the outcome of (4.8) is $v(n_0+1,1,n_2) - v(n_0+1,0,n_2)$, then there are two possibilities for the outcome of (4.7). The two resulting cases are similar to cases (3b) and (3c) analyzed above with $n_1 = 0$.

We will next show that $D_1T_2v(n_0,n_1,n_2) \downarrow n_0$. For any (n_0,n_1,n_2) such that $n_2 > 0$,

$$\begin{aligned} D_1T_2v(n_0,n_1,n_2) &= \max[v(n_0+1,n_1+1,n_2-1)+R_2, v(n_0,n_1+1,n_2)] \\ &\quad - \max[v(n_0+1,n_1,n_2-1)+R_2, v(n_0,n_1,n_2)], \qquad (4.9) \end{aligned}$$

$$\begin{aligned} D_1T_2v(n_0+1,n_1,n_2) &= \max[v(n_0+2,n_1+1,n_2-1)+R_2, v(n_0+1,n_1+1,n_2)] \\ &\quad - \max[v(n_0+2,n_1,n_2-1)+R_2, v(n_0+1,n_1,n_2)] \ (4.10) \end{aligned}$$

and we require (4.10)-(4.9) ≤ 0. There are three possible outcomes for the two max functions in (4.10) as the outcome $v(n_0+2,n_1+1,n_2-1)+R_2-v(n_0+1,n_1,n_2)$ is not feasible due to $D_{0,-2}v \downarrow n_1$.

Case(1): If the outcome of (4.10) is $v(n_0+2,n_1+1,n_2-1)-v(n_0+2,n_1,n_2-1)$, then the only possible outcome for (4.9) is $v(n_0+1,n_1+1,n_2-1)-v(n_0+1,n_1,n_2-1)$

(recall $D_{0,-2}v \downarrow n_0$). (4.10) – (4.9) results in $D_1v(n_0+2, n_1, n_2-1) - D_1v(n_0+1, n_1, n_2-1)$, and this is ≤ 0 by $D_1v \downarrow n_0$.

Case(2): If $v(n_0+1, n_1+1, n_2) - v(n_0+2, n_1, n_2-1) - R_2$ is the outcome of (4.10), $v(n_0+1, n_1+1, n_2-1) + R_2$ must be the outcome of the first max function in (4.9) since $D_{0,-2}v(n_0, n_1, n_2) \downarrow$ in the direction (1,-1,0) and $v(n_0+1, n_1, n_2-1) + R_2$ must be the outcome of the second max function in (4.9) as $D_{0,-2}v(n_0, n_1, n_2) \downarrow n_0$. Therefore ((4.10) – (4.9) equals $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1+1, n_2-1) - v(n_0+2, n_1, n_2-1) + v(n_0+1, n_1, n_2-1) - R_2$
$\leq v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1+1, n_2-1) - v(n_0+1, n_1, n_2) + v(n_0+1, n_1, n_2-1)$
$= D_2v(n_0+1, n_1+1, n_2-1) - D_2v(n_0+1, n_1, n_2-1)$, and this is ≤ 0 by $D_0v \downarrow n_1$.

Case(3): If the outcome of (4.10) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2)$, then there are three psossible outcomes for (4.9). Case(3a): If (4.9) is $v(n_0+1, n_1+1, n_2-1) - v(n_0+1, n_1, n_2-1)$, then (4.10) – (4.9) becomes $D_2v(n_0+1, n_1+1, n_2-1) - D_2v(n_0+1, n_1, n_2-1)$, and this is ≤ 0 by $D_0v \downarrow n_1$. Case(3b): If the outcome of (4.9) is $v(n_0, n_1+1, n_2) - v(n_0+1, n_1, n_2-1) - R_2$, then (4.10) – (4.9) equals $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2) - v(n_0+1, n_1, n_2) + v(n_0+1, n_1, n_2-1) - R_2$
$\leq v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1+1, n_2-1) - v(n_0+1, n_1, n_2) + v(n_0+1, n_1, n_2-1)$
$= D_2v(n_0+1, n_1+1, n_2-1) - D_2v(n_0+1, n_1, n_2-1)$, and this is ≤ 0 again by $D_0v \downarrow n_1$. Case (3c): If, on the other hand, the outcome of (4.9) is $v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2)$, then (4.10) – (4.9) is equal to $D_1v(n_0+1, n_1, n_2) - D_1v(n_0, n_1, n_2)$, and this is ≤ 0 by $D_1v \downarrow n_0$.

For the boundary states at which $n_2 = 0$, we have

$$D_1T_2v(n_0, n_1, 0) = v(n_0, n_1+1, 0) - v(n_0, n_1, 0) \tag{4.11}$$

$$D_1T_2v(n_0+1, n_1, 0) = v(n_0+1, n_1+1, 0) - v(n_0+1, n_1, 0) \tag{4.12}$$

and (4.12) − (4.11) results in $D_1v(n_0+1, n_1, 0) - D_1v(n_0, n_1, 0)$, and this is ≤ 0 by $D_1v \downarrow n_0$.

Next, we will show that $D_1T_3v(n_0, n_1, n_2) \downarrow n_0$. We will make use of the condition that $D_0v(n_0, n_1, n_2) \leq 0$ and the resulting fact that it is not optimal to idle the assembly queue as long as there are orders waiting in the queue. For any state (n_0, n_1, n_2) such that $n_0 > 0$

$$D_1T_3v(n_0, n_1, n_2) = v(n_0-1, n_1+1, n_2) - v(n_0-1, n_1, n_2), \tag{4.13}$$

$$D_1T_3v(n_0+1, n_1, n_2) = v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2), \tag{4.14}$$

and we require (4.14) − (4.13) ≤ 0. (4.14) − (4.13) equals $D_1v(n_0, n_1, n_2) - D_1v(n_0 - 1, n_1, n_2)$, and this is ≤ 0 by $D_1v \downarrow n_0$.

At the boundary states where $n_0 = 0$, we have:

$$D_1T_3v(0, n_1, n_2) = v(0, n_1+1, n_2) - v(0, n_1, n_2), \tag{4.15}$$

$$D_1T_3v(1, n_1, n_2) = v(0, n_1+1, n_2) - v(0, n_1, n_2), \tag{4.16}$$

We need to show (4.16) − (4.15) ≤ 0 and that holds since (4.16) − (4.15) $= 0$.

We next show that $D_1T_4v(n_0, n_1, n_2) \downarrow n_0$. For any state (n_0, n_1, n_2)

$$\begin{aligned} D_1T_4v(n_0, n_1, n_2) &= \max[v(n_0, n_1+2, n_2), v(n_0, n_1+1, n_2)] \\ &\quad - \max[v(n_0, n_1+1, n_2), v(n_0, n_1, n_2)] \end{aligned} \tag{4.17}$$

$$\begin{aligned} D_1T_4v(n_0+1, n_1, n_2) &= \max[v(n_0+1, n_1+2, n_2), v(n_0+1, n_1+1, n_2)] \\ &\quad - \max[v(n_0+1, n_1+1, n_2), v(n_0+1, n_1, n_2)] \end{aligned} \tag{4.18}$$

and we need to show that (4.18) − (4.17) ≤ 0. There are three possible outcomes for (4.18) as the outcome $v(n_0+1, n_1+2, n_2) - v(n_0+1, n_1, n_2)$ is not possible since $D_1v \downarrow n_1$.

Case(1): If the outcome of (4.18) is $v(n_0+1, n_1+2, n_2) - v(n_0+1, n_1+1, n_2)$, then the only possible outcome for (4.17) is $v(n_0, n_1+2, n_2) - v(n_0, n_1+1, n_2)$ since $D_1 v \downarrow n_0$. Hence, (4.18) – (4.17) $= D_1 v(n_0+1, n_1+1, n_2) - D_1 v(n_0, n_1+1, n_2)$, and this is ≤ 0 by $D_0 v \downarrow n_1$.

Case(2): If the outcome of (4.18) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1+1, n_2) (= 0)$, then $v(n_0, n_1+1, n_2)$ must be the outcome of the second max function in (4.17). There are two possibilities for the first max function in (4.17). Case(2b): If $v(n_0, n_1+1, n_2)$ is the outcome of the first max function in (4.17), then (4.18) – (4.17) $= 0$. Case(2b): If the outcome of (4.17) is $v(n_0, n_1+2, n_2) - v(n_0, n_1+1, n_2)$, then (4.18) – (4.17) results in $v(n_0, n_1+2, n_2) - v(n_0, n_1+1, n_2)$ and that is ≤ 0 due to the assumptions of this case.

Case(3): Finally, if the outcome of (4.18) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2)$, the only possible outcome for the first max function in (4.17) is $v(n_0, n_1+1, n_2)$ since $D_1 v \uparrow$ in the direction (1,-1,0), equivalently stated as the condition $D_{0,-1} v \uparrow n_1$. Case(3a): If the outcome of (4.17) is $v(n_0, n_1+1, n_2) - v(n_0, n_1+1, n_2).(= 0)$, then (4.18) – (4.17) $= v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2) \leq 0$ by the case assumption. Case(3b): If, on the other hand, the outcome of (4.17) is $v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2)$, then (4.18) – (4.17) $= D_1 v(n_0+1, n_1, n_2) - D_1 v(n_0, n_1, n_2)$, and this is ≤ 0 by $D_0 v \downarrow n_1$.

Next, we will show that $D_1 T_5 v(n_0, n_1, n_2) \downarrow n_0$. For any state (n_0, n_1, n_2)

$$\begin{aligned} D_1 T_5 v(n_0, n_1, n_2) &= \max[v(n_0, n_1+1, n_2+1), v(n_0, n_1+1, n_2)] \\ &\quad - \max[v(n_0, n_1, n_2+1), v(n_0, n_1, n_2)] \end{aligned} \tag{4.19}$$

$$\begin{aligned} D_1 T_5 v(n_0+1, n_1, n_2) &= \max[v(n_0+1, n_1+1, n_2+1), v(n_0+1, n_1+1, n_2)] \\ &\quad - \max[v(n_0+1, n_1, n_2+1), v(n_0+1, n_1, n_2)] \end{aligned} \tag{4.20}$$

and we require (4.20) – (4.19) ≤ 0. The outcome $v(n_0+1, n_1+1, n_2+1) - v(n_0+1, n_1, n_2)$ is not possible for (4.20) as $D_2 v \downarrow n_1$. Three possible cases remain.

Case(1): If $v(n_0+1, n_1+1, n_2+1) - v(n_0+1, n_1, n_2+1)$ is the outcome of (4.20), then the only possible outcome for (4.19) is $v(n_0, n_1+1, n_2+1) - v(n_0, n_1, n_2+1)$ due to $D_2 v \downarrow n_0$. Then (4.20)–(4.19) becomes $D_1 v(n_0+1, n_1, n_2+1) - D_1 v(n_0, n_1, n_2+1)$ and that is ≤ 0 by $D_1 v \downarrow n_0$.

Case(2): If the outcome of (4.20) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2+1)$, then the only possible outcome for (4.19) is $v(n_0, n_1+1, n_2+1) - v(n_0, n_1, n_2+1)$ due to $D_2 v \downarrow n_0$ and $D_2 v \downarrow$ in the direction (1,-1,0), equivalently stated as the condition $D_{0,-1} v \downarrow n_2$. Hence, (4.20) – (4.19) results in

$$v(n_0+1, n_1+1, n_2) - v(n_0, n_1+1, n_2+1) - v(n_0+1, n_1, n_2+1) + v(n_0, n_1, n_2+1)$$
$$\leq v(n_0+1, n_1+1, n_2) - v(n_0, n_1+1, n_2+1) - v(n_0+1, n_1, n_2) + v(n_0, n_1, n_2+1)$$
$= D_{0,-2} v(n_0, n_1+1, n_2+1) - D_{0,-2} v(n_0, n_1, n_2+1)$, and this is ≤ 0 by $D_{0,-2} v \downarrow n_1$.

Case(3): If the outcome of (4.20) is $v(n_0+1, n_1+1, n_2) - v(n_0+1, n_1, n_2)$, then there are three possible outcomes for (4.19). Case(3a): If (4.19) is $v(n_0, n_1+1, n_2+1) - v(n_0, n_1, n_2+1)$, then (4.20) – (4.19) yields to $D_{0,-2} v(n_0, n_1+1, n_2+1) - D_{0,-2} v(n_0, n_1, n_2+1)$, and this is ≤ 0 by $D_{0,-2} v \downarrow n_1$. Case(3b): If the outcome of (4.19) is $v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2+1)$, then (4.20) – (4.19) becomes $v(n_0+1, n_1+1, n_2) - v(n_0, n_1+1, n_2) - v(n_0+1, n_1, n_2) + v(n_0, n_1, n_2+1)$
$\leq v(n_0+1, n_1+1, n_2) - v(n_0, n_1+1, n_2+1) - v(n_0+1, n_1, n_2) + v(n_0, n_1, n_2+1)$
and that is ≤ 0 following Case(3b). Case(3c): Finally, if the outcome of (4.19) is $v(n_0, n_1+1, n_2) - v(n_0, n_1, n_2)$, then (4.20) – (4.19) equals $D_1 v(n_0+1, n_1, n_2) - D_1 v(n_0, n_1, n_2)$ and that is ≤ 0 by $D_1 v \downarrow n_0$.

The operator T, by definition, is formed by (i) addition and multiplication of positive constants with the functions $T_1 v$ through $T_5 v$ that are shown to be $\downarrow n_0$

and (ii) linear inventory holding and assembly queue backorder costs. Therefore, $D_1 T v \downarrow n_0$ as well. $\square$

Next, we need to show that $v^\star \in V$. Consider a value iteration algorithm to solve the optimal policy for which the inital values $v(n_0, n_1, n_2) = 0$ are used for every state $(n_0, n_1, n_2) \in P$. Conditions (i)-(ix) are satisfied by $v_0(\mathbf{x}, y)$, hence $v_0(\mathbf{x}, y) \in V$. We apply $v_{k+1}(n_0, n_1, n_2) = T v_k(n_0, n_1, n_2)$ for $k = 0, 1, 2, ...$ to determine the relative value functions for successive iterations. Suppose now that the value functions in iteration k satisfy (i)-(ix), i.e. $v_k(n_0, n_1, n_2) \in V$. Then, Lemma 4.1 shows that $v_{k+1}(n_0, n_1, n_2)$ also satisfy (i)-(ix). Therefore $v_{k+1}(n_0, n_1, n_2) \in V$. We have $v^\star(n_0, n_1, n_2)=\lim_{k\to\infty} T^{(k)} v(n_0, n_1, n_2)$ for any $v \in V$ where $T^{(k)}$ is the k^{th} composition of the operator T. Without loss of optimality, we can add the following constraints to the original problem that we cannot admit a product demand when $\max\{R_i\} < b\, n_0/\Lambda$, and we cannot produce a component type-i when $h_i\, n_i/\Lambda > R_i$. For example, if $b_0\;\; n_0/\Lambda > \max\{R_i\}$, this suggests that the amount of backorder cost incurred until the next transition is greater than any potential revenue of R_i that would be received if the next event were a product demand arrival. Similarly, if $h_i\;\; n_i/\Lambda > R_i$, this indicates that the amount of holding cost due to a type-i component incurred during a transition epoch is greater than the potential benefits of selling the product for a revenue of R_i were the next event a demand arrival for product type-i. Hence, the problem can be coverted to a finite state and action space problem. Since the underlying Markov chain is unichain, Theorem 8.4.5 of Puterman [46] ensures that $v^\star \in V$ and the existence of a long-run average profit g which can be determined using a value iteration algorithm.

We conclude the proof by noting that conditions (i)-(v) are sufficient for the structure of the optimal policy. Condition (i) results in the optimal decision that the assembly line never stays idle as long as there are orders waiting in the queue. Conditions (ii) and (iii) imply the characteristics of the optimal production policies. For example, $D_1 \downarrow n_0$ states that if it is optimal not to produce an additional type-1 component in state (n_0, n_1, n_2), it will remain optimal not to produce an additional type-1 component when there is one more order waiting in the assembly queue, (i.e., in the state $(n_0 + 1, n_1, n_2)$). Finally, conditions (iv) and (v) imply the structure of the assembly admission policies. $\square$

CHAPTER V

Conclusions

5.1 Summary and Contributions

This dissertation focuses on the integration of dynamic production and demand management decisions for multiple products facing uncertain demands. The three chapters within the dissertation study three problem settings in the domain of revenue and supply chain management, inquiring into flexibility's role in dynamic pricing, managing exogenous demand for intermediary products, and allocation of shared resources among multiple products in a make-to-order environment.

In Chapter 2, we considered a joint mechanism of dynamic pricing and capacity flexibility to manage demand and supply for multiple products. We studied the optimal dynamic production and pricing decisions for a firm that produces two substitutable items using limited product-dedicated and flexible capacities. Our first contribution in this chapter was to provide a full characterization of the joint optimal production and pricing decisions by assuming a linear additive stochastic demand model that is commonly used in the literature. Under this demand model, we showed that the optimal production policy can be characterized by modified base-stock levels that exhibit distinct forms across two broad regions of the state-space. We presented the optimal policy by classifying the initial inventory level of a product as overstocked if the item requires no further replenishment, as moderately understocked if the available capacity is adequate to bring the inventory to a desired level, and as critically understocked if capacity is restrictive to reach the desired inventory level. Our analysis showed that when at most one item is critically understocked, the modified base-stock level for each product is described by a decreasing function of the inventory level of the other item. However, when both items are critically understocked, it is shown that the modified base-stock level for a product is characterized by an increasing function of the inventory position of both products.

Regarding the optimal pricing strategy with flexible resources, our results showed that a list price is charged for an item if it is moderately understocked. If an item is critically understocked, then a price markup that depends on both inventory levels is applied. When an item is overstocked, a price discount that depends on both inventory levels is given. Our analysis demonstrated that when inventory levels for both items are critically understocked and when the flexible capacity is simultaneously shared between products, the flexible resource resulted in an optimal pricing scheme that maintained a constant price difference between products. At such instances, dynamic pricing only served to adjust the overall level of demand for both products and not to attempt to shift demand from one product to another. Instead, the flexible capacity has been instrumental in restoring the mismatches between the desired and actual inventory level of products.

One of our most significant contributions was that we showed that the availability of a flexible resource gave rise to a certain state space region where the optimal prices charged for the items had a constant price difference. In other words, flexibility helps a firm to maintain stable price differences across items over time even when the optimal price of each item fluctuates over time.

To the best of our knowledge, our work is the first one to consider the effects of flexibility on a dynamic pricing strategy. We believe our results have favorable ramifications from a marketing standpoint as it suggests that even when a firm applies a dynamic pricing strategy, it may still establish consistent price positioning among multiple products if it can employ a flexible replenishment resource. Hence, in the presence of dynamic pricing, flexibility serves as an essential tool to preserve customer's valuations of products over the long run.

In addition, we investigated the economic benefits of a joint strategy versus applying each tool individually. Our results indicate that dynamic pricing and capacity flexibility can be viewed as substitute, but not fully interchangeable approaches. Moreover, we found that dynamic pricing is a more powerful tool if demands are positively correlated while flexibility provides much of the benefits when demands are negatively correlated.

In Chapter 3, we studied a manufacturing firm that assembles a single end-product from many intermediate components. The firm experiences demands for its end-product as well as for any of the intermediary products. The main considerations are the firm's dynamic decisions on how to decide on the production of each component as to coordinate the assembly process and satisfy the exogenous demand for components, when to initiate an assembly operation to convert intermediate components into end-products, and how to set admission rules for demands targeted at intermediary and end-products.

For a general system composed of an arbitrary number of intermediate components, we showed that demand admission for the product and for any of the intermediate products are characterized by state-dependent rationing and admission threshold levels. For example, a demand for an intermediate product will be accepted only if there is a sufficient number of units of that component in inventory. If there are fewer units, the optimal action is to reject the demand and save the component for assembly purposes. We also identified an admission threshold implying that when the assembly queue length gets larger, it may be optimal to reject a demand for the end-product. For component production and assembly decisions, we showed that state-dependent base-stock levels are optimal. That is, if the number of items in inventory is below a certain base-stock level, it is optimal to produce further units of that item. In addition, for each decision type, we showed how the state-dependent thresholds depend on the inventory positions of the other items.

We explored the sensitivity properties of the optimal policy to various problem parameters. We showed that as the end product revenue decreases, the optimal strategy is to accept fewer demand for the end-product and more demand for the intermediate components as well as to produce fewer units of each component and to assemble fewer units of the end-product.

We also provided two extensions for the basic model, one concerning multiple customer classes based on revenue in addition to the classes based on the type of item they request, and the other, investigating the effects of a partial payment scheme on the optimal policy structure. We characterized the structure of the optimal policy for each of these extensions.

Finally, since the optimal policies were rather complex and defined by switching surfaces in a multidimensional space, we also introduced a novel heuristic policy. We tested our heuristic policy against the optimal solution as well as a commonly applied basic heuristic policy. Our heuristic policy has performed better in every single instance compared to the basic heuristic policy. It has also attained profits very close to that obtained by the optimal policy.

Lastly, in Chapter 4, we studied a basic mass customization setting by focusing on a manufacturing firm that assembles two different types of products from two different components. An arriving customer choses an item and the order is assembled only after the order is received and admitted to the assembly queue. We were particularly interested in the manufacturer's decisions on how to decide on the number of component inventories to hold and how to prioritize the different customer classes (based on their product choice and revenue) that share the assembly resource.

We partially characterized the structures of the optimal component production and assembly admission policies. We showed that the optimal production decisions can be defined by a state-dependent base-stock policy and the optimal demand admission rules are defined by a state dependent admission threshold policy. We performed numerous computational tests and conjectured that the partially characterized optimal policy structure over the analyzed specific region extends to the entire state space. Structural results in general assemble-to-order systems are still unknown and we believe our contribution of providing a partial characterization could prove as a building block to derive results for more general systems. Finally, we devised a heuristic solution algorithm that is motivated by the insights we gained from the optimal policy structure. We tested the performance of the algorithm for a number of problem instances with varying problem parameters and number of end-products. We showed that the heuristic policy was very effective in attaining near optimal results.

5.2 Extensions

There is an exciting research potential in the revenue and supply chain management area nourished by rapidly emerging new technology and business practices.

Several interesting research problems for further investigation are presented below:

1. Dynamic Pricing with Lead-time Differentiated Customers:

   Customers may have different preferences regarding the delivery time of a product and a class of customers may even be willing to pay more for a delayed delivery mode. For example, BMW serves customers through a traditional make-to-stock selling channel as well as a make-to-order system in which a vehicle is assembled to the exact specifications requested by a customer. Customers who value a customized product may be willing to wait even at a higher price. Different classes of customers may have different preferences for immediate or delayed but personalized service. In addition, the customers' utilities for either type of service may be influenced by the price and length of the expected delivery period. The decisions given by a manufacturer or service provider regarding how to set the prices for both service types and sequence the capacity to satisfy each class may be a significant contributor to profitability and would constitute an interesting research problem.

2. Dynamic Pricing and Production Control with Consumer Upgrades:

   Along with dynamic pricing, consumer upgrades to a higher quality product also serve as a valuable tool in revenue management to align supply with demand across multiple products. In an immediate research paper, we consider joint price-based and availability-based substitutions by studying a multiple period, two-stage model where in the first stage the firm sets prices and the production targets while the demand is still uncertain, and in the second stage, after the demand is observed, it decides how much (if any) of the customers to upgrade to the higher quality product. The preliminary analysis indicates that a threshold type policy is optimal for the second stage product upgrade decisions. Further research to reveal the complete structure of the optimal policy, devising an easily implementable heuristic, and understanding the economic benefits of product upgrades in a dynamic pricing setting may provide important insights into the operations of firms in various manufacturing and service industries that adapt a customer upgrading strategy.

3. Product Upgrades with Strategic Consumers:

   When considering product upgrades, a natural extension is to adapt a consumer choice model where the customers may act strategically anticipating a potential upgrade offer by the firm. Product upgrades improve customer satisfaction, which in turn improves the loyalty of the customer to the brand. On the other hand, the customers may choose the lower quality product hoping that they will be upgraded to the high quality product, which cannibalizes the demand for the high quality product. Setting optimal prices and upgrade limits, in addition to the capacity levels constitute a challenging and interesting problem.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] N. Agrawal and M.A. Cohen. Optimal material control in an assembly system with component commonality. *Naval Research Logistics*, 48:409–429, 2001.

[2] Y. Akcay and S.H. Xu. Joint inventory replenishment and component allocation optimization in an assemble-to-order system. *Management Science*, 50:99–116, 2004.

[3] G. Aydin and E. Porteus. Joint inventory and pricing decisions for an assortment. *Operations Research*, 56(5):1247–1255, 2008.

[4] S. Benjaafar and M. ElHafsi. Production and inventory control of a single product assemble-to-order systems with multiple customer classes. *Management Science*, 52(12):1896–1912, 2006.

[5] S Benjaafar, M. ElHafsi, C.Y. Lee, and W. Zhou. Optimal control of an assembly system with multiple stages and multiple demand classes. Working paper, University of Minnesota, Mar 2008.

[6] S. Biller, L.M.A. Chan, D. Simchi-Levi, and J.L. Swann. Dynamic pricing and the direct-to-customer model in the automotive industry. Working paper, Massachussetts Institute of Technology, 2005.

[7] R. Billinton and R.N. Allan. *Reliability Evaluation of Engineering Systems: Concepts and Techniques*. Plenum Press, New York, 1983.

[8] G. Bitran and R. Caldentey. Pricing models for revenue management. *Manufacturing & Service Operations Management*, 5:203–229, 2003.

[9] S. Carr and I. Duenyas. Optimal admission control and sequencing in a make-to-stock/make-to-order production system. *Operations Research*, 48(5):709–720, Oct 2000. This paper is magnificent.

[10] O. Ceryan, I. Duenyas, and Y. Koren. Control of a two-stage assemble-to-order manufacturing system under demand and capacity uncertainties. Proceedings of the ASME Conference on Engineering Systems Design and Analysis, July 2008.

[11] O. Ceryan, I. Duenyas, and Y. Koren. Optimal control of an assembly system with demand for the end-product and intermediate components. Technical report, University of Michigan, 2009.

[12] O. Ceryan and Y. Koren. Manufacturing capacity planning strategies. *Annals of the CIRP-Manufaturing Technology*, 58:403–406, 2009.

[13] O. Ceryan, O. Sahin, and I. Duenyas. Managing demand and supply for multiple products through dynamic pricing and capacity flexibility. Technical report, University of Michigan, 2009.

[14] L.M.A. Chan, Z.J.M Shen, D. Simchi-Levi, and J.L. Swann. *Coordination of Pricing and Inventory Decisions: A Survey and Classification*. Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era. Kluwer Academic Publisher, Boston, 2004.

[15] X. Chen and D. Simchi-Levi. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Mathematics of Operations Research*, 29(3):698–723, 2004.

[16] J. Chod and N. Rudi. Resource flexibility with responsive pricing. *Operations Research*, 53(3):532–548, 2005.

[17] M.A. Cohen, N. Agrawal, and V. Agrawal. Winning in the aftermarket. *Harvard Business Review*, May, 2006.

[18] A. Copeland, W. Dunn, and G. Hall. Prices, production and inventories over the automotive model year. NBER Working paper, Board of Governors of the Federal Reserve System, 2005.

[19] G. DeCroix and A. Arreola Risa. Optimal production and inventory policy for multiple products under resource constraints. *Management Science*, 44(7):950–961, 1998.

[20] I. Duenyas and C. Tsai. Centralized and decentralized control of a two stage tandem manufacturing system with demand for intermediate and end products. Working paper, University of Michigan, Aug 2001.

[21] W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003.

[22] R. Evans. Inventory control of a multiproduct system with a limited production resource. *Naval Research Logistics*, 14(2):173–184, 1967.

[23] A. Federgruen and A. Heching. Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3):454–475, 1999.

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

[25] Y. Gerchak and M. Henig. Component commonality in assemble-to-order systems: Models and properties. *Naval Research Logistics*, 36:61–68, Aug 1989. Working paper, University of Michigan.

[26] H. Ghoneim and S. Stidham. Control of arrivals to two queues in series. *European Journal of Operational Research*, 21:399–409, Jan 1985.

[27] P. Glasserman. Allocating production capacity among multiple products. *Operations Research*, 44(5):724–734, 1996.

[28] M. Goyal, S. Netessine, and T. Randall. Deployment of manufacturing flexibility: An empirical analysis of the north american automotive industry. Working paper, University of Maryland, 2006.

[29] A. Ha. Optimal dynamic scheduling policy for a make-to-stock production system. *Operations Research*, 45:42–53, 1997.

[30] A.Y. Ha. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43(8):1093–1103, Jan 1997.

[31] A.Y. Ha. Optimal dynamic scheduling policy for a make-to-stock production system. *Operations Research*, 45:42–53, Jan 1997.

[32] W.H. Hausman, H.L. Lee, and A.X. Zhang. oint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, 109:646–659, 1998.

[33] D.P. Heyman and M.J. Sobel. *Stochastic Models*, volume 2 of *Handbooks in Operations Reseach & Management Science*. Elsevier Science Publishers, BV, 1990.

[34] G. Koole. Monotonicity in markov reward and decision chains: Theory and applications. *Foundation and Trends in Stochastic Systems*, pages 1–76, 2006.

[35] Y. Koren, U. Heisel, F. Jovane, T. Moriwaki, G. Pritschow, G. Ulsoy, and H. Van Brussel. Reconfigurable manufacturing systems. *Annals of the CIRP-Manufaturing Technology*, 48(2):527–540, 1999.

[36] K.L. Kraemer, J. Dedrick, and S. Yamashiro. Refining and extending the business model with information technology: Dell computer corporation. *The Information Society*, 16:5–21, 2000.

[37] C.Y. Ku and S. Jordan. Access control to two multiserver loss queues in series. *IEEE Transactions on Automatic Control*, 42(7):1017–1023, Jan 1997.

[38] C.Y. Ku and S. Jordan. Access control of parallel multiserver loss queues. *Perform Evaluation*, 50:219–231, Jan 2002.

[39] Q. Li and S. Zheng. Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research*, 54(4):696–705, 2006.

[40] S.A. Lippman. Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23:687–710, 1975.

[41] Y. Lu and J.S. Song. Order-based cost optmization in assemble-to-order systems. *Operations Research*, 53(1):151–169, 2005.

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

[43] P. Naor. The regulation of queue size by levying tolls. *Econometrica*, 37(1):15–24, 1969.

[44] W. Narongwanich, I. Duenyas, and J.R. Birge. Optimal portfolio of reconfigurable and dedicated capacity under uncertainty. Working paper, University of Michigan, 2002.

[45] E.L. Plambeck and A.R. Ward. A separation principle for assemble-to-order systems with expediting. Working paper, Stanford University, 2003.

[46] P.L. Puterman. *Markov Decision Processes*. Wiley, New York, 1994.

[47] T. Robertazzi. *Computer networks and systems: queueing theory and performance evaluation*. Springer-Verlag, New York, 2000.

[48] K. Rosling. Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37:565–579, Aug 1989.

[49] A. Sahay. How to reap higher profits with dynamic pricing. *MIT Sloan Management Review*, 48(4):53–60, 2007.

[50] C. Schmidt and S. Nahmias. Optimal policy for a two-stage assembly system under random demand. *Operations Research*, 33:1130–1145, Oct 1985.

[51] J.S. Song, S.H. Xu, and B. Liu. Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes. *Operations Research*, 47:131–149, 1999.

[52] J.S. Song and Z. Xue. Demand management and inventory control for substitutable products. Working Paper, Duke University, 2007.

[53] J.S. Song and P. Zipkin. *Supply Chain Operations: Assembly-to-Order Systems.*, volume 11 of *Handbooks in Operations Reseach & Management Science*. Elsevier Science Publishers, BV, 2003.

[54] S. Stidham. Optimal control of admission to a queueing system. *IEEE Transactions on Automatic Control*, 33(8):705–713, 1985.

[55] D. Topkis. Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Science*, 15:160–176, 1968.

[56] T. Whitin. Inventory control and price theory. *Management Science*, 2(1):61–68, 1955.

[57] Y. Ye. Joint pricing and inventory decision for competitive products. Master's thesis, Massachussetts Institute of Technology, 2008.

[58] K. Zhu and U. Thonemann. Coordination of pricing and inventory control across products. *Naval Research Logistics*, 56(2):175–190, 2009.