ABSTRACT

VILA-PARRISH, ANA RAQUEL. Dynamic Inventory Policies for Short Lifecycle and Perishable Products with Demand Uncertainty. (Under the direction of Julie Simmons Ivy and Russell E. King.)

This dissertation considers a common thread of inventory modeling for short lifecycle and perishable products. These nontraditional products introduce additional elements of complexity, particularly when the demand processes for these goods are nonstationary. We motivate our research with real world context from the consumer electronics and healthcare industries.

In Chapter 2, we consider a firm with a geographically distant supplier that simultaneously decides the order quantity for each of two supply modes (e.g., air shipment and ocean shipment) on a periodic basis. In each period, the firm may order from both sources, order from one source, or order nothing. This research extends the literature on two modes of supply by deriving the optimal solutions under nonstationary product lifecycle (PLC) dependent demand, with lead time differences between the two modes of one or more periods, and a finite sales horizon. This fluctuating demand environment is a distinctive feature of our model since both the demand mean and demand variability may change over time representing the PLC. We develop a Markov decision process (MDP) which determines the optimal ordering quantities per mode under stochastic, nonstationary demand. We consider two cases: Case 1 models demand evolution as an embedded Markov chain, and Case 2 considers the special case where demand evolution is time-dependent. Optimal base

stock policies, defined by two base stock levels, are derived and shown to be both state and time-dependent for Case 2. Our results suggest that optimal shipping strategies (i.e., the distribution of total order quantities between each mode of supply) change throughout the PLC as demand characteristics change. We derive insights from the optimal structure of the time-dependent demand case for the general model. Further, we use numerical examples to explore the relationship between the optimal solutions and the phases of the PLC.

In Chapters 3 and 4, we model and analyze production and inventory policies for perishable pharmaceuticals in the context of the hospital pharmacy. Hospital pharmacies throughout the United States are experiencing drug inventory problems that result in waste and shortages that affect patient outcomes due to delayed procedures and drug substitutions. We consider a pharmacist's decision to order and hold drug products in two inventory stages (i.e., raw material and dispensed form) with varying shelf lives.

In Chapter 3, our research objective is to determine the optimal two-stage inventory and production policy. Since demand for medications is dependent on the patient mix, we define a stochastically changing 'demand state' as a surrogate for patient condition information that is represented by a Markov chain. We evaluate two demand fulfillment scenarios when a shortage occurs which depends on production capabilities. We formulate this problem as a Markov decision problem (MDP) and prove the existence of optimal solutions for both scenarios. Finally, we present numerical examples which demonstrate the behavior of the optimal solutions for various cost structures. We apply this model in Chapter 4 using demand data from a large public hospital. We develop a Markovian demand process and solve for the optimal raw material inventory order and finished good production quantities.

Dynamic Inventory Policies for Short Lifecycle and Perishable Products with Demand Uncertainty

by
Ana Raquel Vila-Parrish

A dissertation proposal submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Industrial Engineering

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Russell E. King	Dr. Julie Simmons Ivy
Chair of Advisory Committee	Co-Chair of Advisory Committee
Dr. Thom J. Hodgson	Dr. Jing-Sheng Song

BIOGRAPHY

Ana Raquel Vila-Parrish was born in Boynton Beach, Florida on August 1, 1979. After relocating to North Carolina in 1995, she graduated from Cary High School and attended North Carolina State University from 1997-2001, earning a Bachelor of Science degree in Industrial Engineering. Following graduation, she and her husband moved to Austin, TX to begin careers at Dell. Her job as a New Product Operations Engineer afforded her the opportunity to travel to Asia and Europe to support manufacturing during new product development. She was awarded a Master of Industrial Engineering at Texas A&M while working fulltime. It was during this time that she decided to leave her job at Dell and pursue further education. After receiving a Master of Art from Duke University, she began her PhD in Industrial and Systems Engineering at North Carolina State University. She received a dissertation research award from the Agency of Healthcare Research and Quality (AHRQ) for her work in inventory and production modeling for perishable pharmaceutical products. Her research interests are in application of decision modeling to supply chain, inventory, and healthcare systems.

ACKNOWLEDGEMENTS

I would sincerely like to thank my advisors Dr. Julie Ivy and Dr. Russell King for allowing me the freedom to find problems that interested me. I truly appreciated the ability to brainstorm and laugh with you both. I would also like to thank my committee members for their valuable comments and suggestions. I would also like to recognize Clarence Smith who has been a mentor to me for the last 10+ years- thanks for all the great ISE memories.

I would like to acknowledge the Agency of Healthcare Research and Quality for providing me the opportunity and sponsorship to conduct research in the healthcare field. I would also like to recognize my clinical mentors, Doctors of Pharmacy Steve Abel and Lynn Eschenbacher who were instrumental in the problem formulation and data collection discussed in Chapters 3 and 4.

I would be remiss if I did not thank my peers (both past and present), especially those in Daniels 307. Laughter was not scarce and for that I thank you.

And finally, to my family: To Mom, Dad, and Tania- I cannot put a value on your never ending encouragement, love, and support. You have played a vital role in who I am today. To Jason- This could not have been possible without you. You and Adrian are a constant source of inspiration.

TABLE OF CONTENTS

LIST OF TABLES viii

LIST OF FIGURES ix

Introduction to Dissertation Research 1

1.1 Introduction 1

1.2 Literature Review for Two Modes of Supply 2

TWO MODES OF SUPPLY 3

MARKOV DEMAND PROCESS 7

1.3 Literature Review for Multi-echelon Pharmaceutical Perishable Inventory 9

SINGLE STAGE PERISHABLE INVENTORY 10

MULTI-ECHELON PERISHABLE INVENTORY 12

1.4 Dissertation Organization 14

The Impact of the Product Lifecycle on Shipment Strategy with Two Modes of Supply 15

2.1 Introduction 16

2.2 Relevant Literature 20

2.3 Problem Description and MDP Formulation 23

2.3.1 MDP Notation 24
2.3.2 Case 1: Uncertainty in the Demand Evolution Process 25
2.3.3 The State Space 26
2.3.4 The Decision Space 26
2.3.5 Case 1: State Transition and Transition Probabilities 27
2.3.6 Case 1: Cost Function and Optimization 28
2.3.7 Case 2: A Special Case When Demand Evolution Is Certain 30
2.4 Optimal Policy Structure 31
2.4.1 Numerical Example 34
2.4.2 Slow Shipment Modes with Lead Time Greater Than One 38
2.5 Implications for the Case of Unknown Transitions 41
2.5.1 Extension of Results from Section 2.4 41
2.5.2 Numerical Example 42
2.6 Conclusions 44
Appendix A 45
A Two-Stage Inventory and Production Model for Perishable Products with Markovian Demand under Various Demand Fulfillment Scenarios 57
3.1 Motivation 58

3.2 Relevant Literature ... 60

3.3 Model Specification ... 63

3.4 Model Assumptions and State Specification ... 65

3.4.1 MDP Cost Formulation for Fulfillment Scenario 1: Demand Fulfilled via External Supplier ... 67

3.5 Optimal Ordering and Production Policy ... 70

3.5.1 Fulfillment Scenario 2: Demand Fulfilled Internally First ... 75

3.6 Numerical Examples ... 78

3.6.1 Impact of Expediting Costs ... 79

3.6.2 Impact of Waste Costs ... 81

3.6.3 Interrelationship Between R and x ... 82

3.7 Conclusion ... 83

Appendix B ... 84

Stochastic Multi-Echelon Perishable Inventory Models for Meropenem Management ... 87

4.1 Introduction ... 87

4.2 Data Analysis ... 90

4.3 Specification of Demand Process ... 93

4.4 Model Formulation ... 98

4.5 Model Preliminaries ... 100

4.5.1 Dynamic Demand Process ... 101

4.5.2 Impact of Expediting Costs ... 103

4.5.3 Cost Impact of Ignoring Patient Population Dynamics ... 106

4.6 Discussion ... 108

Research Contribution and Future Work ... 111

5.1 Research Contribution of Chapter 2 ... 111

5.2 Research Contribution of Chapters 3 and 4 ... 113

Bibliography ... 115

LIST OF TABLES

Table 2-1. Parameter values for the numerical examples 36

Table 2-2. Summary of demand parameters for the numerical experiments 36

Table 3-1: Baseline costs for both internal and external fulfillment scenarios 79

Table 4-1. Clinical indications and corresponding dosing for Meropenem (*FDA approved indications) (Micromedex, 2009) 89

Table 4-2: Summary statistics for total daily dose using the empirically derived demand distributions which correspond to the number of patients in the system. 96

Table 4-3. Transition probabilities for number of patients in the system 97

Table 4-4. Baseline cost parameters (in dollars) 100

Table 4-5. Example of a optimal production decision x^* given the initial state is (z=2, R=35) assuming the realized demand is 4.5 grams 102

Table 4-6. Results of sensitivity analysis on the cost of internal and external expediting, the values noted are the production and raw material quantity (in grams). 105

Table 4-7. Percent change in total expected cost of the dynamic model evaluated with the policies derived by assuming that the demand model is stationary. 108

LIST OF FIGURES

Figure 1-1: A Summary of the Literature Review for Chapter 2 3

Figure 1-2: Summary of Literature Review for Chapters 3 and 4 9

Figure 2-1: State transition diagram for demand including life cycle phases, Markov demand process (states), and corresponding demand characteristics 17

Figure 2-2: A six-state example of the Markovian demand process which represents a typical product lifecycle (PLC)... 26

Figure 2-3: Timeline of events and decisions during period *t* 27

Figure 2-4: Demand distributions with μ=5 and CV = 0.5 and 0.84 35

Figure 2-5: Order quantities for a lifecycle with 3 phases and known transition times for each phase when the state is at a backorder level of 3, (i.e., x = -3). 36

Figure 2-6: Order quantities for a lifecycle with 3 phases and known transition times for each phase when the state is at an inventory level of 3, (i.e., x = +3). 36

Figure 2-7: (Slow Evolution Case, P_1) Optimal order quantities when backlog =3 and demand phase = 2 43

Figure2-8: (Fast Evolution Case, P_2) Optimal order quantities when backlog =3 and demand phase = 2 43

Figure 3-1: Raw material inventory depletion throughout a single cycle.............................. 68

Figure 3-2: Timeline of events for finished goods inventory during a period........................ 68

Figure 3-3: The impact of external fulfillment costs on the optimal initial raw material order quantity (denoted by the black diamond marker) 80

Figure 3-4: Comparison of optimal total expected cost and optimal initial raw material decision (*R*, the diamond marker) as a function of fulfillment scenario and expediting costs. 81

Figure 3-5: Effect of the initial raw material inventory decision on the production decision in period 1. 82

Figure 4-1. Total proportion of dose level used per encounter (with Meropenem use) in the PICU 91

Figure 4-2: Number of patients who are administered the drug over a span of 1 to 10 days 91

Figure 4-3. Total dose of Meropenem administered per day........ 92

Figure 4-4a-e: Empirical probability distributions for total daily dose given there are 1,2,3,4,or 5 candidate patients in the PICU 94

Figure 4-5: Probability mass function for total daily dose in grams, ignoring the number of actual patients in the system 107

CHAPTER 1:

Introduction to Dissertation Research

1.1 Introduction

In this dissertation we consider the complexities in the areas of supply chain and inventory management under demand uncertainty that arise in two separate but related research problems. Specifically, we are interested in finding the optimal inventory and ordering policies for (1) A two supply mode decision problem for products with short product lifecycles (PLC) and (2) A two-stage perishable inventory problem. While these may seem like unrelated areas, we can exploit their natural commonalities. First, a product with a short lifecycle (i.e. less than two years) can be viewed as a perishable product since its value declines at the end of or throughout its sales life. Second, demand uncertainty impacts these types of products to a greater degree than a traditional product. The impact of excess inventory for perishable products equates to waste due to expiration and obsolescence in the case of short lifecycle products. Third, in practice, stationary base stock policies are used to manage both types of products and complexities are often ignored leaving opportunities for cost improvements.

The research questions and contexts presented in this dissertation are inspired by real world problems. The two modes of supply problem is motivated by experiences in the high tech industry where demand evolution throughout the product lifecycle poses increased uncertainty in ordering and inventory decisions. In addition, we investigate the two stage perishable inventory and production decision problem in the context of a hospital pharmacy. The pharmacy inventories raw materials and dispensed forms (e.g. intravenous) of up to 2,000 drugs. We apply a theoretical model to a real problem: the inventory and production management of the drug Meropenem.

In the following sections we review the relevant literature for Chapter 2-4 and briefly discuss the contribution of this dissertation. A more detailed discussion is presented in Chapter 5 and within each Chapter.

1.2 Literature Review for Two Modes of Supply

Figure 1-1 below presents a summary of existing literature on two modes of supply and Markovian demand processes and highlights the contribution of this dissertation research. We trace the tree by subject in the discussion that follows ending with our contribution to the state of the literature.

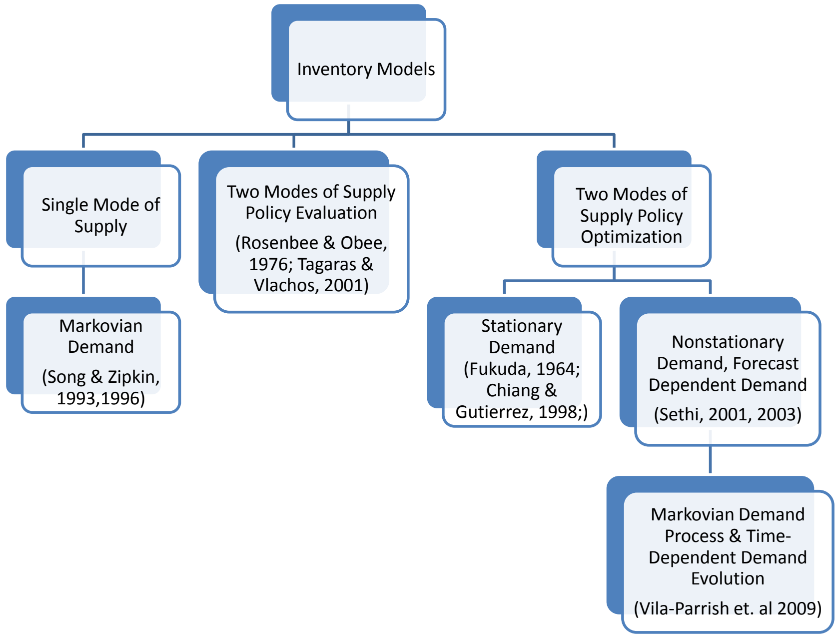


Figure 1-1: A Summary of the Literature Review for Chapter 2

<u>**TWO MODES OF SUPPLY**</u>

The existing literature which considers the two modes of supply problem can be divided into two categories: those that evaluate the ordering decisions and those that optimize the ordering decisions. A review of two modes of supply literature including both evaluation and optimization can be found in Minner (2003). We briefly discuss selected policy evaluation literature then we focus on the optimization literature as our model falls into this group.

Policy Evaluation

Rosenshine and Obee (1976), Moinzadeh and Nahmias (1988), and Tagaras and Vlachos (2001) are among those that consider two modes of supply policy evaluation rather than derive an optimal inventory policy. Rosenshine and Obee (1976) consider a standing order inventory system where a constant order via the regular shipment method is received each period and an emergency order of fixed size may be placed each period. Moinzadeh and Nahmias (1988) analyze the two modes problem using continuous review and an extension of the (s, Q) ordering policy with parameter selection which minimizes long run average costs. Lastly, Tagaras and Vlachos (2001) use a base stock level type policy for the regular mode. Each period a quantity via the regular mode is ordered in order to reach this level while the emergency mode may be used at specific review times which do not occur at the regular review cycle time. Our work is distinct since we consider optimization of policies and do not assume a policy structure.

Optimization: Stationary, Stochastic Demand

Fukuda (1964) extended traditional (one mode) inventory models to allow for delivery of order quantities by both fast (termed emergency) and slow methods simultaneously. However, the results are limited to the case where the difference in lead time between fast and slow methods is exactly one period. Stationary base stock levels are derived for both modes of supply. Whittemore and Saunders (1977) generalize this model to consider two supply modes that differ by n review periods for a product with stochastic demand that is assumed to be independent and identically distributed over successive periods. Based on this assumption, the optimal ordering pair is independent of the time period in which the orders are placed. They derived steady state distribution of surface and air orders which resembles the base stock policy of the single ordering option. Explicit results are found for the case where the difference in lead times is one period. Extending past work to incorporate supply lead times that are not a multiple of a review period, Chiang and Gutierrez (1998) and Chiang (2003) consider a system with long review periods which are possibly longer than the supply leadtimes. In both papers, they analyze the problem in the context of an infinite horizon and derive the optimal solution. Chiang and Gutierrez (1998) derive the optimal control policy which is an extension of the optimal base stock ordering policy for the one supply mode model for the emergency mode which depends on the number of periods until the regular order arrives. Chiang (2003) includes further analysis with positive ordering costs for the emergency shipment method and conclude that an (s,S) policy is optimal in this case. Our research extends the existing literature by considering a nonstationary demand function and unknown demand transitions which mimic the uncertainty of the product lifecycle.

Optimization: Dependent, Stochastic Demand

In contrast to previous research, Sethi et al. (2003) devise a model which incorporates expected demand distributions which are dependent on forecast updates. They prove that for the case where there are fixed costs and the difference in leadtimes is one period, a (s, S) policy is optimal and depends only on the forecast update. While our environment is most similar to Sethi et al. (2003), in that demands are dependent and nonstationary, the specification of our environment is quite different.

Dissertation Research Contribution

Our research explores the impact of changing demand attributes throughout the product lifecycle of a product with a finite sales horizon, the aforementioned research does not capture this reality. We represent the demand process of our fluctuating environment by using a Markov modulated demand structure which is a more flexible model (Song and Zipkin, 1993, 1996). This modeling approach allows us to represent the dynamic demand evolution through the specification of demand distributions.

Our model incorporates these demand characteristics in order to analyze changes in logistics policies during each phase of the PLC. In addition to presenting the results from numerical experimentation for these scenarios, we prove structural properties of these policies that provide insight into the interrelationship between shipment quantities per method and the (1) demand characteristics and (2) demand evolution. In addition to extending the current literature on two modes of supply, this dissertation research extends the literature on Markov demand process. Next, we discuss the literature in this area as well as the contribution to this stream of literature.

MARKOV DEMAND PROCESS

While Markovian properties have been used to represent dynamic capacity, quality, etc., we restrict our attention to those papers which have represented fluctuating demand, which is not independently evolving over time, with Markov modulated demand. The following is a review of select literature in the area of Markovian demand process.

Song and Zipkin (1993) formulate a one mode of supply inventory policy optimization problem which uses Markov-modulated demand to represent a dynamic demand. Their work is motivated by several real-world environments such as new product introduction, projects subject to obsolescence, and projects sensitive to economic conditions. The main result of the paper is the development of quantitative and qualitative forms of the optimal policies under both linear and fixed order costs. Additionally, monotonic properties are developed which allow for an ordering of the base stock levels. Lastly, two algorithms are developed to solve for the optimal policy parameters.

Song and Zipkin (1996) apply the general framework they developed in 1993 to further explore the impact of obsolescence on inventory policies. Numerical examples compare the optimal policy to heuristic policies. It is shown that accounting for uncertainty in timing of obsolescence result in optimal policies that can be significantly different and result in significant cost savings. Treharne and Sox (2002) extend the use of Markovian demand processes by assuming that the demand state is only partially observed. They develop a partially observed Markov decision process (POMDP) in the context of a one mode of supply inventory problem with linear costs and finite horizon. The optimal policy of the POMDP while simple in structure is computationally difficult since the size of the state space includes a vector of probabilities representing the "belief" of the true state of the world. The optimal solution is then compared to several suboptimal control policies under various scenarios.

Our research enriches the literature mentioned above on Markovian demand processes. Like Song and Zipkin (1996) we pay special attention to specifying the demand submodel. However, we consider the entire product lifecycle which is not only inclusive of the obsolescence process but also of the product ramp process. This end to end view of the lifecycle will extend the knowledge of the inventory control policies which depend on the state of the product lifecycle. Our model is appropriate for the demand environment of many innovative products and industries. In these environments there are uncertainties in the adoption of the product, and thus the evolution of demand, as well as in the process by which they ramp down the product in order to introduce the next generation product. As a result of this dynamic product lifecycle demand process steady-state and infinite horizon models may not be appropriate modeling methods; therefore we consider a finite horizon setting. We also extend the current research using Markovian demand processes and single mode of supply to consider two modes of supply.

1.3 Literature Review for Multi-echelon Pharmaceutical Perishable Inventory

Figure 1-2 below identifies the contribution of this research with respect to the current literature. We trace the tree by subject in the discussion that follows ending with our contribution to the state of the literature.

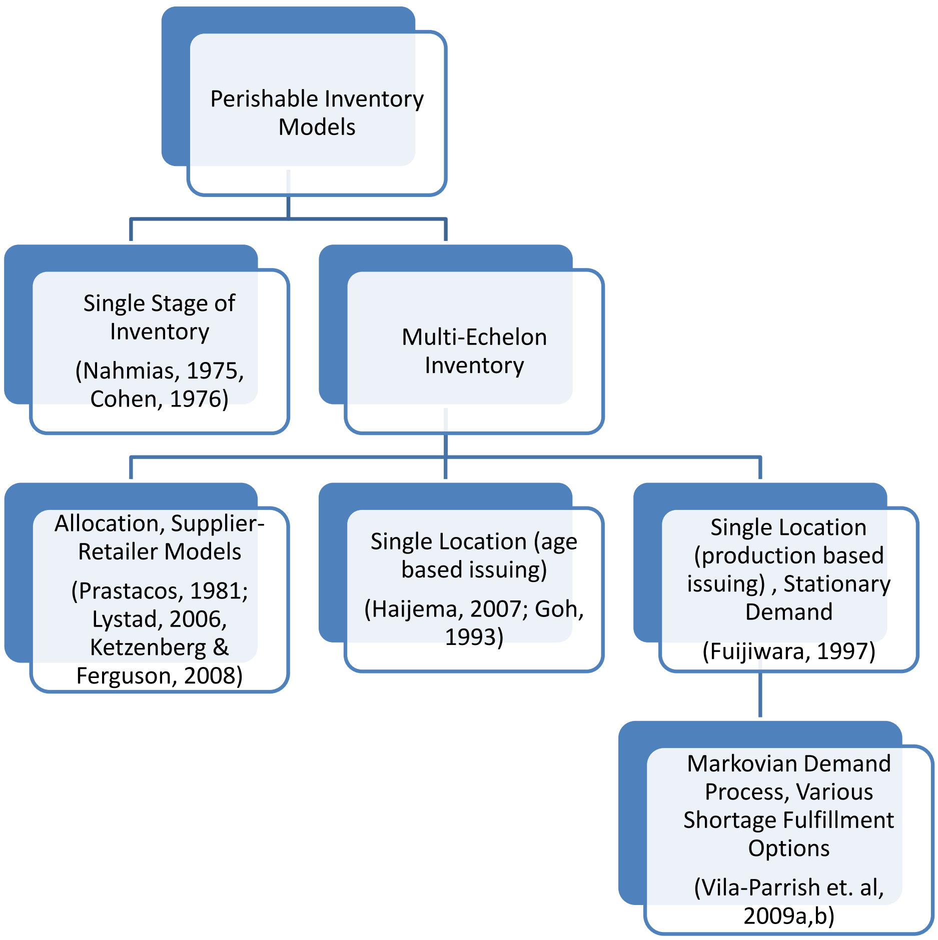


Figure 1-2: Summary of Literature Review for Chapters 3 and 4

For a complete review of perishable inventory literature, which includes deteriorating products, refer to Nahmias's (1982) comprehensive literature review or Giri and Goyal (2001) for a more recent review.

SINGLE STAGE PERISHABLE INVENTORY

Periodic Review, Deterministic Demand

Dave (1979) considers a discrete time inventory model. A deterministic model with a known and constant demand and scheduling period is prescribed for a perishable item that deteriorates as a function of time. A total average cost per unit time equation was formulated and the optimal order level, lot-size per cycle, and the number of units that deteriorate during a cycle are determined.

Periodic Review, Stochastic Demand

A general framework developed by Nahmias (1975) where the number of periods of fixed lifetime was not restricted and costs such as ordering, holding, stock out, and outdating were charged. Nahmias's analysis was in the setting of periodic review and his mode of analysis used the functional equation approach of dynamic programming. In the finite horizon dynamic problem the demand distributions were assumed to be stationary and the results hold in the nonstationary case if and only if the demands are stochastically increasing over time. Ultimately the model determines a single ordering decision of the perishable good. Since these models were mathematically intensive or intractable, subsequent work assumed a policy instead of determining the optimal policy. Many of the policies chosen had their origins as optimal solutions to nonperishable inventory problems.

In 1976, Brodheim et al. evaluated a class of inventory policies for perishable products motivated by issues in blood bank management. The class of policies is characterized by scheduled deliveries of fixed amounts of blood at regular intervals to replenish most of the blood used. The inventory age distribution forms a Markov chain and has a stationary transition probability. Using the Poisson distribution upper and lower bounds of the probability of shortage are plotted as a function of the fraction of usage replaced daily.

Extending the work of Nahmias, Cohen (1976) restricts his analysis to a stationary critical order policy that is characterized by an order up to level. The vector of the disposition of inventory is shown to be a Markov chain with a stationary demand distribution. Cohen was able to demonstrate existence of an invariant distribution for the disposition process though a closed form expression for the expected outdates in the general m-period were not derived. In parallel, Chazan and Gal (1977) also formulated an inventory model as a finite Markov chain. Instead of the fixed delivery quantity policy used by Brodheim et. al (1976), they assume a policy which maintains a constant inventory level. This model is similar to Cohen's (1976) but their results demonstrate a tighter upper bound and equivalent lower bound. Stationary Poisson demand is used in the evaluation of the bounds of expected outdating. In our research extends this literature to consider a multi-echelon inventory system that is subject to a Markovian demand process. Below, we will discuss the applicable multi-echelon inventory and our contribution to the stream of existing literature.

<u>MULTI-ECHELON PERISHABLE INVENTORY</u>

Multi-Echelon Perishable Inventory Allocation (age/location based issuing)

We define age/location based issuing as the method in which the perishable product's lifetime is accounted. In these age/location models, the remaining lifetime and location of the product defines the echelon and the inventory is tracked at each location (i.e., retailer). Prastacos (1981) extends his previous single period models to a multi-period single-suppler *n*- retailer allocation model for perishable goods. The model uses a one period optimal policy to establish bounds on the long run expected cost of a multi-period problem. Lystad (2006) extends this model to include backlogging. Lastly, Ketzenberg and Ferguson (2008) assume a single-supplier, single-retailer environment where the retailer orders 0 or *Q* units.

Multi-Echelon Perishable Inventory Single Location (age based issuing):

In the following papers, age is used as the echelon definition. Haijema et al. (2007) considers two demand types: a patient demand for fresher (aka younger) blood platelet or a patient demand which can receive either older or young platelets. The echelons are defined as the age of the blood product. Haijema et al. expanded the state space of his Markov chain to include the day of the week in addition to the inventory age vector. However, the demand from week to week was assumed to be iid. The demand for 'young' platelets is more certain than that of 'any' due to the nature of the use of each type (oncology vs. trauma). A two dimensional order up to policy is obtained which characterizes two vectors: one for total inventory level and one for 'young' only. Goh et al. (1993) evaluate product substitution policies when there are two levels of inventory: 'new' and 'old'. In our research our echelons are not defined by age and thus movement from one echelon to the next does not happen automatically with the passage of time. Instead we define a serial perishable inventory system where the Stage 1 raw material is produced into Stage 2 finished goods inventory. We make a production decision which converts the material and further affects the lifetime of the product. We define this issuing process as production based and discuss the relevant literature below.

Multi-echelon Perishable Inventory (production based issuing):

This type of issuing involves progression to downstream echelons as a result of further production-based activities. Fujiwara et al. (1997) models such a case motivated by the perishable foods industry, where there are two-stages of inventory: whole product (e.g. meat carcass) and sub-products (e.g. sirloin steak) produced from the whole product. Fujiwara et al. (1997) assumes that the whole product's shelf life is much longer than its sub-products due to the cold storage process. Therefore they consider a single cycle where the cycle length is the shelf life of the whole product and is made up of multiple periods which represent the shelf life of the sub-product. They derive a mathematical model for determining the base stock level of the whole product and a stationary base stock level for each sub-product when demand occurs at the second stage. The demand is assumed to be stationary.

We extend two-stage perishable inventory research to consider production activities that affect the remaining shelf life of the product. We formulate the problem as a Markov decision problem (MDP) an further define and model two ordering and production scenarios relevant for pharmaceutical drug management. We define a stochastic demand function that is a property of the system instead of an exogenously defined random variable. This demand function captures the complexity of the hospital setting and the interrelationship between the patient's condition, the patient's demographics, and drug utilization. Further in contrast to Fujiwara et al. (1997), we assume a cost minimization model with no second stage salvage value (i.e. expired products are may not be used), demand that is nonstationary, and shortages that can be fulfilled in two different ways. We show that the optimal production and inventory policies exist and are dependent on the shortage fulfillment method.

1.4 Dissertation Organization

The remainder of the dissertation is organized as follows: Chapter 2 is the first dissertation paper: *The Impact of the Product Lifecycle on Shipment Strategy with Two Modes of Supply*, Chapter 3 is the second paper: *A Two-Stage Inventory and Production Model for Perishable Products with Markovian Demand and Various Demand Fulfillment Scenarios*, Chapter 4 is third paper: *Stochastic Multi-Echelon Perishable Inventory Models for Meropenem Drug Management*, Chapter 5 discusses the research contributions of this dissertation and suggests direction for future work.

CHAPTER 2:

The Impact of the Product Lifecycle on Shipment Strategy with Two Modes of Supply

Abstract: We consider a firm with a geographically distant supplier that simultaneously decides the order quantity for each of two supply modes (e.g., air shipment and ocean shipment) on a periodic basis. In each period, the firm may order from both sources, order from one source, or order nothing. This research extends the literature on two modes of supply by deriving the optimal solutions under nonstationary demand with lead time differences between the two modes of one or more periods and a finite sales horizon. This fluctuating demand environment is a distinctive feature of our model since both the demand mean and demand variability may change over time representing the product lifecycle (PLC). We develop a Markov Decision Process (MDP) which determines the optimal ordering quantities per mode under stochastic, nonstationary demand. We consider two cases: Case 1 is the general model which models demand evolution as an embedded Markov chain, and Case 2 considers the special case where demand evolution is time-dependent. Optimal base stock policies, defined by two base stock levels, are derived and are shown to be both state and time-dependent for Case 2. When the fast delivery has zero lead time, i.e., arrives in the

same period in which it is ordered, we prove that the policy structure holds regardless of the leadtime difference between the two modes. This result suggests that optimal shipping strategies (i.e., the distribution of total order quantities between each mode of supply) change throughout the PLC as demand characteristics change. We derive insights from the optimal structure of the time-dependent demand case for the general model. Further, we use numerical examples to explore the relationship between the optimal solutions and the phases of the PLC.

2.1 Introduction

In this paper, we consider a firm whose supplier is geographically distant and must determine how to effectively use two different modes of supply. The firm receives product from the supplier, completes final customization in their facility, and ships the product to the customer within a short (less than one week) lead time (Kapuscinski et al. 2004, Karawarwala and Matsuo 1996). We assume that the time to complete final customization and production is negligible in comparison to the supplier lead time (Kapuscinski et al. 2004). Thus, the focus is on the shipment method decision between the supplier and the firm.

The objective of this research is to develop inventory, ordering, and shipping policies that incorporate the impact of the product lifecycle (PLC) on demand uncertainty and variability. We define *demand evolution* as the process which describes the timing, uncertainty, and variability in demand over time. *Demand evolution* is modeled in two ways. In Case 1 a Markov modulated demand structure is used to represent the fluctuating demand environment (Song and Zipkin 1993). Under this structure, the timing of demand evolution is uncertain. Figure 2-1 is a graphical depiction of the three life cycle phases and their corresponding demand characteristics for Case 1. The circles in Figure 2-1 represent the demand states of the Markov chain in each phase of the PLC. The arrows signify the probabilistic transition between states symbolizing the uncertainty in the timing of demand evolution. As depicted by the demand states (i.e., density functions), we consider both changes in the variability and mean of demand in addition to uncertainty in the timing of the evolution of the PLC. We model a firm that observes more variable demand during the ramp and decline phases than in the maturation phase. This higher variability is due to uncertainty in the adoption of the new product in the ramp phase and uncertainty in demand during the decline phase.

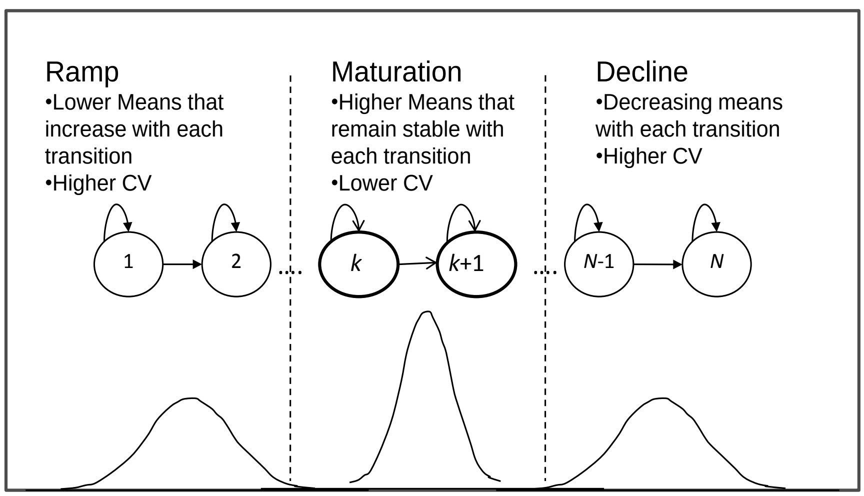


Figure 2-1: State transition diagram for demand including life cycle phases, Markov demand process (states), and corresponding demand characteristics

The second model, Case 2, assumes certainty in demand evolution and changes in the demand environment are time dependent. Although the timing of PLC transitions is certain; the demand itself is still a stochastic element of the model. The case of known PLC transitions represents a product whose demand can be shaped (e.g., by pricing) by the firm to follow a particular product lifecycle demand curve. Regardless of the type of demand evolution, our system's inventory level is governed by a Markov process which transitions probabilistically according to the current inventory level, the shipment decisions, and the realized demand. Under Case 1, we have an embedded Markov chain. In contrast to Sethi (2001, 2003), these Markovian structures allow us to focus on the specification of the demand environment and the impact of the type of demand evolution on the solutions.

Partnering with geographically distant suppliers, common in the high-tech industry, increases the uncertainty in ordering and production decisions due to long lead times. In this environment, determining the method in which a domestic firm's product will travel from the foreign supplier to the firm's facility is particularly important. Additionally, in contrast to functional products (the focus of most of the current literature), these innovative products have fluctuating demand characteristics. While functional products have longer sales seasons and more stable demand, the products we consider have high demand uncertainty and variability, a short selling season, high stockout cost, and high obsolescence cost (Lee 2002). Due to these PLC characteristics, firms use fast shipment methods, such as air freight, in addition to regular (i.e., slower) shipment methods. Firms attempt to cost-effectively utilize both shipment types as an integral part of their logistics strategy during all phases of the

PLC. In a case study conducted by Beyer and Ward (2001), Hewlett-Packard shipped 85% of a particular component of a network server using fast (air) methods. Further, while the case study indicates that firms commonly use advanced planning systems (APS) to determine order quantities and shipment plans, these systems typically do not determine inventory policies. Instead, the firm must specify these as deterministic inputs. The combination of static inputs and changing PLC demand characteristics expose the firm to risks such as stockouts and excess inventory.

Attention to the differences in demand characteristics during the product lifecycle phases (i.e., ramp, maturation, and decline) is critical to successful inventory management of innovative products. The ramp and decline phases of the product life cycle exhibit different behavior than the maturity phase; yet they are often ignored in analyses, resulting in stationary inventory and ordering policies (Kapuscinski et al. 2004). During the initial product ramp up, shortages can hamper the adoption of a new technology in the market; by ignoring the costs at the end of product life, excess inventory during the decline phase can have detrimental financial ramifications. Beyer and Ward (2001) note that ignoring the startup phase can "compromise market share whereas excess inventory at the end of the product's life can consume all the product line's profits." This environment is common in the electronics industries, such as the PC industry, where product lifecycles are short, products become obsolete (i.e., have little value at the end of the sales horizon), and exhibit high demand variability.

The remainder of the paper is organized as follows. In Section 2.2, we review the relevant literature. In Section 2.3, we formulate the Markov decision process (MDP) model for both cases of demand evolution. In Section 2.4 we prove the optimal policy for the case where the timing of demand evolution is known and the lead time difference between the two modes is greater or equal to one period. In Section 2.5 we relax the assumption that the timing of demand evolution is known. Section 2.6 presents a discussion of the results and the direction of future work.

2.2 Relevant Literature

The literature that considers the two modes of supply problem can be divided into two categories: those that evaluate ordering decisions and those that optimize them. A recent review of two modes of supply literature including both evaluation and optimization can be found in Minner (2003).

Fukuda (1964) extended inventory models to allow for delivery of order quantities by both fast (termed emergency) and slow methods simultaneously. However, the results are limited to the case where the difference in lead time between fast and slow methods is exactly one period and demand is stationary. Whittemore and Saunders (1977) generalized this model to consider two supply modes that differ by *n* review periods for a product with stochastic, stationary demand that is assumed to be independent and identically distributed over successive periods. Based on this assumption, the optimal ordering pair is independent of the time period in which the orders are placed. They derived steady state distribution of surface and air orders, which resembles the base stock policy of the single ordering option. Similarly, explicit results are found for the case where the difference in lead times is one period. Extending Fukuda's work to consider multiple delivery modes (greater than two), Feng et al. (2006) prove that a base stock solution structure is optimal for the first two consecutive modes but not otherwise. Like Fukuda (1964) they assume stationary demand, and therefore their solutions, unlike ours, are stationary. Veeraraghavan and Scheller-Wolf (2008) derived a heuristic dual index policy for two modes of supply, which is simpler to compute (and near optimal) assuming stationary, stochastic demand and infinite horizons. Of the papers that consider two modes of supply, there is a line of research that has explored the impact of various types of supply lead times on the ordering policies. Chiang and Gutierrez (1998) and Chiang (2003) considered an infinite horizon system with long review periods that may be longer than the supply lead times. Our research differs from these papers by considering a nonstationary demand function with unknown demand transitions (in Case 1) that mimic the uncertainty of the product lifecycle. We assume that the PLC has a finite horizon. In addition to specifying a different environment, we derive the optimal solution for two modes of supply which differ by $\lambda \geq 1$ period, when the fast mode has zero lead time, for a product with stochastic, state-dependent demand.

Similar to our research, Sethi et al. (2001, 2003) devise a model that incorporates advance information via forecast dependent expected demand distributions. In Sethi et al. (2001) it is shown that a modified base stock policy is optimal when at the beginning of each period a forecast update is made. Sethi et al. (2003) extends this model to include a fixed cost of ordering. In both research papers the lead time difference between the two modes of supply is one period. While our environment is similar to Sethi et al. (2001, 2003) in that both have state-dependent and nonstationary demand, there are some key differences. We develop an MDP to evaluate two types of demand evolution: Markovian (Case 1) and time-dependent (Case 2). The Markovian demand evolution model is an embedded Markov chain within the system level MDP. Our research focuses on the structure of the optimal policy for both types of demand evolution, and we demonstrate that the results depend on the type of demand evolution. This modeling approach allows us to represent the dynamic demand evolution through the specification of demand distributions which correspond to the PLC phases shown in Figure 2-1. We prove that the structure of the policy is maintained when lead time differences are greater than one period and the fast mode order quantity is delivered immediately.

In this paper, we also extend literature that considers only one mode of supply in a fluctuating environment. Song and Zipkin (1993, 1996) developed a Markovian demand model to represent a dynamic demand process for a single mode of supply inventory policy optimization problem. Our work is also related to Papachristos and Katsaros' (2008) research which considered an inventory model for one mode of supply with a lead time of λ periods in a fluctuating demand environment. We extend their results to consider two modes of supply where one mode has a fixed lead time of λ periods.

2.3 Problem Description and MDP Formulation

In this section, we formulate the two modes of supply decision problem as a Markov decision process (MDP). The optimal policy defines the order quantities for each of the two modes of supply given the state-of-the-world as defined by the inventory level and the demand state (in Case 1). In our analysis, the state space is dependent on the demand evolution. As defined in Section 2.1, *demand evolution* is the process which describes the timing, uncertainty, and variability in demand over time. Demand evolution is assumed to be either certain or uncertain (specifically Markovian). In Case 1, the demand evolution is Markovian and the state space has two elements: i_t, the demand phase (state), and x_t, the inventory level. For example, the *demand phase* is defined as the state in the Markov demand process associated with demand distribution i in period t that has mean $\mu(i_t)$ and standard deviation $\sigma(i_t)$. According to the Markovian assumption, we assume that the firm knows the current demand state but does not know the demand state in the next period. We represent the uncertainty in both the time and level of change in demand by the state transition probability matrix. In Case 2, we assume that demand evolution is certain; however, the state space is governed by a Markov process that is defined by the inventory level.

Formulating this problem as a discrete-time MDP allows us to determine the two modes of supply optimal policy that results in the minimum total expected cost per period, using a stochastic sequential-decision model defined by a set of states and decisions

(Puterman 1994). In the subsections that follow, we define the states, actions, and objective function of the MDP model for both Cases 1 and 2.

2.3.1 MDP Notation

We begin by describing the necessary notation used throughout the paper. The following summarizes the parameters and variables for Cases 1 and 2.

t	= {1, 2, …, T} where T is the horizon of the inventory problem
A	= finite collection of possible demand states {1, 2, …, N}
i_t	= the demand phase (state) during period t
x_t	= inventory level at the start of period t (may be a vector of in transit material)
(x_t, i_t)	= state of the MDP (Case 2) where if demand evolution (Case 1) is certain the state is defined solely as (x_t)
d_t	= the demand in period t, $d_t \geq 0$, dependent on i_t
$\psi_k(.)$	= the conditional density function of d_t when $i_t = k$
$\Psi_k(.)$	= the distribution function corresponding to $\psi_i(.)$
f_t	= the order quantity via the fast method in period t
s_t	= the order quantity via the slower method in period t
l_f	= the lead time of the fast shipment method
l_s	= the lead time of the slower shipment method, $l_f < l_s$
P	= the probability transition matrix associated with the demand states, where its elements, $p_{kj} = P[i_{t+1} = j \mid i_t = k]$ represent the probability that the demand state is k at time t and j at time $t + 1$
Q	= the probability transition matrix associated with the inventory level
c_0	= the variable cost of the fast shipment method
c_1	= the variable cost of the slower shipment method, $c_0 > c_1$
h_t	= unit holding cost where at time t, when t = T we refer to h_T as the end of life holding cost
b_t	= unit backorder cost where at time t, when t = T we refer to b_T as the end of life backorder cost

The commonality between both demand evolution models is the Markovian nature of the inventory level process. The Markov chain associated with the inventory level is independent of the type of demand evolution. The current inventory level changes as a result of the ordering decisions and the realized demand. In Case 1, we embed an additional Markov chain (that of the demand states) which we define in the following section.

2.3.2 Case 1: Uncertainty in the Demand Evolution Process

Modeling demand as a Markov process allows us flexibility in capturing demand uncertainty for products that face obsolescence and for new products with uncertain conditions during the product life cycle (Song and Zipkin 1993, 1996). Case 1 uses an embedded Markov chain to model demand that is nonstationary and state-dependent. Recall that we refer to the *demand phase* i_t as the state in the Markov demand process associated with demand distribution i in period t, which has mean $\mu(i_t)$ and standard deviation $\sigma(i_t)$. We assume demand is stochastic, non-stationary, and is defined according to the evolution of the product life cycle. Using the Markovian assumption, the probability of being in a demand phase in period t+1 depends only on the demand phase occupied in period t. Figure 2-2 represents a simple example of the demand process with six states. States 1 and 2 belong to the ramp phase (i.e., lower means, higher variability), states 3-4 represent the maturity phase (i.e., higher means, lower variability), and 5-6 the decline phase (i.e., lower means, higher variability). The demand process described by Figure 2-2 has a special structure in order to represent the product life cycle evolution. Each demand phase represents a time ordering; therefore, in general, the transition probabilities from demand state k to demand state j, p_{kj}, are non-zero for all $j \geq k$ and zero otherwise. The last state (e.g., state 6 in Figure 2-2) is an absorbing state signifying the end of the sales life/horizon.

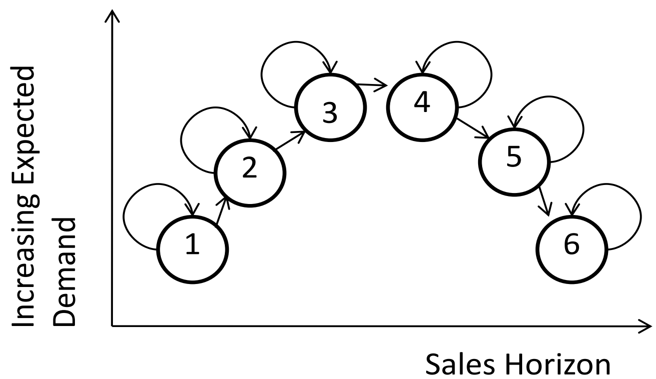


Figure 2-2: A six-state example of the Markovian demand process which represents a typical product lifecycle (PLC)

2.3.3 The State Space

As defined in Case 1, when the transition between lifecycle phases, and thus demand distributions, is uncertain, the state space is multi-dimensional. The first component of the state space, i_t, represents the current demand phase; and the second component denotes inventory level, x_t. In contrast, if the transition times are known (i.e., Case 2), the state space is composed solely of the inventory level since demand phase is known with certainty and dependent on the time period.

2.3.4 The Decision Space

Each period the decisions to be made are: (1) the order quantity f_t to be shipped via the fast method and (2) the order quantity s_t to be shipped via the slow method. The decisions are clearly dependent on the state of the system, i.e., the inventory level and demand. For l_s =1 and l_f =0, the decisions and events are shown in the order in which they occur in Figure 2-3. At the beginning of period t, the state is observed and orders may be placed via both supply methods. The order quantity via the fast supply method is assumed to arrive immediately, then demand occurs and is fulfilled (to the extent possible) prior to the end of the period, and unsatisfied demands are fully backlogged. The order quantities (s_t, f_t), are determined simultaneously at the beginning of period t.

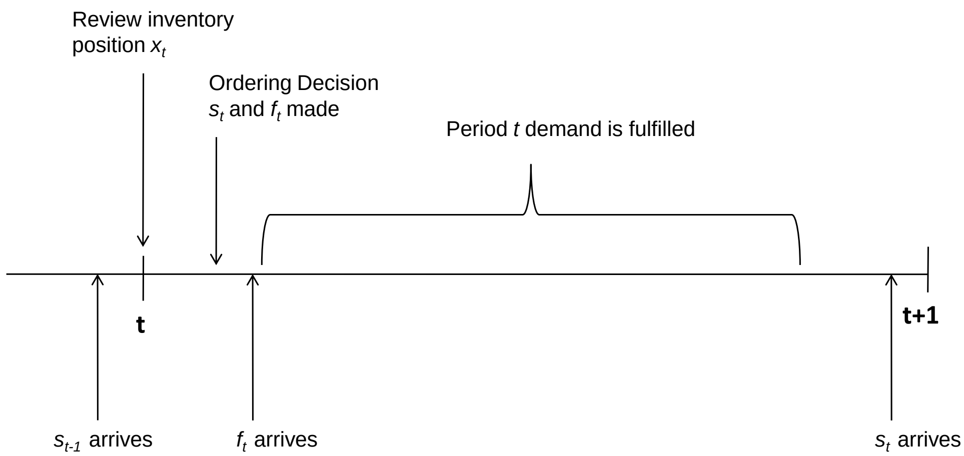


Figure 2-3: Timeline of events and decisions during period t

2.3.5 Case 1: State Transition and Transition Probabilities

In the discussion that follows we define the state transition and transition probabilities for Case 1. If demand evolution is Markovian, then the state transition depends on the actions (s_t, f_t), inventory level (x_t), current demand phase (i_t), and the actual demand. We define the state in period t as $S_t = (x_t, i_t)$, and the decisions made in this period are defined as s_t and f_t. If at time t the demand phase is j, the demand in period t (d_t) has distribution $\psi_j(\cdot)$ and is conditioned on the current demand phase, j. If the demand phase in period t+1 transitions from phase j to phase k with probability p_{jk} and the inventory level becomes x_{t+1}, the next state will be $S_{t+1} = (x_{t+1}, i_{t+1}=k)$. Therefore, for Case 1 the state transition is divided into two parts. Part (a) defines the inventory level transition ($l_s = 1$, $l_f = 0$), and (b) defines the demand phase transition as:

$$(a)\ x_{t+1} = x_t + f_t + s_t - d_t$$

$$(b)\ i_{t+1} = k$$

$$P\{i_{t+1} = k \mid i_t = j\} = \begin{cases} p_{jk} & \geq 0 \text{ for all } k \geq j \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{j=1}^{N} p_{kj} = 1$$

Now that the state space and transitions have been described for Case 1, in the following section we define MDP cost function.

2.3.6 Case 1: Cost Function and Optimization

We define the cost function for the case where the demand phase is part of the state definition. Assume that for the discussion that follows, $l_f = 0$ (i.e., fast orders are received immediately) and $l_s = 1$ (i.e., slow orders are received for use in the next period). Given the state of the system in period t is (x_t, i_t,), the single period expected cost under decision (s_t^i, f_t^i), $G_t(x_t, i_t)$ is defined as

$$G_t(x_t, i_t) = \left[c_1 s_t^i + c_0 f_t^i + L_t(x_t + f_t^i, i_t) \right] \tag{2.1}$$

where $L_t(x_t, i_t)$ is defined as

$$L_t(x_t, i_t) = \begin{cases} \int_0^{x_t} h_t(x_t - q)\psi_{i_t}(q)dq + \int_{x_t}^{\infty} b_t(q - x_t)\psi_{i_t}(q)dq, & x_t > 0 \\ \int_0^{\infty} b_t(q - x_t)\psi_{i_t}(q)dq, & x_t \le 0 \end{cases}$$

The first two terms are the linear costs for the slow and fast shipping methods, respectively, and the third term is the total expected holding and backlog costs. For innovative technology products, holding costs often reflect price erosion that occurs in place of, or in addition to, traditional physical storage costs (Kurawarwala and Matsuo, 1996). Accordingly, the stochastic dynamic program that accompanies the MDP incorporates equation (2.1) as follows, where we define $J_t(x_t, i_t)$ as the total expected cost over *t* periods given that you are in state (x_t, i_t). We assume that the sales life is *T* periods and that the final costs are assessed in period *T*+1.

$$\begin{aligned} &J_t(x_t, i_t) = \min_{s_t^i, f_t^i \ge 0} \left[G_t(x_t, i_t) + \alpha \sum_{i=1}^{N} p_{i,j} \int_0^{\infty} J_{t-1}(x_t + f_t^i + s_t^i - q, j)\psi_j(q)dq \right] \\ &t = 0,1,\ldots,T \\ &J_0(x_1) = h\max(0, x_1) + b\min(0, x_1) \end{aligned} \tag{2.1}$$

The second term in (2.1) is the expected cost with t-1 periods to go as a result of the decision in the current period. At the end of the horizon ($t = 0$), final backorder and holding costs are assessed.

The MDP is solved using backwards recursion (Puterman 1994), and optimal decisions ($s_t^*(x_t, i_t)$, $f_t^*(x_t, i_t)$) are found for each possible state in each period. The last period to make a decision to ship via the fast method is defined as period T-l_f, and, similarly, period T-l_s is the last decision period for slow shipment decisions. At the end of the finite horizon, end of life (EOL) costs, which are comprised of excess and obsolescence costs, are assessed. Obsolescence cost penalties are commonly associated with innovative products because new products replace older technology rendering them less valuable in the market. At the end of the horizon, unfulfilled orders may be fulfilled with the new product or may become a lost sale, resulting in an undesirable situation for the firm and the customer. In the following section, we derive the optimal policies and provide a numerical example for a special case of Case 1.

2.3.7 Case 2: A Special Case When Demand Evolution Is Certain

As defined in Section 2.3, Case 2 is a special case of Case 1 where the demand phase transitions are time dependent and thus certain. For Case 2 the demand phase transition matrix can be represented by an identity matrix. The evolution model used in Case 2 is assumed to have nonstationary demand, time-dependent demand evolution, and a finite horizon. For example, the demand distribution associated with period t has mean $\mu(t)$ and standard deviation $\sigma(t)$. In period t+1 the demand distribution has mean $\mu(t+1)$ and standard deviation $\sigma(t+1)$ with certainty.

When demand evolution is certain, the state transitions are dependent on the actions, inventory level, and the actual demand. Assuming that the shipments arrive as shown in Figure 2-3, the state in period t is defined as $S_t = (x_t)$ and the decisions made are s_t and f_t. The state in period t+1 is defined as:

$$x_{t+1} = x_t + f_t + s_t - d_t$$

The demand in period t takes the value of d_t which has distribution $\psi_t(\cdot)$ and is conditioned on the current time period.

We use this special case to derive the optimal policy and gain insights into the solution structure for both cases. In the discussion that follows we assume that the fast shipment arrives immediately while the slow shipment arrives at the end of the period. We relax the restriction of our model to accommodate slow shipment lead times greater than one in Section 2.4.2. In Section 2.4 we derive the structure of the optimal policy and in Section 2.4.1 provide a representative numerical example.

2.4 Optimal Policy Structure

Let us consider a new product for which there are three product lifecycle phases: ramp, maturity, and decline. We assume that we know exactly when the transitions between the three phases occur. Thus our demand evolution process is deterministic, but the distribution of demand is time-dependent and stochastic. Equations (2.1) and (2.2) are modified by reducing the state definition to include only the inventory level, since demand is not state-dependent. We specify the demand distributions to characterize the PLC as defined in Section 2.1.

We can rewrite the end of period cost equation defined by (2.2) for *n* periods to go as follows in (2.3) and solve for the optimal shipment quantities via each method.

$$J_n(x_n) = \min_{s_n \geq 0,\, f_n \geq 0} \left\{ c_0 f_n + c_1 s_n + L(x_n, f_n) + \alpha \int_0^{\infty} J_{n-1}(x_n + f_n + s_n - q)\psi_{n-1}(q)dq \right\} \tag{2.2}$$

If we make the substitutions (1) $w_n = x_n + f_n$ called the *immediate cumulative base stock level*, (2) $v_n = x_n + f_n + s_n$ called the *total cumulative base stock level*, and (3) $c = c_0 - c_1$, then we can rewrite equation (2.3) as follows and solve for the two base stock levels:

$$J_n(x_n) = \min_{v_n \geq w_n \geq x_n} \left\{ c(w_n - x_n) + L_n(w_n) + c_1(v_n - x_n) + \alpha \int_0^{\infty} J_{n-1}(v_n - q)\psi_{n-1}(q)dq \right\} \tag{2.3}$$

Extending the work of Fukuda (1964), we prove the optimal inventory policy is a two-level base stock policy. Lemma 1 summarizes the structure of the optimal policy and solution to the problem defined in Section 2.4.1.

Lemma 1: *A two-level base stock policy is optimal with optimal levels in period t* w_t^* and v_t^*, *called the immediate cumulative base stock level and total cumulative base stock level, and order quantities* $f_t(x_t)$, $s_t(x_t)$ *which are state and time dependent.*

The result leads us to the following theorem.

Theorem 1: *For* $t \geq 2$, *the optimal ordering policy is a state- and time-dependent base stock policy with two base stock levels* $w^*_t = x_t + f_t$ and $w^*_t = x_t + f_t + s_t$. *The optimal order quantities with t periods remaining and the inventory position is x are denoted as* $(s_t^*(x_t), f_t^*(x_t))$.

Case A: $v_t^* > w_t^*$

$$f^*_t(x_t) = \begin{cases} w_t^* - x_t & \text{if } x_t < w_t^* \\ 0 & \text{otherwise} \end{cases}$$

$$s^*_t(x_t) = \begin{cases} v_t^* - w_t^* & \text{if } x_t < w_t^*, \text{ for } t = 1,..,T - l_s \\ v_t^* - x_t & \text{if } w_t^* \leq x_t \leq v_t^* \\ 0 & \text{otherwise} \end{cases}$$

Case B: $v_t^* \leq w_t^*$

$$f^*_t(x_t) = \begin{cases} v_t^* - x_t & \text{if } x_t < v_t^* \\ 0 & \text{otherwise} \end{cases}$$

$$s^*_t(x_t) = 0 \qquad \text{for all } x$$

The proof of Theorem 1 derives the solution technique for the optimal base stock levels defined by Theorem 1. First, the solution w_t^* for the *immediate base stock level* is determined. Second, the solution for the slow shipment mode is determined by evaluating the cost function at w_t^*. If the derivative of the cost function is negative, at w_t^*, then we can find a $v_t^* > w_t^*$ (i.e., Case A) that is the optimal *total cumulative base stock level*. If the derivative is non-negative, its optimal solution is to not order via the slower mode. For details of the proof, refer to the Appendix.

Theorem 1 proves that the optimal order quantities are state dependent because the solutions are contingent upon the current inventory level. Further, the optimal solutions are dependent on the time period because the demand is time-dependent. These are important observations since the demand is dynamic throughout the PLC and therefore the quantity shipped by each method should be dynamic. We prove that the nonstationary, state-dependent base stock levels are optimal, which extends the stationary proof by Fukuda (1964).

2.4.1 Numerical Example

We provide a numerical example to demonstrate the state dependency of the optimal solution. The variables used for this example are summarized in Tables 2-1 and 2-2. The parameters of the demand distributions are defined in Table 2-1 and their probability density functions are shown in Figure 2-4. The transition periods are indicated graphically by the vertical lines in Figures 2-5 and 2-6. The first and third demand distributions, representing the ramp and decline phases, are characterized by high variability, CV = 0.84; and the second, representing the maturity phase, has lower variability, CV = 0.5. In order to isolate the effect of changes in CV and shape, both phases are assumed to have the same mean. We use a negative binomial distribution because it is discrete, bounded by zero, and can be made to take on a variety of shapes by selecting its parameters (*r*, *p*).

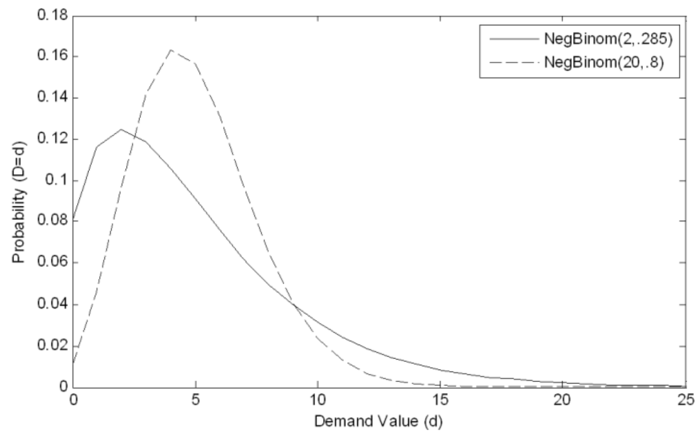


Figure 2-4: Demand distributions with μ=5 and CV = 0.5 and 0.84

Throughout our numerical experimentation we explored a variety of horizon lengths (from five to twenty periods) and costs. For this example, we have a fourteen-period decision problem where period fifteen is the end of the decision horizon where final costs are assessed.

Table 2-1. Parameter values for the numerical examples

Parameter	Value
l_s	1
l_f	0
c_s	1
c_f	10
h	0.01
b	20

Table 2-2. Summary of demand parameters for the numerical experiments

	Ramp	Maturity	Decline
r	2	20	2
p	0.285	0.8	0.285
Mean	5	5	5
Cv	0.84	0.5	0.84

The results were collected for a large range of inventory levels that did not constrain the selection of optimal solutions. Select results (i.e., optimal decisions) are shown in Figures 2-5 and 6 for a backlog and inventory (respectively) level of 3.

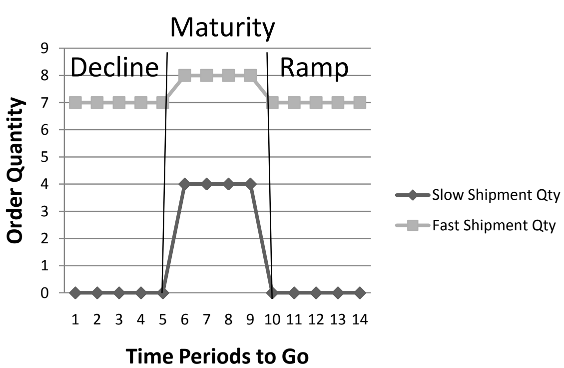


Figure 2-5: Order quantities for a lifecycle with 3 phases and known transition times for each phase when the state is at a backorder level of 3, (i.e., x = -3).

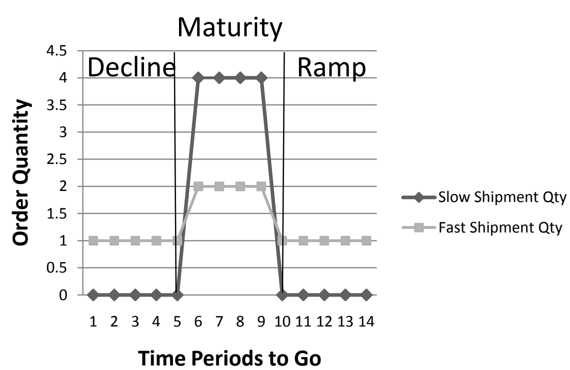


Figure 2-6: Order quantities for a lifecycle with 3 phases and known transition times for each phase when the state is at an inventory level of 3, (i.e., x = +3).

The example shows that as the CV decreases (i.e., as we move from ramp to maturity) we choose to ship a larger fraction of the total shipment quantity ($s + f$) via the slow method. Figure 2-5 shows that optimal cumulative base stock level in the maturity phase is eight units, and in the ramp and decline it is four units. If we ignore the ramp and decline and use a stationary policy defined by the base stock levels of the maturity phase, we may have excess inventory in the pipeline. In contrast, if we use a stationary policy as defined by the optimal base stock levels of the ramp or decline phases the firm loses the cost advantage of shipping more volume using the cheaper, slower mode. When the inventory level is three units (Figure 2-6), we observe that; though the optimal solutions are distinct from those shown in Figure 2-5. Carrying a higher level of inventory (i.e., three in this case) results in higher utilization of the slow shipping method during the maturity phase. Recall that the inventory state is kept static in the example in order to isolate the impact of the changes in demand variability, due to the dynamics of the PLC, on the solutions.

Although this is a simple example, it allows us to clearly illustrate and glean insights on the impact of the PLC on shipment policies. First, the optimal solutions are dependent on the variability and shape of demand in the corresponding PLC phase. During phases of high variability the solution emphasizes the use of the accelerated method of shipment, while during the maturation phase both accelerated and regular shipment methods are utilized. Therefore, a firm with a product in the maturity phase should be able to take advantage of shipping a larger percentage of its overall order quantity via the slower, cheaper mode compared to the quantity ordered during the ramp and decline phases. The stability of the demand during the maturity phase lessens the risk of higher inventory in the pipeline. Second, we observe that the behavior of the solution (i.e., a change in optimal base stock level) changes as the phase shifts due to the one-period lead time difference between the two modes and certainty in demand evolution. When the lead time of the slow shipment increases or there is uncertainty in the demand evolution, the optimal base stock level could change many periods before the actual shift change occurs. Lastly, carrying higher levels of inventory may result in the ability to utilize the slow shipment method when demand variability is more stable. This is an intuitive result since carrying more inventory on hand decreases the probability that you will need additional products immediately to fulfill demand. In the following section we extend the model from Section 2.4 to allow for a larger lead time difference between the slower and faster modes of delivery.

2.4.2 Slow Shipment Modes with Lead Time Greater Than One

In this section the results from Theorem 1 are extended to consider the case where the lead time for the slower shipment method is a constant λ periods. The shipping lead times are assumed to be a multiple of the review period, and the faster shipment method is assumed to arrive in the same period that it is ordered (i.e., zero lead time). Papachristos and Katsaros (2008) prove that there is a base stock policy for a similar inventory model in a fluctuating environment with only one shipment option and a constant λ period lead time. We extend the results in Section 2.4, with one period lead time to consider longer lead times, and the results from Papachristos and Katsaros (2008) to include two modes of supply.

To extend the PLC-demand model from Section 2.4.1, the state space is now a vector of inventory levels. For example, in period *t*, the state is the vector $(x_{t,} s_1, \ldots, s_{\lambda-1})$ where x_t is the inventory level on hand and $s_1, \ldots, s_{\lambda-1}$ are the orders that will arrive at the beginning of each of the following λ-1 periods. In contrast to Papachristos and Katsaros (2008), our decision in the n^{th} period includes two decision variables for each state, s_n and f_n, at unit prices c_1 and c_0, respectively. The decision s_n becomes the last element in the inventory vector and is renamed s_λ due to its remaining lead time. The cost formulation derived in (2.3) for a problem with *n* periods remaining is:

$$J_n(x_n, s_1, \ldots, s_{\lambda-1}) = \min_{s_n \geq 0,\, f_n \geq 0} \left\{ \begin{array}{l} c_0 f_n + c_1 s_n + L(x_n + f_n + s_1) + \\ \alpha \int_0^\infty J_{n-1}(x_n + f_n + s_1 - q, s_2, \ldots, s_\lambda) \psi_{n-1}(q) dq \end{array} \right\} \tag{2.5}$$

As assumed in Section 2.4.1, the first two terms correspond to the linear procurement costs, the third term is the holding/backlog function, and the fourth term is the discounted expected cost function when there are *n*-1 periods remaining. To simplify the notation (as in Papachristos and Katsaros 2008) we define $u_m = s_1 + s_2 + \cdots + s_m$ for all $m = 1 \ldots \lambda\text{-}1$. Demand is assumed to be a function of time (i.e., nonstationary) and the demand distribution in each period is known with certainty. For any period *t* the following substitutions are made: (1) $w_t = x_t + f_t$ called the *immediate cumulative base stock level*, (2) $v_t = x_t + f_t + u_{\lambda-1} + s_\lambda$ called the *total cumulative base stock level*, and (3) $c = c_0 - c_1$; then equation (2.1) can be rewritten as follows and solved for the two base stock levels, w_t, v_t.

$$J_n(x_n, s_1, ..., s_{\lambda-1}) = \min_{\substack{v_n \geq w_n \geq 0 \\ w_n \geq x_n}} \left\{ \begin{array}{l} c_0(w_n - x_n) + L(w_n + s_1) + c_1(v_n - w_n - u_{\lambda-1}) + \\ \alpha \int_0^\infty J_{n-1}(v_n - \sum_{m=2}^{\lambda} s_m - q, s_2, ..., s_\lambda)\psi_{n-1}(q)dq \end{array} \right\} \tag{2.6}$$

In the following Lemma, the structure of the optimal policy and solutions for each mode of supply under the assumptions introduced in Section 2.4 are presented.

Lemma 2: *Order up to levels with two mode-of-delivery based levels is optimal when the fast delivery mode has zero lead time. The optimal solutions in period t, defined as the immediate cumulative base stock level and total cumulative base stock level, are* w_t^* and v_t^* *with order quantities* $f_t(x_t, s_1, ..., s_{\lambda-1}), s_t(x_t, s_1, ..., s_{\lambda-1})$ *that are state and time dependent.*

The existence of the optimal order up to levels follows from Sethi et al. (2001) and Feng et al. 2006. In general, the form of the policy is not base stock in structure but optimal order quantities can be found that minimize the total expected cost function. As the lead time for the slower shipment increases, the firm must consider how demand will evolve during the shipment time. The ordering decisions made in the current period will be made with the intent of fulfilling future demand. When demand evolution is uncertain, as we will discuss in the following section, the exposure to cost incurred because of inaccurate order quantities increases with the extent of uncertainty. Lastly, we are interested in finding conditions which lead to the optimality of base stock policies in our future research.

2.5 Implications for the Case of Unknown Transitions

The results from Theorem 1 (for the special case) provide insight for the general problem where transition times are uncertain. For the general case, the evolution of the demand process is uncertain. We define this uncertainty in the timing of demand phase transitions by a transition matrix, *P*, which can represent various demand evolution scenarios.

2.5.1 Extension of Results from Section 2.4

In this case equations (2.3) and (2.4) can be rewritten as (2.12) and (2.13), respectively, which expand the state space to include the demand phase while the cost equation is modified to include the transition probabilities. Note that the demand function at time $t + 1$ is conditional on the demand phase in period t. For example, if the current demand phase is k, the system transitions to demand phase j with probability p_{kj}. We define the substitutions (1) $w_t^i = x_t + f_t$, the *immediate cumulative base stock level* when demand is in state i at time t, (2) $v_t^i = x_t + f_t + s_t$, the *total cumulative base stock level* when demand is in state i at time t, and (3) $c = c_0 - c_1$, and rewrite equation (2.4) as follows to solve for the two base stock levels.

$$J_n(x_n, i_n) = \min_{s_n^i \geq 0,\, f_n^i \geq 0} \left\{ c_0 f_n + c_1 s_n + L(x_n + f_n, i_n) + \alpha \sum_{j=1}^{N} p_{i_n j} \int_0^\infty J_{n-1}(x_n + f_n + s_n - q, j) \psi_j(q) dq \right\} \quad (2.12)$$

$$J_n(x_n, i_n) = \min_{\substack{v_n^i \geq w_n^i \geq 0 \\ w_n^i \geq x_n}} \left\{ c(w_n^i - x_n) + L(w_n^i, i_n) + c_1(v_n^i - x_n) + \alpha \sum_{j=1}^{N} p_{i_n j} \int_0^\infty J_{n-1}(v_n^i - q, j) \psi_j(q) dq \right\} \quad (2.13)$$

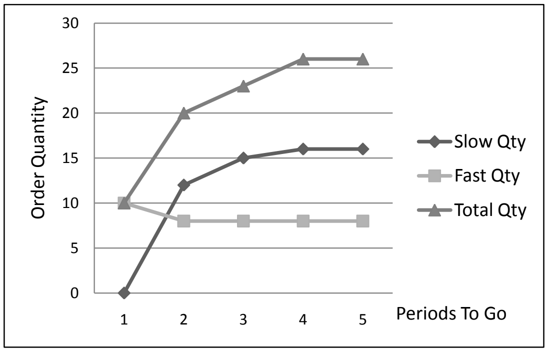
As this formulation suggests, the optimal base stock levels are dependent on both the inventory position and the current demand phase. In the following section we present a numerical example using the new formulation to solve for the optimal base stock levels. The changes in the cost formulations (i.e., equations 2.12 and 2.13) may significantly impact the solutions.

2.5.2 Numerical Example

This section presents a numerical example to illustrate the behavior of the optimal logistics policy. The following transition matrices were defined to represent different types of demand evolution (Treharne and Sox, 2002). The transition matrices P_1 and P_2 characterize a demand process with a slower demand evolution (represented by smaller probabilities of moving to the next PLC phase), and a demand process with faster demand evolution, (represented by larger probabilities of moving on to the next PLC phase).

$$P_1 = \begin{bmatrix} .8 & .1 & .1 \\ 0 & .8 & .2 \\ 0 & 0 & 1 \end{bmatrix} \qquad P_2 = \begin{bmatrix} .2 & .4 & .4 \\ 0 & .2 & .8 \\ 0 & 0 & 1 \end{bmatrix}$$

Figures 2-7 and 2-8 show the total order quantity, slow order quantity, and fast order quantity for a backlog level of three and demand state 2 (i.e., maturity) as a function of the number of periods to go. The *P* matrices used in this case were P_1 and P_2 representing a slow and fast upward trend, respectively. The lead time of the slow shipment method is assumed to be one period and the demand parameters are defined in Table 2-2.



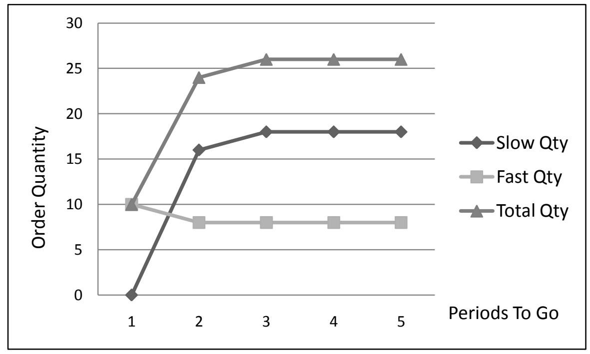


Figure 2-7: (Slow Evolution Case, P_1)
Optimal order quantities when backlog =3
and demand phase = 2

Figure2-8: (Fast Evolution Case, P_2)
Optimal order quantities when backlog =3
and demand phase = 2

The numerical results show that the optimal fast shipment decision does not differ between these cases. This is reasonable because the optimal order quantity shipped via the fast method is intended to impact the present period. Therefore changes in the demand distribution in subsequent periods are most cost effectively fulfilled via the slow shipment method. When the current demand state is 1 (ramp phase) and the demand evolves according to P_1, the optimal slow shipment order quantity is greater than or equal to when demand evolves according to P_2. This observation suggests that the slow shipment method may be more cost effective when demand evolves slowly. When the current demand state is 2 the order quantity shipped via the slow method, under transition matrix P_1, is less than the quantity shipped when demand evolves according to P_2. Since demand state 3 is an absorbing state, the demand evolution defined by P_2 defines a larger probability of demand evolving to its final end of produce life phase. Finally, when the demand state is 3, the order quantities are equal since state 3 is an absorbing state and therefore the demand evolution is irrelevant. These results are consistent with the demand evolution characteristics.

There are similarities between the optimal policies when demand evolution is certain and uncertain. In both cases, the fast order quantity orders up to a state- and time-dependent base stock level. However, the optimal slow shipment order quantity varies greatly. When the evolution is uncertain, the order quantities differ from the results of the certain evolution case because of the uncertainty in the demand for the next period. In addition to base stock levels that are dependent on the inventory level after the fast order and the number of periods to go (as in the case when demand transitions are known), the optimal slow order quantity is dependent on the current demand state.

2.6 Conclusions

This paper extends the literature on two modes of supply by deriving the optimal solutions under nonstationary demand, lead time differences between the two modes of more than one period, and a finite sales horizon. We specify demand evolution and demand characteristics according to the structure of a PLC. The problem is formulated as an MDP which allows us to model the dynamics of the inventory and demand model. Two cases are defined: Case 1 is the general model which models demand evolution as an embedded Markov chain, and second is a special case (i.e., Case 2) where demand evolution is time-dependent. These two cases represent different product environments with different levels of PLC demand uncertainty and progression.

Optimal base stock policies, defined by two base stock levels, are derived and are shown to be both state- and time-dependent. We prove that the state dependency of the optimal solutions holds regardless of the lead time difference between the two modes. We return to the general case in order to derive insights from the optimal structure of the special case. Numerical experiments show that the optimal solutions may differ significantly between the two cases. While past research has focused on stationary demand, our results show that it is important to consider the type of demand evolution as well as the demand variability.

Given the nonstationary structure of the optimal policy, ordering strategies that use steady state solutions may not perform as well as the dynamic policies derived in this paper. This is particularly true during the ramp and decline phases of the PLC, which are often ignored. The implications of these results are that different strategies (i.e., the split of total order quantities per mode of supply) should change throughout and according to the product lifecycle evolution. Potential areas for future research include understanding the cost performance implications of considering uncertainty versus certainty in the demand process.

Appendix A

Proof of Theorem 1

In this section we will extend the proof of Fukuda (1964), to determine the optimal ordering policies for the case where demand is nonstationary and time dependent. In order to shorten the notation define $L_t(x)$ as the expected single period holding and backlog cost given that the inventory position is x and the demand phase has probability density function ψ_t in period t.

$$L_t(x_t)=\begin{cases}\int_0^{x_t} h_t(x_t-q)\psi_t(q)dq+\int_{x_t}^{\infty} b_t(q-x_t)\psi_t(q)dq, & x_t>0\\ \int_0^{\infty} b_t(q-x_t)\psi_t(q)dq, & x_t\le 0\end{cases} \tag{A1}$$

For simplicity of notation, from this point forward, without loss of generality, we omit the time subscript *t* on the holding and backorder costs terms in (A1). Using the Leibniz Rule we show that $L_t'(x)$ is increasing in x.

$$L_t'(x)=\begin{cases} h\Psi_t(x)-b[1-\Psi_t(x)] & x\ge 0\\ -b & x<0\end{cases}$$

Using induction arguments we start with the single period model. In the single period model we can only ship via the immediate delivery method, f_1 (i.e. $s_1=0$) where the subscript represents the current period. As before in Section 2.3, we define the cost function of a model consisting of a single period as:

$$J_1(x_1)=\min_{f_1\ge 0}\ c_0 f_1 + L_1(x_1+f_1) \tag{A2a}$$

Defining $w_t = x_t + f_t$ as the base stock level after the receipt of the fast shipment but before demand occurs; we rewrite the single period cost function as follows:

$$J_1(x_1)=\min_{w_1\ge x_1}\ c_0(w_1-x_1)+L_1(w_1) \tag{A2b}$$

To obtain the solution to the immediate base stock level for the one period problem, $\hat{x}_1$ we obtain from (A2b) the following expression, which includes all the terms that are functions of w.

$$F_1(w_1)=c_0 w_1 + L_1(w_1) \tag{A3}$$

Using the results from the derivative of (1), $F_1'(w_1)$ is increasing in w, and as $w \to \infty$ it approaches $c_0 + h > 0$. For $w \leq 0$, $F_1'(w_1) = c_0 - b$, thus we assume that

(A4)

In the one period problem there exists a unique positive w_1^*, such that $F_1'(w_1^*) = 0$. Therefore if the inventory position, x_1, is less than w_1^* we can see that we order up to w_1^* and if $x_1 \geq w_1^*$ we order nothing. It follows that $J_1(x_1)$ is given by the following expression:

$$J_1(x_1) = \begin{cases} c_0(w_1^* - x_1) + L_1(w_1^*) & x_1 < w_1^* \\ L_1(x_1) & x_1 \geq w_1^* \end{cases} \tag{A5}$$

Note that these results make no assumptions about demand and therefore apply to all demand phases $J_1(x_1)$ is convex in x and $J'_1(x_1) \geq -c_0$.

Next, we evaluate the two-period problem. The decision in the second period now includes two decision variables for each state, namely s_2 and f_2 at unit prices c_1 and c_0, respectively. We assume that $c_0 > c_1$. We formulate the model as

$$J_2(x_2) = \min_{s_2 \geq 0,\, f_2 \geq 0} \left\{ c_0 f_2 + c_1 s_2 + L_2(x_2 + f_2) + \alpha \int_0^\infty J_1(x_2 + f_2 + s_2 - q)\psi_1(q)dq \right\} \tag{A6}$$

As before, the first two terms are the linear procurement costs, the third term is the holding/backlog function, and the fourth term is the discounted expected cost function when there is one period remaining. We assume demand is a function of time (i.e., nonstationary), but the demand distribution in each period is known with certainty. If we make the substitutions (1) $w_t = x_t + f_t$ called the *immediate cumulative base stock level*, (2)

$w_t = x_t + f_t + s_t$ called the *total cumulative base stock level*, and (3) $c = c_0 - c_1$, then we can rewrite equation (A6) as follows and solve for the two base stock levels:

$$J_2(x_2) = \min_{v_2 \geq 0, w_2 \geq x} \left\{ c(w_2 - x_2) + L_2(w_2) + c_1(v_2 - x_2) + \alpha \int_0^\infty J_1(v_2 - q)\psi_t(q)dq \right\} \quad \text{(A7)}$$

Now let us consider the following minimization equation which is the two-period cost equivalent to (A2b):

$$\tilde{L}_2(x_2) = \min_{w_2 \geq x} \; c(w_2 - x_2) + L_2(w_2) \quad \text{(A8)}$$

Since the right hand side of (A8) is identical to (A2b) except for that c_0 is replaced by c, we may immediately conclude that, under assumption (A4), there exists a unique positive solution to (A9), w_2^*, such that

$$\tilde{L}_2'(x_2) = c + L_2'(w_2^*) = c_0 - c_1 + L_2'(w_2^*) = 0 \quad \text{(A9)}$$

It follows that $w_2 = w_2^*$ for $x_2 < w_2^*$, and $w_2 = x_2$ for, $x_2 \geq w_2^*$; $\tilde{L}_2(x_2)$ consequently is given by

$$\tilde{L}_2(x_2) = \begin{cases} c(w_2^* - x_2) + L_2(w_2^*) & x < w_2^* \\ L_2(x_2) & x \geq w_2^* \end{cases} \quad \text{(A10)}$$

With these properties established we return to equation (A7) and evaluate two cases to finish evaluating the two-period problem.

Case (a) $x_2 \geq w_2^*$

As was shown in equation (A8) we will have $w_2 = x_2$ (i.e., we do not need to order via the fast shipment method) and (A7) is rewritten as

$$J_2(x_2) = L_2(x_2) + \min_{v_2 \geq 0}\left\{c_1(v_2 - x_2) + \alpha\int_0^\infty J_1(v_2 - q)\psi_1(q)dq\right\} \tag{A11}$$

The holding and backlog costs are assessed at the inventory level x_2, and our problem becomes a minimization on the total cumulative base stock level, v_2 alone.

Case (b) $x_2 < w_2^*$

In this case we will have $w_2 = w_2^*$ as was shown by the evaluation of equation (A8), if an additional restriction that $v_2 \geq w_2$ does allow w_2 to assume the optimal value w_2^* meaning that $v_2 \geq w_2^*$. Otherwise we must have $w_2 = v_2 < w_2^*$ due to the convexity of $\tilde{L}(x)$. Once again, the minimization is solely over the total cumulative base stock level since we have the optimal immediate base stock level from (A8) in this case. Therefore, we have the following two expressions:

1. If $v_2 \geq w_2^*$,

$$J_2(x_2) = c(w_2^* - x_2) + L_2(w_2^*) + \min_{v_2 \geq x}\left\{c_1(v_2 - x_2) + \alpha\int_0^\infty J_1(v_2 - q)\psi_1(q)dq\right\} \tag{A12}$$

2. If $v_2 < w_2^*$

$$\begin{aligned} J_2(x_2) &= \min_{w_2^* \geq v_2 \geq x_2}\left\{c(v_2 - x_2) + L_2(v_2) + c_1(v_2 - x_2) + \alpha\int_0^\infty J_1(v_2 - q)\psi_1(q)dq\right\} \\ &= c(w_2^* - x_2) + L_2(w_2^*) + \min_{w_2^* \geq v_2 \geq x}\left\{c_1(v_2 - x_2) + c\ (v_2 - w_2^*) + L_2(v_2) - L_2(w_2^*) + \alpha\int_0^\infty J_1(v_2 - q)\psi_1(q)dq\right\} \end{aligned} \tag{A13}$$

If we define $\Gamma_2(v_2)$ by:

$$\begin{aligned}\Gamma_2(v_2) &= c(v_2 - w_2^*) + L_2(v_2) - L_2(w_2^*) \qquad && v_2 < w_2^* \\ &= 0 && v_2 \geq w_2^* \end{aligned} \tag{A14}$$

Then by the use of $\tilde{L}_2(x)$ and $\Gamma_2(v_2)$, we may represent expressions (A11), (A12), and (A13) by a single expression which depends on the relationship between v_2, w_2^*, and the current inventory position, x_2.

$$J_2(x_2) = \tilde{L}_2(x_2) + \min_{v_2 \geq x_2}\left\{ c_1(v_2 - x_2) + \Gamma_2(v_2) + \alpha \int_0^\infty J_1(v_2 - q)\psi_1(q)dq \right\} \tag{A15}$$

It is essential to notice that $\Gamma(v_2)$ is a continuous convex function of v in order to achieve the minimization. To determine the optimal ordering level for the two-period model, we obtain from (A15) the following expression which includes all terms that are a function of v_2 (i.e., the variable that is the subject of interest in the minimization defined in (A15)):

$$F_2(v_2) = c_1(v_2) + \Gamma(v_2) + \alpha \int_0^\infty J_1(v_2 - q)\psi_1(q)dq \tag{A16}$$

The derivative $F_2'(v_2)$ approaches $c_1 + \alpha h > 0$ as v_2 approaches infinity, and at $v_2 = w_2^*$ the derivative is given by

$$F_2'(w_2^*) = c_1 + \alpha \int_0^\infty J_1'(w_2^* - q)\psi_1(q)dq \tag{A17}$$

Two cases may now be considered, according to the sign of $F_2'(w_2^*)$:

Case (a) $F_2'(w_2^*) < 0$

Since $F_2(v_2)$ is convex in v_2 due to the convexity of $\Gamma(v_2)$ and $J_2(v_2)$, there exists a unique positive v_2^* such that $v_2^* > w_2^*$ and

$$F_2'(v_2^*) = 0 = c_1 + \alpha \int_0^\infty J_1'(v_2^* - q)\psi_1(q)dq \tag{A18}$$

It follows that we have in (A15) $v_2 = v_2^*$ if $x_2 <$ v_2^* and $v_2 = x_2$ if $x_2 \geq v_2^*$, so v_2^* is the solution to the total cumulative base stock level. Therefore, choosing appropriate forms of $\tilde{L}_t(x_2)$ and $\Gamma_t(v_2)$ from (A10) and (A14), respectively, we obtain from (A15)

$$\begin{aligned}
J_2(x_2) &= c_0(w_2^* - x_2) + c_1(v_2^* - w_2^*) + L_2(w_2^*) + \alpha \int_0^\infty J_1(v_2^* - q)\psi_1(q)dq && x_2 < w_2^* \\
&= c_1(v_2^* - x_2) + L_2(x) + \alpha \int_0^\infty J_1(v_2^* - q)\psi_1(q)dq && w_2^* \leq x_2 < v_2^* \\
&= L_2(x_2) + \alpha \int_0^\infty J_1(x_2 - q)\psi_1(q)dq && x_2 \geq v_2^*
\end{aligned} \tag{A19}$$

From (A19) we derive that the optimal policy is to order:

$$\begin{aligned}
f_2 &= \begin{cases} w_2^* - x_2 & x_2 < w_2^* \\ 0 & otherwise \end{cases} \\
s_2 &= \begin{cases} v_2^* - w_2^* & x_2 < w_2^* \\ v_2^* - x_2 & \hat{x}_2 \leq x_2 < v_2^* \\ 0 & x_2 \geq v_2^* \end{cases}
\end{aligned} \tag{A20}$$

The derivative of $J_2(x_2)$ is given by

$$\begin{aligned}
J_2'(x_2) &= -c_0 && x_2 < w_2^* \\
&= -c_1 + L_2'(x_2) && w_2^* \leq x_2 < v_2^* \\
&= L_2'(x_2) + \alpha \int_0^\infty J_1'(x_2 - q)\psi_1(q)dq && x_2 \geq v_2^*
\end{aligned} \tag{A21}$$

It is noticed that $J_2(x_2)$ is convex in x_2, $J_2'(x_2) \geq -c_0$ for all x.

Case (b) $F_2'(w_2^*) \geq 0$

We first recall that for $x_2 < w_2^*$ $F_2'(x_2)$ is given by

$$F_2'(x_2) = c_0 + L'(x_2) + \alpha \int_0^\infty J_1'(x_2 - q)\psi_2(q)dq$$

Therefore, there exists a unique v_2^* such that $w_2^* = v_2^*$ (and as is required by the sign of the derivative of F at w_2^*) and

$$F_2'(v_2^*) = 0 = c_0 + L'(v_2^*) + \alpha \int_0^\infty J_1'(v_2^* - q)\psi_2(q)dq$$

By the use of this v_2^* we obtain from (A15)

$$\begin{aligned} J_2(x_2) &= c_0(v_2^* - x_2) + L_2(v_2^*) + \alpha \int_0^\infty J_1(v_2^* - q)\psi_1(q)dq \qquad & x_2 < v_2^* \\ &= L_2(x_2) + \alpha \int_0^\infty J_1(x - q)\psi_1(q)dq & x_2 \geq v_2^* \end{aligned} \tag{A22}$$

We observe from (A22) that $w_2 = v_2 = v_2^*$ for this case and it follows that the optimal policy is to order:

$$\begin{aligned} f_2 &= \begin{cases} v_2^* - x_2 & x_2 < v_2^* \\ 0 & x_2 \geq v_2^* \end{cases} \\ s_2 &= 0 \qquad\qquad \forall x_2 \end{aligned} \tag{A23}$$

The derivative of $J_2(x_2)$ is given by

$$J_2'(x_2) = -c_0 \qquad\qquad x_2 < v_2^*$$

$$= L'(x_2) + \alpha \int_0^\infty J_1'(x_2 - q)\psi_1(q)dq \qquad\qquad x_2 > v_2^* \tag{A24}$$

$J_2(x_2)$ is convex in *x*, $J_2'(x_2) \geq -c_0$ for all *x*.

This establishes the desired result and the two-period model is complete.

In order to complete the induction we assume that optimal policy derivation holds true for the *n*-period problem. We rewrite the key equations necessary to proceed to period *n*+1.

We can express (A6) for the general n-period problem as (A25) and likewise express (A15) as equation (A26).

$$J_n(x_n) = \min_{s_n \geq 0,\, f_n \geq 0} \left\{ c_0 f_n + c_1 s_n + L_n(x_n + f_n) + \alpha \int_0^\infty J_{n-1}(x_n + f_n + s_n - q)\psi_{n-1}(q)dq \right\} \tag{A25}$$

$$J_n(x_n) = \tilde{L}_n(x_n) + \min_{v_n \geq x_n} \left\{ c_1(v_n - x_n) + \Gamma_n(v_n) + \alpha \int_0^\infty J_{n-1}(v_n - q)\psi_{n-1}(q)dq \right\} \tag{A26}$$

We define the following which corresponds to (A17)

$$F_n'(w_n^*) = c_1 + \alpha \int_0^\infty J_{n-1}'(w_n^* - q)\psi_{n-1}(q)dq$$

Case (a) $F_n'(w_n^*) < 0$

The cost equation (A27) is written as in (A19) and the corresponding policy follows in (A28).

$$J_n(x_n) = c_0(w_n^* - x_n) + c_1(v_n^* - w_n^*) + L_n(w_n^*) + \alpha \int_0^\infty J_{n-1}(v_n^* - q)\psi_{n-1}(q)dq \qquad x_n < w_n^*$$
$$= c_1(v_n^* - x_n) + L_2(x_n) + \alpha \int_0^\infty J_{n-1}(v_n^* - q)\psi_{n-1}(q)dq \qquad w_n^* \le x_n < v_n^*$$
$$= L_n(x_n) + \alpha \int_0^\infty J_{n-1}(x_n - q)\psi_{n-1}(q)dq \qquad x_n \ge v_n^* \tag{A27}$$

$$f_n = \begin{cases} w_n^* - x_n & x_n < w_n^* \\ 0 & otherwise \end{cases}$$
$$s_n = \begin{cases} v_n^* - w_n^* & x_n < w_n^* \\ v_n^* - x_n & w_n^* \le x_n < v_n^* \\ 0 & x_n \ge v_n^* \end{cases} \tag{A28}$$

Case (b) $F_n'(w_n^*) \ge 0$

The cost equation (A29) is written as in (A22) and the corresponding policy follows in (A3).

$$J_n(x_n) = c_0(v_n^* - x_n) + L_n(v_n^*) + \alpha \int_0^\infty J_{n-1}(v_n^* - q)\psi_{n-1}(q)dq \qquad x_n < v_n^*$$
$$= L_n(x_n) + \alpha \int_0^\infty J_{n-1}(x_n - q)\psi_{n-1}(q)dq \qquad x_n \ge v_n^* \tag{A29}$$

$$f_n = \begin{cases} v_n^* - x_n & x_n < v_n^* \\ 0 & x_n \ge v_n^* \end{cases}$$
$$s_n = 0 \qquad \forall x_n \tag{A30}$$

We now briefly describe and analyze the n+1 period problem in order to proceed with the induction. To begin, $J_{n+1}(x_{n+1})$ will be defined by expressions identical to (A25) except Jn is replaced by J_{n+1}. This expression is further simplified to an expression identical to (A26), with J_n replaced by J_{n+1}, through the same arguments used in the two-period problem based on cases (a) and (b). Then, $F_{n+1}(v)$ will be defined by an expression identical to $F_n(v)$ except J_n is replaced by J_{n+1}. If $F'_{n+1}(w_{n+1}{}^*) < 0$, we may immediately conclude that there exists a unique $\hat{y}_{n+1}$ such that $v_{n+1}{}^* > w_{n+1}{}^*$ and $F'_{n+1}(v_{n+1}{}^*) = 0$. Else, $J_{n+1}(x_{n+1})$ is given by an expression identical to (A29) with $v_n{}^*$ and $J_n(x_n)$. The resulting optimal policy is identical to (A30) except the solution $v_n{}^*$ is replaced with $v_{n+1}{}^*$. It will be noticed that $J_{n+1}(x_{n+1})$ is convex in x.

We now have all the necessary ingredients of an inductive proof, and we summarize the preceding results as

Theorem 1:

For each $n \geq 2$ there exists a unique positive x_{n} which, together with $w_n{}^*$ determined by (A9), uniquely determines the optimal policy for the n period model as follows:

Case A If $v_n{}^* > w_n{}^*$ it is optimal to order:

For $x_n < w_n{}^*$, amount $w_n{}^* - x_n$ at c_0 and amount $v_n{}^* - w_n{}^*$ at c_1;

For $w_n{}^* \leq x_n \leq v_n{}^*$, amount $v_n{}^* - x_n$ at c_1;

For $x_n \geq v_n{}^*$, none.

Case B If $v_n{}^* \leq w_n{}^*$ it is optimal to order:

For $x_n < v_n{}^*$, amount $v_n{}^* - x_n$ at c_0;

For $x_n \geq v_n{}^*$, none;

Furthermore, the following properties hold:

(i) v_n^* is a unique root of an equation

$$c_1 + \Gamma'(v_n) + \alpha \int_0^\infty J'_{n-1}(v_n - q)\psi_{n-1}(q)dq = 0$$

(ii) Case A $v_n^* > w_n^*$:

$$\begin{aligned} J'_n(x_n) &= -c_0 && x_n < w_n^* \\ &= -c_1 + L'(x_n) && w_n^* \le x_n < v_n^* \\ &= L'(x_n) + \alpha \int_0^\infty J'_{n-1}(x_n - q)\psi_{n-1}(q)dq && x_n \ge w_n^* \end{aligned}$$

Case B $v_n^* \le w_n^*$:

$$\begin{aligned} J'_n(x_n) &= -c_0 && x_n < w_n^* \\ &= L'(x_n) + \alpha \int_0^\infty J'_{n-1}(x_n - q)\psi_{n-1}(q)dq && x_n > w_n^* \end{aligned}$$

CHAPTER 3:

A Two-Stage Inventory and Production Model for Perishable Products with Markovian Demand under Various Demand Fulfillment Scenarios

Abstract: We study a two-stage perishable inventory problem motivated by challenges seen in hospital pharmacy medication management. Inpatient hospital pharmacies stock pharmaceutical drugs in two stages, namely raw material: Stage 1 and the finished good: Stage 2 (i.e. a dispensed form such as intravenous, IV). The raw material is converted through one or more production steps into the final finished good. While both stages of material are perishable, the medication in the finished good form is highly perishable compared to the raw material. Demand for the medication depends on the population and the condition of patients in the hospital. In this research we use a stochastically changing 'demand state' as a surrogate for this patient condition information. We model these changes as a Markov chain to determine the optimal, state-dependent two-stage inventory and production policy. We define and evaluate two ordering and production scenarios which depend on production capabilities. We prove the existence of optimal solutions for both scenarios and show numerical examples which demonstrate the behavior of the optimal solutions for various cost structures.

3.1 Motivation

In this paper we study a two-stage perishable inventory problem motivated by challenges in hospital pharmacy medication management. Inpatient hospital pharmacies stock pharmaceutical drugs in two stages, namely raw material: Stage 1 and the finished good: Stage 2 (i.e., a dispensed form such as an IV). The raw material is converted through one or more production steps into the final finished good. The medication, in the finished good form, is highly perishable in comparison to the raw material. Thus, production decisions affect the shelf life of the final product. Multi-echelon inventory research for perishable products has been limited, in comparison to multi-echelon research for products with infinite shelf lives. Patient information, such as diagnosis codes, can be mapped to a distribution of demand for dosage of a drug. Since a patient's condition may change, we use a stochastically changing 'demand state' (modeled as a Markov chain) as a surrogate for patient condition information. The goal of the optimal inventory and production policy is to minimize waste, expediting, purchasing/production, and holding costs by incorporating the patient demand process.

In a national survey of 374 U.S. pharmacy directors on the impact of drug shortages in acute care hospitals, 75% of respondents indicated they were forced to either purchase the drug off-contract from their current vendor, borrow the drug from another institution, or purchase the drug from an alternative vendor at an increased price (Baumer et al., 2004). In addition to the negative impact on purchasing costs, two-thirds of respondents reported delayed or canceled medical procedures due to drug shortages. The results of the survey attributed an inter-quartile range of approximately $33 - $300 million in additional costs to the U.S. healthcare system to these shortages (Baumer et al., 2004).

Other sources have described specific impacts of drug shortages and overages on hospital costs and efficiencies. In a large scale study spanning one year in a large tertiary hospital, Gillerman and Browning (2000) found of six anesthesia drugs they tracked, waste costs from five of them contributed to 26% of the total cost of the departmental pharmaceutical budget and 2% of the total of annual hospital expenditures. Handfield (2007) found that there appears to be a strong relationship between missed doses and medical errors. Drug substitutions due to stockout, which have different dosing regimens, could result in medical errors. Other serious consequences of traditional pharmacy purchasing include missed contract compliance, excess inventory levels, frequent stock-outs and costly deliveries, workflow interruptions and expensive rework, and increased health system labor requirements (Alverson, 2003). These problems are exacerbated by the large number (typically more than 2,000) of perishable drugs in the hospital and the manual inventory stock keeping and daily order entry.

While some drug shortages are uncontrollable (e.g. due to a natural disaster), improper inventory management can result from the clinical versus procurement expertise of those managing the inventory (Alverson, 2003). In an attempt to improve their drug delivery system, many hospitals have implemented automated systems such as Pyxis machines (Handfield, 2007). While these machines allow for greater efficiency compared to drug delivery in patient specific trays, they are labor intensive and inefficient without the implementation of proper inventory management policies. Inaccurate inventory policies necessitate refilling the machine several times per day due to machine stock outs. Handfield (2007) found that many hospitals have not experienced the anticipated benefits of optimizing inventory after deploying Pyxis; in fact, some have reported decreased performance.

Problems related to the management of perishable inventories arise in many sectors including blood banks, foodstuffs, chemicals, and drugs (see Nahmias, 1982 for a complete review). Motivated by inventory issues of perishable medications in the hospital pharmacy, we focus on a two-stage model which consists of raw materials, the upstream echelon, and finished goods, the downstream echelon. The modeling framework we propose is general enough to be used for perishable products that are further processed into perishable sub-products (e.g., fresh grocery items and pharmaceuticals). We consider two decisions: (1) an ordering decision at Stage 1 and (2) a production decision to convert Stage 1 product into Stage 2 inventory.

In the following discussion we present a brief review of single-stage, periodic review perishable inventory problems and then focus our discussion on multi-echelon systems.

3.2 Relevant Literature

Perishable inventory research to date has made simplifying assumptions deemed necessary to make the problem tractable but limit its applicability to more complex inventory problems such as the pharmacy problem. We refer the reader to Goyal and Giri (2001) for the most recent review of perishable inventory models. The majority of perishable inventory literature has focused on a single inventory echelon. Early research (refer to Veinott (1960) and Dave

(1979)) considered a single inventory stage with deterministic demand. Those that model the stochastic nature of the demand function assume that it is essentially independent of the system (exogenously defined) and stationary (e.g., Lian and Liu, 1999b). Most of the research in this area has evaluated presumed inventory policies such as an (s, S) or fixed order quantity policy under stochastic, stationary demand.

Inventory research for perishable products with multiple echelons has been limited. In general, multi-echelon research for perishable products can be classified by issuing method, where issuing is the manner by which product moves from an upstream echelon to a downstream echelon. One issuing method defines the echelons by inventory location, and the product is issued from one location to another. One area of this research has focused on the allocation of blood products from a distribution center to customers (refer to Cohen et. al (1981) and Prastacos (1981)). Ketzenberg and Ferguson (2008) extend the inventory research to the context of perishable foods and model a two-stage serial supply chain where the retailer follows an EOQ ordering policy. They evaluate the value of information sharing between the supplier and retailer. A second issuing method considers a product which moves automatically to the next inventory stage due to its age (refer to Goh et al. (1993) and Lystad (2006)). Much of the blood platelet research falls under this category where the blood is classified as ‘any’ or ‘young’ due to the patient requirements (Haijema et al. 2007). Haijema et al. (2007) extend previous research by including a production lead time of one period and assuming different demand characteristics for each day of the week.

In contrast to the aforementioned research, we not only make a production decision (for the finished good) but also an inventory decision for the raw material. Demand is non-stationary and dependent on the patient characteristics, which predict demand for the finished good. In the medication inventory environment, we assume that as product moves between echelons the shelf life of the product changes drastically. This third type of issuing involves progression to downstream echelons as a result of further production-based activities. Fujiwara et al. (1997) models such a case motivated by the perishable food industry, where there are two-stages of inventory: whole product (e.g., meat carcass) and sub-products (e.g., sirloin steak) produced from the whole product. Fujiwara et al. (1997) assume that the whole product's shelf life is much longer than its sub-products due to the cold storage process. Therefore they consider a single cycle where the cycle length is the shelf life of the whole product and is made up of multiple periods that represent the shelf life of the sub-product. They derive a mathematical model for determining the base stock level of the whole product and a stationary base stock level for each sub-product when demand occurs at the second stage.

We consider a two-stage perishable inventory problem where the production activities affect the remaining shelf life of the product. We formulate the problem as a Markov decision process (MDP) which we describe in more detail in Section 3.3. In addition, we define and model two ordering and production scenarios relevant for pharmaceutical drug management. We define a stochastic demand function that is a property of the system instead of an exogenously defined random variable. This nonstationary demand function captures the complexity of the hospital setting and the interrelationship between the patient's condition, the patient's demographics, and drug utilization. Further, in contrast to Fujiwara et al. (1997), we assume a cost minimization model with no second-stage salvage value (i.e., expired products are not used), demand that is nonstationary, and shortages that can be fulfilled in two different ways. We prove that the production decision is state dependent (i.e., may vary through time) versus Fujiwara's stationary production policy. We also show that the optimal production and inventory policies are dependent on the shortage fulfillment method.

In Section 3.3, we define the model and notation for a single perishable product under two different fulfillment scenarios. We prove the existence of an optimal solution to the first stage's initial inventory level and the production quantity of raw material into finished good. In Section 3.4 we provide some numerical examples which show the behavior of the solutions as costs vary. We conclude the research in Section 3.5 with a discussion of the results.

3.3 Model Specification

In the hospital pharmacy setting, medication inventories are generally stored in one of two stages. Stage 1 is composed of raw materials that have longer shelf lives, six months to one year. Stage 2 is the finished good (e.g., IV) made from processing the raw material and has a shelf life of less than one day. Patient demand is for the Stage 2 items. Regardless of the stage of inventory, the shelf life is deterministic and finite. We assume that the pharmacist reviews the inventory level of the raw material and the state of the demand process on a periodic basis (e.g., once per shift) and plans the production of Stage 2 products.

We now define the necessary costs for this problem. First, there is a holding cost h_1 for Stage 1 inventory per unit per period. There is a unit ordering cost c_1 for the raw material and per unit production cost c_2 for the finished good. If there is a stockout at Stage 2, then there is an additional production cost of b_1 if there is sufficient inventory in Stage 1. If there is a stockout of both Stage 1 and Stage 2 material, the remaining patient demand is fulfilled via an emergency channel at a penalty cost of b_2. In the context of hospital pharmacy operations, stockouts are fulfilled from either a supplier or another hospital for an additional expedited fee. We define two scenarios that dictate the type of expediting (internal vs. external) that can be done in order to fulfill shortages. The first scenario restricts expediting to external suppliers, and the second scenario permits internal (i.e., production from on-hand raw material) and external expediting. Finally, there is a waste cost w_1 (w_2) associated with expired Stage 1 (Stage 2) material.

In the following section we model the problem as a single cycle, of length T, with the cycle length defined by the shelf life of the raw material. The cycle is subdivided into n periods, which represent the frequency of the periodic reviews and production decisions during the cycle. The period length is defined as the shelf life of the finished good.

3.4 Model Assumptions and State Specification

In this section, we develop the MDP model for the two-stage, perishable pharmaceutical inventory problem with the objective of determining the inventory and production policy that minimizes total expected cost. The following assumptions apply to our model.

1. The inventory levels are strictly non-negative (no backlogging), and shortages must be expedited via an internal or external source depending on the model under consideration.
2. Raw material (Stage 1 inventory) is ordered prior to the start of the cycle.
3. Finished goods (Stage 2 inventory) may be produced each period.
4. The shelf lives of both inventory stages are finite.
5. The shelf life of the finished good is equivalent to the length of a period.
6. Customer demand is a nonstationary Markov demand process (i.e., state dependent).

The MDP state definition involves two components: the raw material inventory level and the state of the demand process. We define i_t^r as the inventory level of raw material with t ($t = 1,\ldots, n$) periods remaining. Given an initial raw material order quantity R for an n period problem (i.e., with n periods remaining), the relationship between the raw material inventory and finished goods production decision, x_n, is defined by:

$$i_n^r = R$$
$$i_{n-1}^r = R - x_n$$

The production decision (x_t) is constrained to be less than or equal to the inventory position of the raw materials, $x_t \le i_t^r$. Therefore the quantity of finished goods produced in period t is capacitated by the quantity of raw material available in period t.

The demand process, Z_t, is defined by K states that capture patient information that can be correlated to patient demand. For example, the demand state could represent the number of patients in a particular unit. The number of patients in that unit changes stochastically as patients arrive and depart each period. Alternatively, the demand state could represent a patient characteristic that maps directly to demand, for example, platelet count. Demand is assumed to be nonstationary and represented by a Markov chain. A transition probability matrix, Q, represents the demand state transition probabilities. In any period, the probability of transitioning from demand state i to demand state j is $q_{ij} = P(Z_{t+1} = j| Z_t = i)$. When the state of the Markov demand process (i.e., demand state) is z_t, we assume that demand is distributed according to the pdf $f_{Z_t}(\cdot)$. Finally, we define α_j as the probability that the demand process is in state j in the first period.

We define the 'system' to represent the dynamics of the Markovian demand and raw material inventory processes. If we define the state in period t as $S_t = (Z_t, i_t^r)$, then system level transition probabilities can be expressed the probability that the state transitions from state i to j as $p_{ij} = P(S_{t+1} = j| S_t = i)$. During each period, three events determine the system state transitions: (1) Stage 2 production, (2) customer demand fulfillment, and (3) demand state transition.

We assume that there are two scenarios for the fulfillment of insufficient finished goods. **Scenario 1**: The first involves fulfillment of demand via an external source. This situation occurs when the production lead time for converting the raw material to a finished good is too long (or resources are not available during the period for additional internal production). **Scenario 2**: The second option involves producing additional finished goods internally, from existing raw material, to fulfill the excess demand. In this case, if the remaining raw material is insufficient to fulfill the backlog, the difference must be fulfilled via the external source. Contrary to the first case, the second option is viable for drugs that can be produced quickly and can only be utilized if there is sufficient raw material on hand. These two cases are defined further in the next section.

3.4.1 MDP Cost Formulation for Fulfillment Scenario 1: Demand Fulfilled via External Supplier

As mentioned in Section 3.1, depending on the type of medication, shortages may have to be fulfilled by an external supplier or hospital. This situation arises for various reasons. First, there may be a resource constraint which prohibits additional production runs, or second, the production time makes it impossible to produce additional units and dispense them in the same day. Figure 3-1 depicts how the inventory position of the raw material changes during each period. A timeline during a generic period *t* is shown in Figure 3-2 for a single cycle problem horizon.

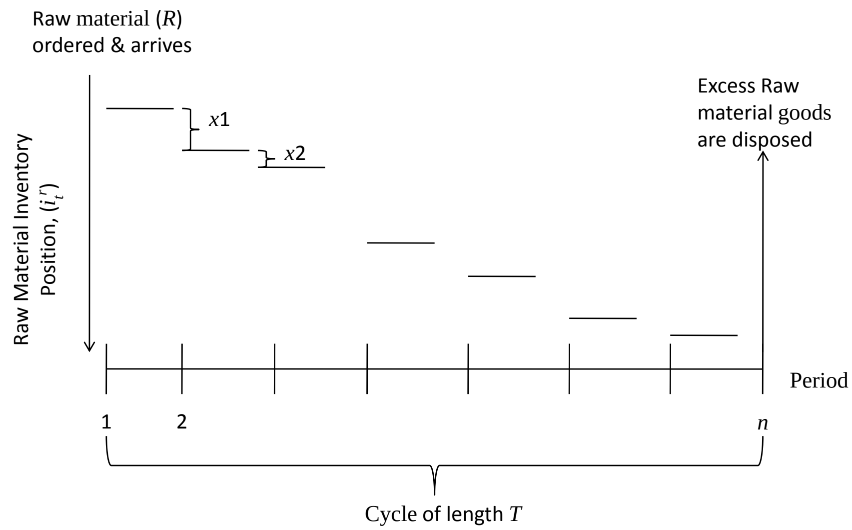


Figure 3-1: Raw material inventory depletion throughout a single cycle

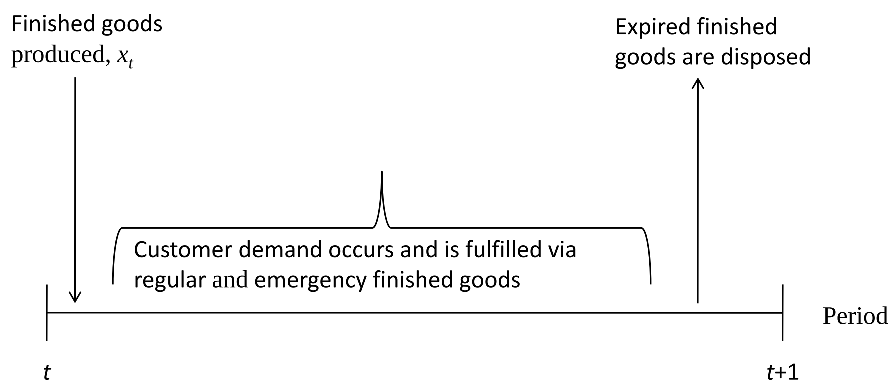


Figure 3-2: Timeline of events for finished goods inventory during a period

Equation (3.1) defines the single period expected cost function, g_t, given that the state at time t is (z_t, i_t^r), and excess demand is fulfilled via an external source. Recall that prior to the start of the cycle the raw material order decision is made. We define the cost of this initial ordering decision as g_n in Equation (3.1a). The linear cost per unit of raw material ordered at time zero is c_1. Equation (3.1b) is the single period expected cost given the production decision with t periods remaining is x_t. There are n-1 production decision periods in an n period problem.

$$\begin{aligned}
&(a)\ \ g_n(\cdot,R) = c_1 R \\
&(b)\ \ g_t(z_t, i_t^r, x_t) = \left\{ c_2 x_t + h_1(i_t^r - x_t) + \int_0^{x_t} w_2(x_t - m_t) f_{z_t}(m)dm + \int_{x_t}^{\infty} b_2(m_t - x_t) f_{z_t}(m)dm \right\} \quad for\ x_t \le i_t^r \\
&t = 1,\ldots,n-1
\end{aligned} \tag{3.1}$$

Assuming an n period cycle, with n periods remaining (i.e., prior to the start of the cycle) the only decision is the initial raw material ordering quantity (Equation 3.2a). The total expected cost equation, v_t, for t periods remaining is defined by Equation 3.2b. The first term (of Eqn. 3.2b) is the single period expected cost, and the second term is the expected total cost with t-1 periods remaining. In the final period of the cycle no decisions are made but a final disposal cost, w_1, per unit of leftover raw material is incurred (Equation 3.2c).

$$\begin{aligned}
&(a)\quad v_n(\cdot,R) = c_1 R + \sum_{j=1}^{K} \alpha_j v_{n-1}(j, i_{n-1}^r = R) \\
&(b)\quad v_t(z_t, i_t^r) = \min_{x_t \le i_t^r} \left\{ g_t(z_t, i_t^r) + \sum_{j=1}^{K} q_{z_t j} E[v_{t-1}(j, i_{t-1}^r = i_t^r - x_t)] \right. \qquad t = 1,\ldots,n-1 \\
&(c)\quad v_0(z_0, i_0^r) = w_1 * i_0^r
\end{aligned} \tag{3.2}$$

We use backwards recursion to solve this finite horizon MDP for the production decision for each state. With n periods remaining, the optimal solution to the initial raw material ordering problem corresponds to the R^* associated with minimum cost, where R^* = argmin$\{v_n(\cdot, R)\}$ Therefore, we are interested in finding the optimal initial ordering quantity R^* and, each period, the optimal production quantity $x_t^*(z_t, i_t^r)$.

3.5 Optimal Ordering and Production Policy

We now prove, by induction, that optimal solutions exist for the initial raw material ordering decision and finished good production decision. We extend the proof of Gallego and Hu (2004) to include two stages of perishable inventory. While the proof is similar, our problem is distinct. They define a single-stage (nonperishable) inventory model where demand is fulfilled by the quantity ordered and the inventory on hand. In our system, demand is fulfilled by the second stage perishable product and any expedited production. Finally, our production and ordering decisions are linked by the multi-echelon nature of the problem. We begin the proof with a single-cycle, single period problem. Given a production quantity x_1, the single period/single-cycle cost function is:

$$L_1(z_1, R, x_1) = \left\{ c_1 R + c_2 x_1 + w_1(R - x_1) + \int_0^{x_1} w_2(x_t - m) f_{z_1}(m)dm + \int_{x_1}^{\infty} b_2(m - x_1) f_{z_1}(m)dm \right\} \quad for\ x_1 \le R \tag{3.3}$$

The total expected cost function is defined as: $v_1(z, R) = \min_{x \le R}\{L_1(z, R, x)\}$. We take the partial derivatives of Equations (3.3) with respect to the decision variables R and x in order to derive the relationship between the decision variables and the cost function. Going forward we omit the time subscript.

$$L_R(z,R,x) = \frac{\partial L(z,R,x)}{\partial R} = c_1 + w_1$$

$$L_x(z,R,x) = \frac{\partial L(z,R,x)}{\partial x} = c_2 - w_1 + w_2 F_z(x) - b_2(1 - F_z(x))$$

In the following set of Lemmas we prove the convexity of the total expected cost function and the existence of an optimal raw material order quantity and production quantity decisions.

Lemma 1: *For a fixed (z, R) there is a finite $x_1^*(z, R)$ which exists such that L is convex in x and* $L_x(z,R,x_1^*(z,R)) = 0$

Proof:

Using the Dominated Convergence Theorem (DCT) we can observe that the following relationships exist.

$$\lim_{x \to -\infty} L_x(z,R,x) = c_2 - w_1 - b_2 < 0$$
$$\lim_{x \to \infty} L_x(z,R,x) = c_2 - w_1 + w_2 > 0$$

For these relationships to hold, b_2 must be greater than the production cost (i.e., expediting must be more costly than production) and the cost w_2 must be greater than or equal to w_1, i.e., the finished good is at least as valuable as the raw material. Therefore, for a fixed (z, R) the single period cost function $L(z,R,x)$ is convex in x and an optimal solution to the minimization function $v_1(z,R)$ exists and is defined as $x_1^*(z, R)$.

Lemma 2: *Define R(z) as the raw material order quantity when the demand process is initially in state z. For a fixed demand state, z, there is a finite $R^*(z)$ which exists such that*

$L_R(z, R^*(z), x) = 0$

Proof:

$$L_R(z,R,x) = \frac{\partial L_R(z,R,x)}{\partial R} = c_1 + w_1 \text{ is convex for any } R$$

Therefore, for a fixed *z* and any production quantity, *x*, the single period cost function $L(z,R,x)$ is convex in *R*, and an optimal solution to the minimization function $v_1(z,R)$ exists and is defined as $R^*(z)$.

Theorem 1: *For a given demand state, z, there exists an optimal solution to the raw material inventory decision, $R^*(z)$ and the finished goods production quantity $x_1^*(z, R^*)$, such that the value function $v_1(z, R^*, x_1^*(z, R^*))$ is minimized.*

Proof: Since $v_1(z,R) = \min_{x \le R}\{L_1(z,R,x)\}$ and $L_1(z,R,x)$ has been shown (in Lemmas 1 and 2) to be convex in *R* and *x* (for a fixed *z*), then convexity of the total expected cost function is preserved. Thus, the existence of optimal solutions for the single period problem is proven. Lemmas 1 and 2 and Theorem 1 are used to extend Theorem 1 for the multi-period problem.

Lemma 3: For any fixed *z*:

(1) $v_n(z_n, i_n^r = R)$ is convex in (*R*, *x*)

(2) There exists an optimal solution $x_n^*(z_n, i_n^r)$ such that $v_n'(z_n, i_n^r, x_n(z_n, i_n^r)) = 0$ for all *n*.

(3) There exists an optimal solution $R(z)$ that minimizes the total expected cost over n periods

Proof: As shown by Lemmas 1 and 2, Lemma 3 holds when $n = 1$. We assume that it holds for a problem with n-1 periods. We define $H_{n-2}(z_{n-2}, i_{n-2}^r = R - x_{n-1})$ as the total cost for the n-2 period remaining problem and $L_{n-1}(z_{n-1}, i_{n-1}^r = R, x_{n-1})$ as the expected single period cost at when n-1 periods are remaining and production decision is x_{n-1}. The total expected cost with n-1 periods remaining is defined as:

$$
\begin{aligned}
v_{n-1}(z_{n-1}, i_{n-1}^r) &= \min_{x_{n-1} \le i_{n-1}^r} \{L_{n-1}(z_{n-1}, i_{n-1}^r, x_{n-1}) + \sum_{z_{n-2}=1}^{K} q_{z_{n-1}z_{n-2}} E[v_{n-2}(z_{n-2}, i_{n-2}^r = i_{n-1}^r - x_{n-1})]\} \\
&= \min_{x_{n-1} \le i_{n-1}^r} \{L_{n-1}(z_{n-1}, i_{n-1}^r, x_{n-1}) + H_{n-2}(z_{n-2}, i_{n-2}^r)\}
\end{aligned} \tag{3.4}
$$

We assume that v_{n-1} (Equation 3.4) satisfies condition 1 through 3 of the lemma. In any period the raw material inventory level can be written as a function of the initial decision R and the periodic production decision x.

We will now show by induction that it holds for an n period problem. Let us define the following:

$$
\begin{aligned}
H_{n-1}(z_{n-1}, i_{n-1}^r = R - x_n) &= \sum_{z_{t-1}=1}^{K} p_{z_n z_{n-1}} E[v_{n-1}(z_{n-1}, i_{n-1}^r = R - x_n)] \\
H_{n-1,R}(z_{n-1}, i_{n-1}^r = R - x_n) &= \frac{\partial H_{n-1}(z_{n-1}, i_{n-1}^r = R - x_n)}{\partial R} \\
H_{n-1,x}(z_{n-1}, i_{n-1}^r = R - x_n) &= \frac{\partial H_{n-1}(z_{n-1}, i_{n-1}^r = R - x_n)}{\partial x_{n-1}}
\end{aligned}
$$

We assume that v_{n-1} is convex, for a fixed z, and thus H_{n-2} and L_{n-1} must be convex in R and x. For the n period problem, the expectation of v_{n-1} (i.e., H_{n-1}) is added to the function L_n. Convexity of H_{n-1} is preserved since it is the expectation of a convex function. We prove that L_n is convex and thus v_n is convex in (R, x). Following the same procedure as in Lemma 1 we take the partial derivative of L_n, with respect to x, to show the existence of the optimal solution $x_n^*(z, R)$.

$$\lim_{x \to -\infty} L_{n,x}(z_n, i_n^r) = c_2 - b_2 < 0$$
$$\lim_{x \to \infty} L_{n,x}(z_n, i_n^r) = c_2 - h_1 + w_2 > 0$$

From the relationships above we can show that for a given (z, R) there is an optimal finished good production quantity $x_n^*(z, R)$, and Lemma 3 part 2 is proven. Next, we prove Part 3 of Lemma 3. Similarly we focus on the partial derivative of L_n, with respect to R, and show that the conditions derived in Lemma 2 still hold: $L_{n,R}(z, R) = c_1 + h_1 > 0$. There exists an optimal initial raw material inventory level, R, for the n period problem. Finally, given the proof of Lemma 3 part 1 and 2, we explicitly state that v_n is convex in (R, x) since Equation (3.4) is the addition of two convex functions. Part 1 of the Lemma is proved. We state the results of Lemma 3 formally as Theorem 2.

Theorem 2: *For a given demand state, z, there exists an optimal solution to the raw material inventory decision* $R^*(z)$ *and the finished goods production quantity,* $x_n^*(z, R)$, *such that the value function* $v_n(z, R^*, x_n^*(z, R^*))$ *is minimized.*

Proof: Since $v_n(z_n, R) = \min_{x_n \le R}\{L_n(z_n, R) + H_{n-1}(z_{n-1}, R - x_n)\}$ and $L(z, R, x)$ and $H(z, R - x)$ have been shown (in Lemmas 1, 2, and 3, respectively) to be convex in *R* and *x* (for a fixed *z*), then convexity of the total expected cost function is preserved for the *n* period problem. The optimal finished good production quantities that are dependent on the system state and an initial raw material order quantity exist.

It may be more realistic to assume that while the raw material quantity is a continuous decision variable, the production decision is a discrete variable such as number of IV bags or doses. The following Proposition must hold in order to prove that a discrete optimal production decision exists. It is shown in Appendix B that the second difference of the total expected cost function under this scenario is non-negative and thus an optimal solution exists for the discrete production decision. We define *Z* as the set of integers and *X* as the subset of feasible discrete production values.

Proposition 1: If $\kappa(x)$ is a real-valued function of $x \in X$ where *X* is a bounded subset of *Z*, and if $\Delta^2\kappa(x) \ge 0$ for all $x \in X$, then every local minimum of $\kappa(x)$ is also a global minimum of $\kappa(x)$ on *X*.

Proof: See Appendix B.

In the following section we extend the model developed in Section 3.3.2 to include internal as well as external expediting options.

3.5.1 Fulfillment Scenario 2: Demand Fulfilled Internally First

As discussed in Section 3.1, some medication shortages can only be fulfilled externally due to resource or production time constraints. However, many medications can be produced and dispensed quickly. In this case we now assume that excess demand in a period can be met from additional production up to the limit of available raw material. Any demand beyond this is fulfilled from external sources. The difference in total expected cost of Scenario 1 versus Scenario 2 provides a bound on what managers should spend to become capable of external expediting. These insights are discussed further in Section 4. In order to represent this scenario, we redefine Equations (3.1b) and (3.2) as Equations (3.5) and (3.7). The holding cost term in Equation (3.5) includes the expected internal expediting cost that is produced from the raw material inventory level. We define the number of units of finished goods produced by internal expediting (i.e., after demand is realized) at time t as a_t. The expected quantity of internally expedited finished good in state (z_t, i_t^r) with production decision x_t is expressed as: $E[a_t \,|\, z_t, i_t^r, x_t] = \int_{x_t}^{i_t^r} (m - x_t) f_{z_t}(m)dm$. Going forward this is written as $E[a_t]$ to simplify notation. The fourth term has been added to represent the expected cost of internal expediting given that the planned production level is x_t.

$$g_t(z_t, i_t^r, x_t) = \left\{ \begin{array}{l} c_2 x_t + h_1(i_t^r - x_t - E[a_t]) + \int_0^{x_t} w_2(x_t - m) f_{z_t}(m)dm \\ + \int_{x_t}^{i_t^r} b_1(m - x_t) f_{z_t}(m)dm + \int_{i_t^r}^{\infty} b_2(m - x_t) f_{z_t}(m)dm \end{array} \right\} \tag{3.5}$$

Since raw material may be made into additional finished goods after demand, d_t, is realized, we redefine how the raw material inventory level changes from period to period. Equation

(3.6) captures the new material balance dynamics of the raw material inventory process. The quantity a_t is constrained by the quantity of raw material after the initial finished good is produced.

$$\begin{aligned} & i_{t-1}^r = i_t^r - x_t - a_t \\ & where\ a_t = \max(\min(d_t - x_t, i_t^r - x_t), 0) \end{aligned} \tag{3.6}$$

Using the revised single period expected cost equation (3.5) and the revised definition of i_{t-1}^r, the recursive cost function for this scenario is updated accordingly (Equation (3.7)) and is solved using backwards recursion. The system state space is now Markovian in the raw material inventory position as well as the demand state. The raw material inventory position is Markovian because of the internal expediting quantity that is dependent on the realization of demand from the stochastic, state-dependent demand function. The state transitions are more uncertain than those of the problem defined in Section 3.2.

$$\begin{aligned} & (a)\ v_n(\cdot, R) = c_1 R + \sum_{z_{n-1}=1}^{K} \alpha_{z_{n-1}} v_{n-1}(z_{n-1}, R) \\ & (b)\ v_t(z_t, i_t^r) = \min_{x_t \le i_t^r} \{ g_t(z_t, i_t^r) + \sum_{k=1}^{K} p_{z_t k} E[v_{t-1}(k, i_{t-1}^r = i_t^r - x_t - a_t)] \} \\ & t = 1, \ldots, n-1 \\ & (c)\ \ v_0(z_0, i_0^r) = w_1 * i_0^r \end{aligned} \tag{3.7}$$

In the following theorem we prove that the multi-period cost function is convex and therefore an optimal solution for R and x exist.

Theorem 3: *For a given demand state, z, there exists an optimal solution to the raw material inventory decision $R^*(z)$ and the finished goods production quantity $x_n^*(z, R)$, such that the value function $v_n(z, R^*, x_n^*(z, R^*))$ is minimized.*

Proof: The proof follows directly from the derivation of Theorem 2. The cost of external expediting must be greater than the cost of internal expediting in order for the conditions to be true. This is a reasonable ordering of the expediting costs.

Theorem 3 proves that: (1) the optimal finished good production decision is dependent on the system state, and (2) an optimal initial raw material ordering quantity for the multi-period problem, dependent on the initial state of demand, exists. Now that the existence of optimal solutions has been derived, we present a set of numerical examples. In the following section we compare the optimal ordering and inventory policies under the two cases outlined in Sections 3.2 and 3.3. We vary the cost parameters in order to gain insight into the behavior of the optimal solutions.

3.6 Numerical Examples

In this section we present a series of examples that demonstrate the effect of the cost parameter values and fulfillment method on the behavior of the optimal solution. The cost parameters for the baseline situation are defined in Table 3-1. For ease of discussion we assume that there are two demand states: (A) a discrete uniform (0, 6) and (B) a discrete uniform (0, 3). We can think of the demand states in the following way. The first demand state represents a patient with more variable demand and a larger upper limit on dose (i.e., an unstable patient). The second demand state represents a patient with less variable and lower demand (i.e., a more stable patient condition). The elements of the demand state transition matrix are p_{AA}=0.9, p_{AB} =0.1, p_{BA}=0.2, p_{BB}=0.8. For the following set of examples, we assume that the initial demand state is (A). A demand value of 1 is interpreted as one unit of the drug, for example, one injection or IV bag.

Table 3-1: Baseline costs for both internal and external fulfillment scenarios

	c_1	c_2	h_1	w_1	w_2	b_1	b_2	N
External Fulfillment	5	1	0.1	1	10	0	100	20
Internal Fulfillment	5	1	0.1	1	10	5	100	20

3.6.1 Impact of Expediting Costs

To explore the impact of expediting on the optimal decisions of both raw material and periodic production, we vary the parameters b_1 and b_2. Beginning with the case of external fulfillment (external source), we set b_2 = 50, 100, 200. Figure 3-3 demonstrates that the optimal initial raw material level (i.e., denoted by the black diamond) decreases as the cost of expediting decreases. The key cost relationship that drives this behavior is the decreasing ratio between shortage and finished good waste costs.

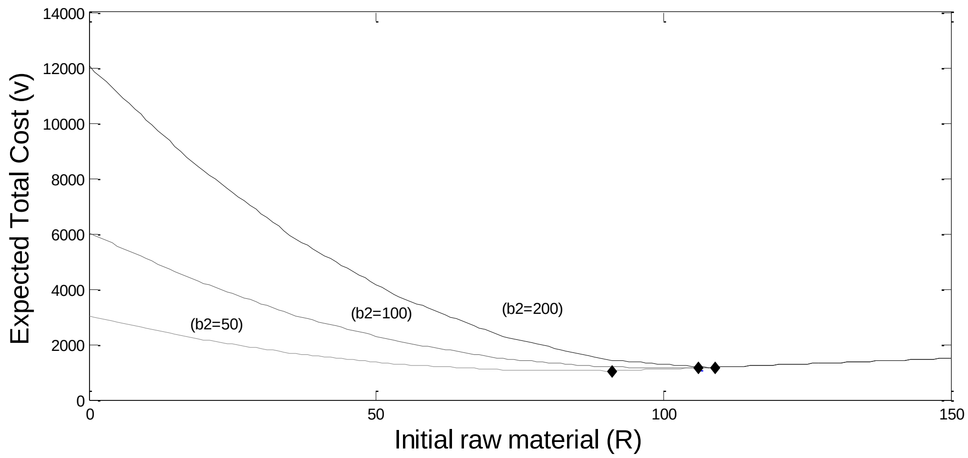


Figure 3-3: The impact of external fulfillment costs on the optimal initial raw material order quantity (denoted by the black diamond marker)

Next, we investigate the impact of having both internal and external fulfillment method capabilities on the total expected cost and the optimal initial raw material ordering decision. Keeping b_2 fixed at 100, we vary the cost of additional internal fulfillment b_1 from its baseline value of 5 to 50. As b_1 increases, we observe that the initial raw material inventory level increases. Figure 3-4 shows the difference in the optimal solution (R, denoted by the diamond marker) and total expected cost for the baseline costs of Fulfillment Scenario 1 (gray line) and two cases of Fulfillment Scenario 2.

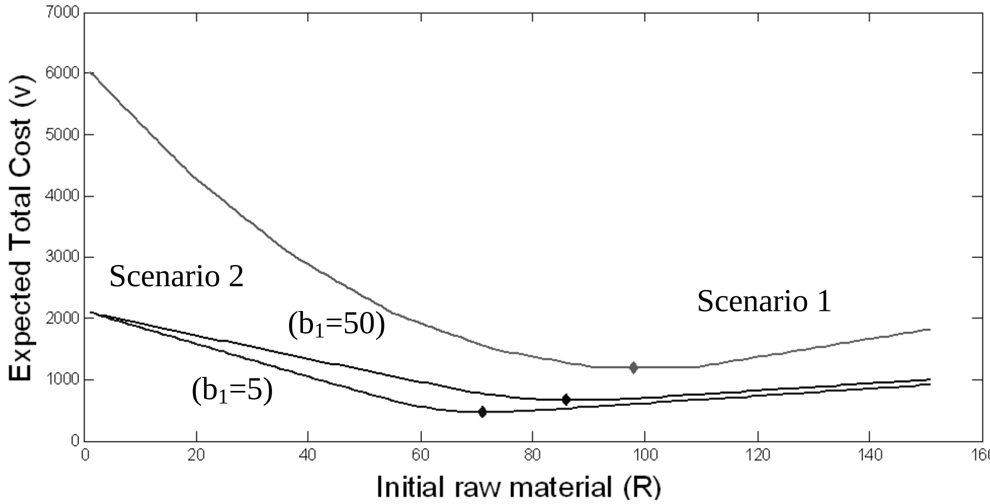


Figure 3-4: Comparison of optimal total expected cost and optimal initial raw material decision (*R*, the diamond marker) as a function of fulfillment scenario and expediting costs.

While it is intuitive that the change in total expected cost, between fulfillment scenarios, decreases as the cost of internal expediting increases, the degree of change can be insightful. The cost gap provides a bound on the cost that a manager should spend to have internal expediting capabilities.

3.6.2 Impact of Waste Costs

Depending on the product, waste costs (due to expiration) may be significant. Some products require special disposal (e.g., chemotherapy drugs), which increases costs. Waste costs are charged for unused raw material at the end of the cycle and the unused finished goods at the end of each period. Keeping all other costs fixed, as the waste cost of finished good increases, w_2 = 10, 20, 30, the production quantity x, of finished good, decreases. By varying the waste cost of the raw material in an increasing fashion, w_1 = 1, 4, 8, the optimal initial raw material level is non-increasing and the expected total cost is non-decreasing. It is intuitive that the echelon waste cost impacts the decision at that echelon. In the following section we describe how the two decisions are interrelated.

3.6.3 Interrelationship Between *R* and *x*

The two decision variables *R* and *x* are linked as the raw material inventory level in any period is a constraint on the production decision. Figure 3-5 demonstrates how a suboptimal choice for *R* impacts the optimal production decision. The crosses plotted on Figure 3-5 indicate the optimal production decision x_1^* given an initial order quantity *R*.

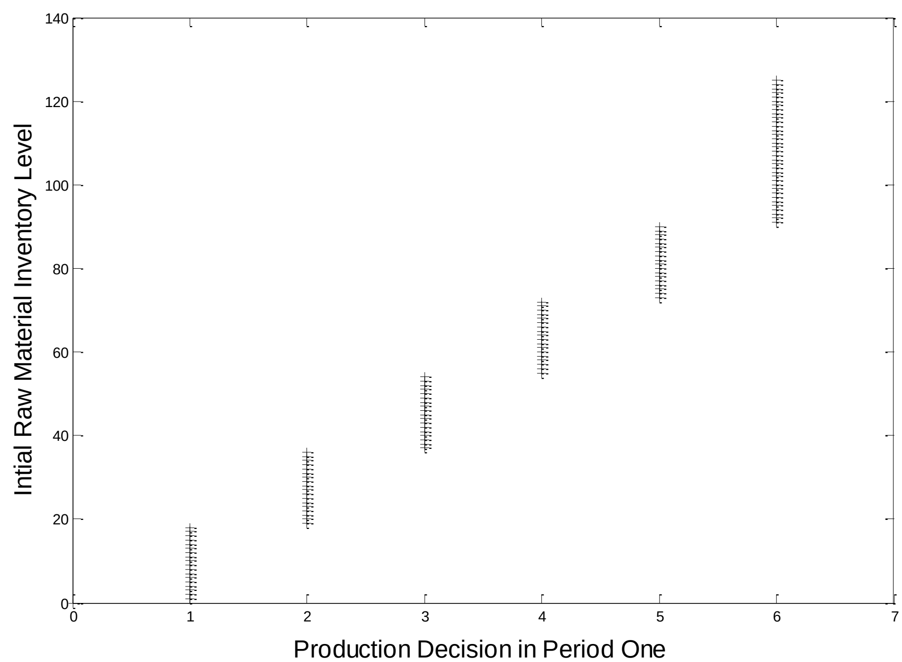


Figure 3-5: Effect of the initial raw material inventory decision on the optimal production decision, denoted by the ‘+’, in period 1

The optimal values for (R, x_1) are (105, 6). As shown in Figure 3-5, there is a relationship between the two decision variables. There are threshold values for R such that if the initial raw material inventory is less than this value, the optimal production decision does not change. For example, for $R < 20$ the production quantity, x_1^*, equals 1 unit, for $19 < R < 37$, x_1^*, equals 2 units, etc. The interdependency between the production decision and the initial raw material order decision is clearly linked. In the following section we discuss our results and conclude with our research contribution.

3.7 Conclusion

We develop a two-stage (i.e., multi-echelon) perishable inventory model. In our model the production decision which converts the raw material (in Stage 1) to the finished good (in Stage 2) significantly increases the perishable nature of the product. We extend the literature by defining a Markovian demand process that represents the uncertainty in patient condition progression which directly impacts the dose required by the patient and hence drug demand. We prove that the optimal production decision in each period is dependent on the state of the Markov demand process and raw material inventory level. Thus, unlike Fujiwara et al. (1997), our optimal production decision is nonstationary.

We consider two different demand fulfillment strategies for responding to a shortage. For the first scenario, shortages must be fulfilled via an external source. In this case the raw material inventory is depleted by the planned production and is unaffected by the quantity expedited. The second scenario extends the model to allow for additional production from internal raw material as well as the option to expedite from the external source, if needed.

The existence of optimal solutions to the initial raw material inventory order and periodic production decision for both scenarios is proven. These optimal solutions are dependent on the fulfillment scenario. The difference in the total expected cost between scenarios can be interpreted as the value of investing in internal expediting capabilities.

The numerical examples illustrate the behavior of the optimal solutions for a series of cost structures and the fulfillment scenarios. As we have demonstrated, the optimal results are influenced by the tradeoff between waste and expediting costs. Therefore accurately capturing these costs for a given drug or drug class is important for ensuring that the output of the model represents the real system.

Appendix B

Recall that L is our single period cost function (Equation 3.3) when the demand is in state z, the raw material inventory is R, and the production decision is x. Now if we assume that the decision R can be continuous but the decision x is discrete, we must prove that there still exists an optimal solution for x. The following Proposition defines the analog to Lemma 1.

Proposition 1: If $\kappa(x)$ is a real-valued function of $x \in X$ where X is a bounded subset of Z, and if $\Delta^2\kappa(x) \geq 0$ for all $x \in Z$, then every local minimum of $\kappa(x)$ is also a global minimum of $\kappa(x)$ on X.

Proof:

In order to prove that the proposition is true, we must develop the first difference equation. We define the first difference as where $\Delta f(x) = f(c + h) - f(c)$ where h is defined as the increment. Rewriting equation (3.3) as its discrete counterpart below, the first difference in developed as a function of the decision variable *x*.

$$
\begin{aligned}
L_1(z_1,R,x_1) &= c_1R + c_2x_1 + w_1(R-x_1) + \sum_{m_1=0}^{x_1} w_2(x_1-m_1)p(m_1) + \sum_{m_1=x_1}^{\infty} b_2(m_1-x_1)p(m_1) \\
f(c+h) &= L_1(z_1,R,c+h) \\
&= c_1R + c_2(c+h) + w_1(R-(c+h)) + \sum_{m_1=0}^{(c+h)} w_2((c+h)-m_1)p(m_1) + \sum_{m_1=(c+h)}^{\infty} b_2(m_1-(c+h))p(m_1) \\
f(c) &= L_1(z_1,R,c) \\
&= c_1R + c_2(c) + w_1(R-c) + \sum_{m_1=0}^{c} w_2(c-m_1)p(m_1) + \sum_{m_1=c}^{\infty} b_2(m_1-c)p(m_1) \\
&= c_1R + c_2(c) + w_1(R-c) + \sum_{m_1=0}^{c} w_2(c-m_1)p(m_1) + \sum_{m_1=c}^{\infty} b_2(m_1-c)p(m_1) \\
f(c+h)-f(c) &= c_2h - w_1h + w_2\sum_{m_1=0}^{c} hp(m_1) + \sum_{m_1=c}^{c+h} w_2(c+h-m_1)p(m_1) - b_2(m_1-c)p(m_1) \\
&\quad -b_2\sum_{m_1=c+h}^{\infty}(h)p(m_1) \\
\frac{f(c+h)-f(c)}{h} &= c_2 - w_1 + w_2\sum_{m_1=0}^{c} p(m_1) - b_2\sum_{m_1=c+h}^{\infty} p(m_1) + \frac{\sum_{m_1=c}^{c+h} w_2(c+h-m_1)p(m_1) - b_2(m_1-c)p(m_1)}{h}
\end{aligned}
$$

The last term can be further simplified as follows.

$$
\begin{aligned}
\frac{\sum_{m_1=c}^{c+h} w_2(c+h-m_1)p(m_1) - b_2(m_1-c)p(m_1)}{h} &= \frac{w_2(c+h-c)p(c) - b_2(c-c)p(m_1) + w_2(c+h-(c+h))p(c) - b_2(c+h-c)p(m_1)}{h} \\
&= \frac{w_2hp(c) - b_2hp(c+h)}{h} = w_2p(c) - b_2p(c+h)
\end{aligned}
$$

$$
\begin{aligned}
\frac{f(c+h)-f(c)}{h} &= c_2 - w_1 + w_2F(c) - b_2\sum_{m_1=c+h}^{\infty} p(m_1) + w_2p(c) - b_2p(c+h) \\
&= c_2 - w_1 + w_2F(c) - b_2(1-F(c+h)) + w_2p(c) - b_2p(c+h) \\
&\text{as } h \to 0 \\
&= c_2 - w_1 + w_2F(c) - b_2(1-F(c)) + w_2p(c) - b_2p(c) \\
&= c_2 - w_1 - b_2 - (w_2+b_2)F(c) + w_2p(c) - b_2p(c)
\end{aligned}
$$

Defining the second difference as $\Delta^2 f(x) \equiv f(c+2h) - 2f(c+h) + f(c)$, we next evaluate the resulting second difference for the single period problem in order to determine if an optimal production decision exists.

$$
\begin{aligned}
f(c+2h) - 2f(c+h) + f(c) = {} & c_1R + c_2(c+2h) + w_1(R-(c+2h)) + \sum_{m_1=0}^{(c+2h)} w_2((c+2h)-m_1)p(m_1) \\
& + \sum_{m_1=(c+2h)}^{\infty} b_2(m_1-(c+2h))p(m_1) - 2c_1R - 2c_2(c+h) - 2w_1(R-(c+h)) \\
& + -2\sum_{m_1=0}^{(c+h)} w_2((c+h)-m_1)p(m_1) + -2\sum_{m_1=(c+h)}^{\infty} b_2(m_1-(c+h))p(m_1) \\
& + c_1R + c_2(c) + w_1(R-c) + \sum_{m_1=0}^{c} w_2((c+h)-m_1)p(m_1) + \sum_{m_1=c}^{\infty} b_2(m_1-c)p(m_1) \\
= {} & w_2hp(c+h) + b_2hp(c+h)
\end{aligned}
$$

$$
\frac{f(c+2h) - 2f(c+h) + f(c)}{h^2} = \frac{w_2p(c+h) + b_2p(c+h)}{h} \geq 0
$$

Thus, Proposition 1 is proven for the single period problem.

The multi-period problem extends the single period total expected cost function by the addition of positive expected cost. It immediately follows that Proposition 1 holds for the *n* period problem and thus the second difference for the multi-period total expected cost function *v* has a positive second difference.

CHAPTER 4:

Stochastic Multi-Echelon Perishable Inventory Models for Meropenem Management

Abstract: In this paper we apply a two stage perishable inventory model to the production and ordering decision of a pharmaceutical drug administered in an inpatient setting. We expand this model to the management of Meropenem drug therapy. Based upon analysis of data from a large public hospital we derive demand characteristics and distributions for total daily dose. The demand distribution for the total daily dose of Meropenem correlates directly to the number of patients in the hospital. We formulate the demand process as a Markov chain where the state is defined as the number of patients in the system. We evaluate the difference in cost performance from the Markovian demand model with that of a model that ignores this relationship. Finally, we conduct sensitivity analysis on cost parameters with levels of uncertainty.

4.1 Introduction

The inventory management of perishable pharmaceuticals within the hospital pharmacy is a complex task. In fact, the impact of mismanagement extends beyond healthcare costs (due to waste and shortages) to patient care and utilization of resources (Baumer et al. 2004). Due primarily to suboptimal inventory systems, hospital pharmacies average a low 10.2 inventory turns per year, lose contract compliance opportunities, and continue costly process inefficiencies (Alverson 2003). In an attempt to improve their drug delivery system, many hospitals have implemented automated systems such as Pyxis machines but have not experienced success due to their labor intensiveness, inefficiencies, and lack of proper inventory management policies (Handfield 2007). The demand for medication is based on the patient population within the hospital which is uncertain and varies through time. Due to this stochastic demand, inventory levels are set artificially high in order to hedge against uncertainty. To add to the complexity, syringes and IVs that are partially used cannot be used on subsequent patients. Because of the high waste costs for certain drugs, preparation of doses in the dispensed form (i.e., syringe or IV) in advance of demand has decreased (Fraind et al. 2002). In a study of IV preparation in European hospitals, many medication errors, such as the use of the wrong diluents, incorrect administration rate, and products that were not mixed, were found (Cousins et al. 2005). Fraind et al. (2002) and Cousins et al. (2005) propose having pharmacy departments or suppliers prepare these IVs in advance in order to mitigate some of the mistakes. The research question becomes: How much to prepare? A more careful analysis of demand patterns reveals opportunities to make inventory and preparation policies which are cost-effective. The documented issues with outdating and shortage can be attenuated by the implementation of a robust inventory management policy, which is driven from a more intelligent forecast based on expected patient demand, and can be reached by the understanding of the relationship between patient condition and demand.

We present an application of a two-stage perishable inventory model for Meropenem drug treatment. A vial of Meropenem is termed the raw material (i.e., Stage 1) and an intravenous (IV) Meropenem (plus diluents) is the finished good (i.e., Stage 2). Meropenem is an ultra-broad spectrum antibiotic, administered intravenously to treat a wide variety of bacterial infections. This drug was chosen because is costly (~$30 per gram) and has been difficult to control, resulting in excess waste costs according to our partner hospital. Another desirable attribute is that its administration is protocol driven: patients are given 0.5 gram, 1 gram or 2 grams at a time, depending on their condition. Table 4-1 presents examples of patient conditions requiring Meropenem and the corresponding standard dose used to treat the condition.

Table 4-1. Clinical indications and corresponding dosing for Meropenem (*FDA approved indications) (Micromedex, 2009)

Indication	**Dosing***
Bacterial meningitis	6 grams/day IV divided every 8 hr
Infection of skin AND/OR subcutaneous tissue, complicated	500 mg IV every 8 hr; diabetic foot: 1 gram every 8 hr
Infectious disease of abdomen, complicated	1 gram IV every 8 hr

Though Table 4-1 shows what appears to be a simple mapping of condition to dose, there are other factors such as serum creatinine level (that indicates renal function), platelet count, and complete blood count which influence the dose and/or the frequency of administration. The length of administration is also stochastic. In the following section we discuss the dataset and data analysis.

4.2 Data Analysis

We partnered with a large urban hospital that has over 350 inpatient beds and more than 14,000 adult admissions per year. Institutional Review Board (IRB) approval was granted for this study. For the purpose of this study we analyze encounter data for patients who received Meropenem during their length of stay (LOS) in 2008. An encounter is defined as each time a patient received the drug. The dataset includes ~8,000 encounters. In this paper, we analyze patients in the progressive intensive care unit (PICU). The demand in the PICU is the largest of all units, representing approximately 17% of total demand. The PICU data includes 1,427 patient encounters for 154 distinct patients. As described in Table 4-1, there are three dose levels. Figure 4-1 shows the proportion of each dose level used in the PICU.

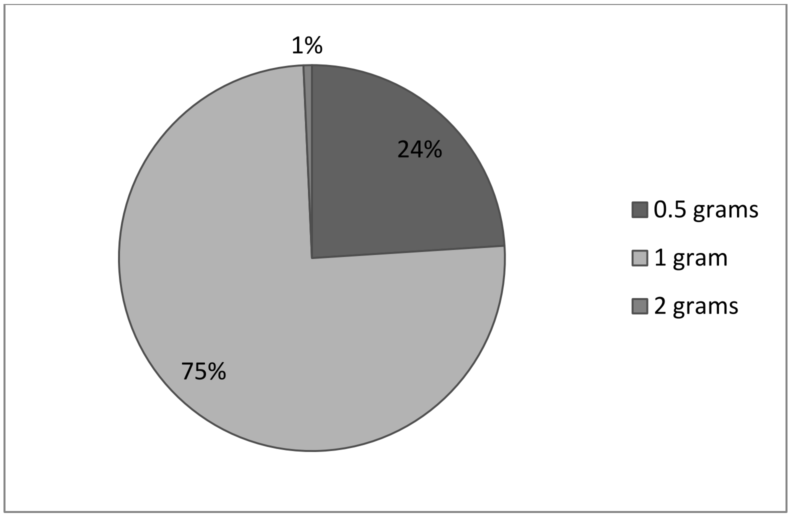


Figure 4-1. Total proportion of dose level used per encounter (with Meropenem use) in the PICU

As is evident by Figure 4-1, the majority of patients receive a 1 gram dose at each administration. However, patients typically do not receive just one administration in their course of treatment and the dose level may change over time. Next we analyze the variability in the length (in days) of Meropenem treatment. Figure 4-2 illustrates the number of patients who receive Meropenem therapy from 1 to 10 days.

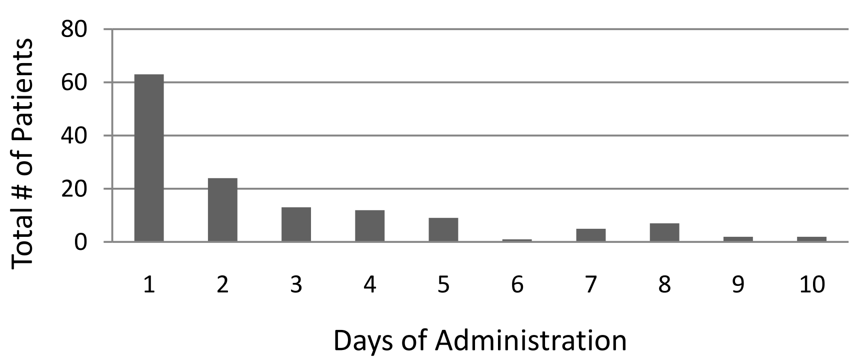


Figure 4-2: Number of patients who are administered the drug over a span of 1 to 10 days

Figure 4-2 suggests the number of days that a patient receives Meropenem is highly variable. Lastly we are interested in determining whether a correlation between the number of patients in the system and total dose exists. Plotting the total dose (in grams) and number of patients in the PICU for each service date (in Figure 4-3) reveals a relationship between the two.

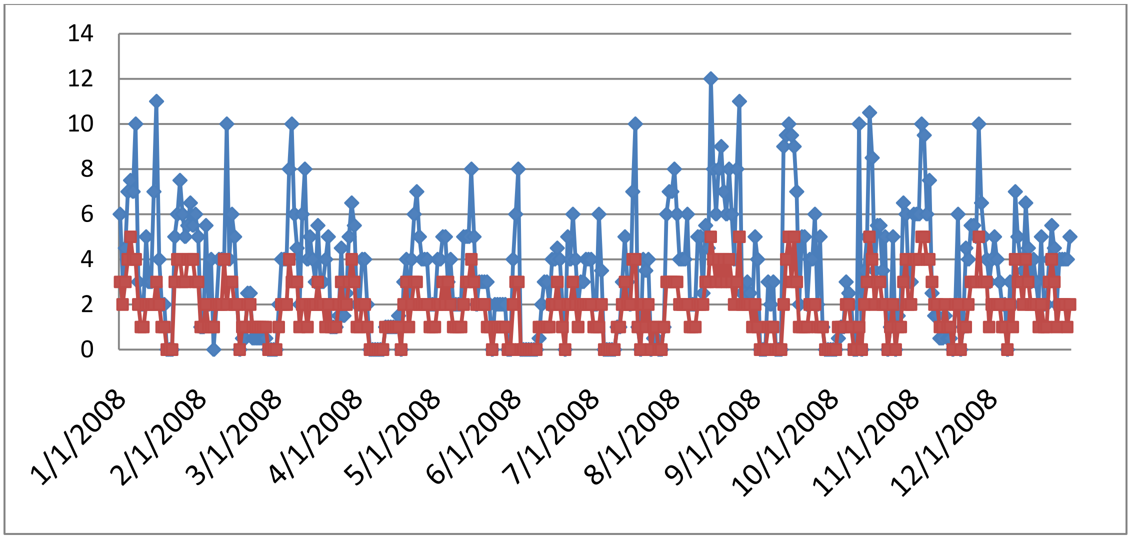


Figure 4-3. Total dose of Meropenem administered (blue) and total number of patients in the system per day (red)

Figure 4-3 shows that the total dose per day (total administered to the patient population) ranges from 0 grams to 12 grams and that the total number of patients ranges from 0 to 5 per day. A simple regression using the number of patients to predict total dose results in an R^2 of .75 and p value of approximately zero. Therefore, we conclude that the number of patients in the system is a significant predictor of the total dose demand. Each day new patients may arrive to the PICU, current patients may depart or continue to stay in the hospital. This change in the number of patients is the motivation for the development of a

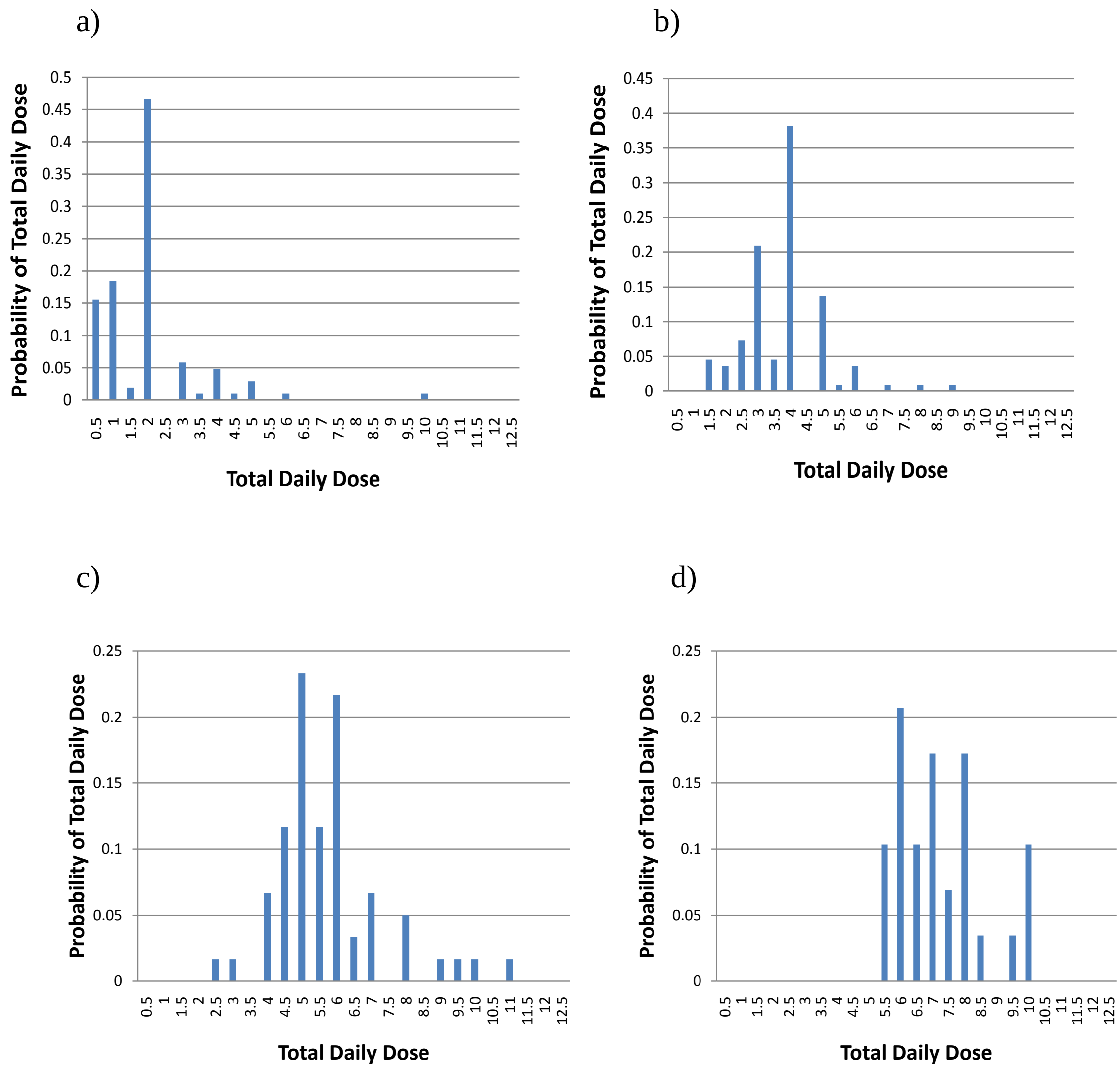
Markov chain which represents the changes in patient population and consequently changes in the distribution for total patient demand.

The subset of data analysis presented in Figures 4-1 to 4-3 show the variability in the demand process. In Section 4.3 we present the analysis supporting the development of a time-homogenous Markov chain which represents the changes in patient demand. The two-stage perishable inventory problem is defined as a Markov decision problem (MDP) in Section 4.4. Section 4.5 compares optimal and sub-optimal heuristic inventory and preparation policies. Finally, in Section 4.6 we discuss the results and future research.

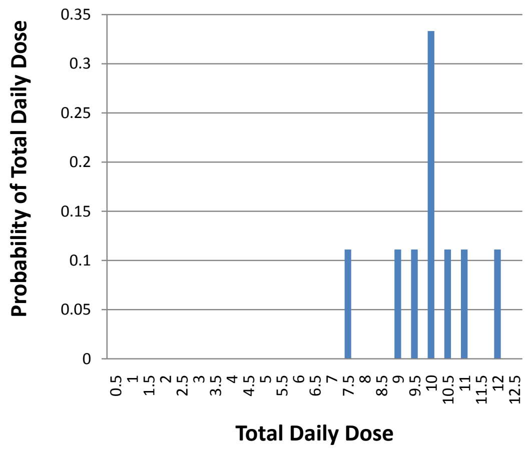
4.3 Specification of Demand Process

For this study, we use the patient care unit (PCU) as a surrogate for patient condition. Specifically, we model the distribution of patients in the PICU that receive Meropenem. As found in Section 4.2 demand for the drug is highly dependent on the number of patients in the system. We develop a Markovian demand process that represents the dynamics of the change in the number of patients and their link to demand for a single drug. Figures 4-4a to 4-4e show the empirical discrete demand distributions as a function of the number of patients in the PCIU, which can be 1, 2, 3, 4, or 5 patients in the system.

Figure 4-4a-e: Empirical probability distributions for total daily dose given there are 1,2,3,4,or 5 (figures a –e, respectively) candidate patients in the PICU


a)
b)
c)
d)
Probability of Total Daily Dose
Total Daily Dose

e)



We assume if there are no patients then demand is zero with certainty. Figure 4-4a-e show that the coefficient of variability changes depending on the number of patients in the system. Table 4-2 show summary statistics according to the number of patients in the system. The coefficient of variability (CV) decreases as the number of patients increase due to an averaging effect of demand. The 95% confidence intervals show no overlap between the patient groups and thus the difference in demand distributions are statistically different at the 0.05 level.

Table 4-2: Summary statistics for total daily dose using the empirically derived demand distributions which correspond to the number of patients in the system.

# of Patients	1	2	3	4	5
Mean	1.971	3.805	5.733	7.241	9.944
St. Dev	1.389	1.244	1.558	1.373	1.261
CV	0.705	0.327	0.272	0.190	0.127
Confidence Level(95.0%)	0.272	0.235	0.403	0.522	0.969

We define the state, z, for the patient Markov chain as the number of Meropenem candidate patients in the PICU (e.g., state 0 represents zero patients, state 1 represents one patient, etc….). There is a transition matrix Q where q_{ij}= P(j patients are present on day t| i patients are present on day t-1). For the purpose of this analysis we assume that an epoch is one day (24 hours). The empirically derived transition probabilities are shown in Table 4-3.

Table 4-3. Transition probabilities for number of patients in the system

state	**0**	**1**	**2**	**3**	**4**	**5**
0	0.574	0.259	0.148	0.019	0.000	0.000
1	0.155	0.524	0.282	0.029	0.010	0.000
2	0.063	0.243	0.460	0.180	0.054	0.000
3	0.017	0.117	0.300	0.316	0.167	0.083
4	0.000	0.034	0.138	0.448	0.277	0.103
5	0.000	0.000	0.111	0.333	0.444	0.112

We observe that the transition probabilities indicate that it is unlikely that many patients would be discharged from one day to the next (i.e., $p_{30} < p_{31}$). Similarly, since this drug is used for complex patients, it is unlikely that the patient population would jump from zero patients to five patients in one day. This environment lends itself to the Markovian assumption which dictates that knowledge of the present state is the only information needed to predict the future state. Given the derived demand distributions and Markovian demand process, we use these as inputs for the model we specify in the next section.

4.4 Model Formulation

In this section we use the Markov decision process formulated in Vila-Parrish et al. (2009) as the framework for our analysis. In this model, there are two decisions: (1) an initial raw material order quantity, R and (2) a periodic finished good production quantity, x. Each period the pharmacist reviews the Stage 1 inventory level of the raw material (i_t^r) and the state of the demand process (z) and plans the production of Stage 2 finished good (x). Next, we briefly define the costs parameters used in this model.

First, there is a holding cost of h_1 for raw material inventory per unit per period. There is a cost associated with producing in advance of demand, c_2, for the finished good. We assume that internal and external expediting are viable alternatives to fulfill demand shortages. Expediting costs are, b_1 and/or b_2, depending on the expediting method utilized, internal and external expediting, respectively. We assume there is no backlogging allowed since the patient's health requirements must be met. We define external expediting as a method to fulfill demand in excess of Stage 2 inventory by which excess demand is fulfilled from either a supplier or another hospital for an additional fee. Internal expediting refers to finished good production that occurs within the hospital but after demand. Lastly, there is a waste cost, w_1 (w_2) associated with expired Stage 1 (Stage 2) material. Equation (4.1) is composed of the production cost, and the expected holding, waste, and internal and external expediting costs given the production decision is x_t. Equation (4.2a) shows the total expected cost with n periods remaining in the horizon. This cost is waited by a factor of α_j which is the initial probability that the patient demand process is in state j. Equation (4.2b) is the recursive cost function for this problem and Equation (4.2c) is the waste cost incurred for raw material remaining at the end of the cycle.

$$g_t(z_t, i_t^r, x_t) = \left\{ \begin{array}{l} c_2 x_t + h_1(i_t^r - x_t - \int_{x_t}^{i_t^r} (m - x_t) f_{z_t}(m) dm) \\ + \int_0^{x_t} w_2 (x_t - m) f_{z_t}(m) dm + \int_{x_t}^{i_t^r} b_1 (m - x_t) f_{z_t}(m) dm \\ + \int_{i_t^r}^{\infty} b_2 (m - x_t) f_{z_t}(m) dm \end{array} \right\} \tag{4.1}$$

$$\begin{array}{l} (a)\ v_n(\cdot, R) = \sum_{z_{n-1}=1}^{K} \alpha_{z_{n-1}} v_{n-1}(z_{n-1}, R) \\ (b)\ v_t(z_t, i_t^r) = \min_{0 \le x_t \le i_t^r} \{ g_t(z_t, i_t^r) + \sum_{z_{t-1}=1}^{K} p_{z_t k_{t-1}} E[v_{t-1}(k_{t-1}, i_{t-1}^r = i_t^r - x_t - a_t)] \} \\ t = 1, \ldots, n-1 \\ (c)\ v_0(z_0, i_0^r) = w_1 * i_0^r \end{array} \tag{4.2}$$

We model the problem as a single cycle, of length T, with the cycle length defined by the shelf life of the raw material. The cycle is subdivided into n periods which represent the frequency of the production decisions during the cycle and the shelf life of the finished good. Depending on the diluents used and the storage conditions, the shelf life of the IV solution varies from 1 hour to 24 hours. For the purpose of the analysis, we assume that the shelf life is 24 hours (e.g., Sodium Chloride Injection 0.9% at 39 degrees F) (Micromedex, 2009). Next, Section 4.5 we present numerical analysis of the production and raw material inventory decisions and the impact of cost assumptions on these decisions.

4.5 Model Preliminaries

Using the model described in Section 4.4 and the parameters derived in Section 4.2, we evaluate an 11 period problem with 10 production decision epochs. Table 4-4 shows the baseline parameter values used in the numerical example. The waste cost of the finished good and of the raw material are derived from the real costs incurred at our partner institution. The external expediting costs are composed of a cost premium and the management cost of ordering and receiving the product. There is also uncertainty in the cost of the internal expediting. We conduct sensitivity analysis to evaluate the differences in the optimal Stage 2 production decisions as a result of changes in expediting costs. All other parameters (i.e., transition rates and demand functions) are equivalent to those defined in Tables 4-1 and 4-3.

Table 4-4. Baseline cost parameters (in dollars)

Parameter Name	c_2	h_1	w_1	w_2	b_1	b_2	**N**
Baseline Value	1	0.1	30	32	5	50	11

We evaluate the gap in cost performance between the full-scale, dynamic problem and a simpler model which ignores the fluctuations in the number of patients. In the following sections, through examples, we evaluate the implications of assuming that demand is Markovian versus stationary and discuss the resulting optimal decisions.

4.5.1 Dynamic Demand Process

In this section we present the production and ordering results from the Markovian demand process model. The sequence of optimal raw material ordering and production decisions results in an optimal dynamic production policy. Using the baseline costs, the optimal initial raw material ordering quantity is 35 grams and the total expected cost is $734.17. Table 4-5 shows various feasible sample paths of the system state space (i.e., raw material inventory level and patient demand process) from period t to t+1and the corresponding optimal production decision. In order to show the raw material inventory transition, we assume that the realized total demand is 4.5 grams.

Table 4-5. Example of a optimal production decision x^* given the initial state is (z=2, R=35) assuming the realized demand is 4.5 grams

n periods remaining					*n*-1 periods remaining		
A	B	C	E	F	G	H	I
# of Patients (z)	Raw Material Inventory ($R=r_n$)	Optimal Production (x)	Demand	P(Demand)	# of Patients (z)	Raw Material Inventory (r_{n-1})	Optimal Production (x)
2	35	x_n = 2.5 grams	0	0%	0	30.5	x_{n-1} = 0 grams
			0.5	0%	1	30.5	x_{n-1} = 2 grams
			1	4.55%	2	30.5	x_{n-1} = 2.5 grams
			1.5	3.64%	3	30.5	x_{n-1} = 4 grams
			2	7.27%	4	30.5	x_{n-1} = 5.5 grams
			2.5	20.91%	5	30.5	x_{n-1} = 7.5 grams
			3	4.55%			
			3.5	38.18%			
			4	0%			
			$\boxed{4.5}$	13.64%			
			5	0.91%			
			5.5	3.64%			
			6	0.00%			
			6.5	0.91%			
			7	0%			
			7.5	0.91%			
			8	0%			
			8.5	0.91%			

As these results in Table 4-5 show, the optimal finished good production policy is nonlinear with respect to the number of patients in the system. This snapshot illustrates the necessary parameters for the practitioner to make an informed production decision, namely the raw good inventory (r_n) and the number of patients (z_n). Next we walk the reader through a feasible path assuming a demand of 4.5 grams is realized in period *t*. Given the starting state with *n* periods remaining (z_n=2, R=r_n=35), it is optimal to produce 2.5 grams of finished goods (column C). If a demand of 4.5 grams (Column D) is realized this results in 2 additional grams of finished goods produced via internal expediting. With *n*-1 periods remaining, the raw material inventory level is 30.5 grams (Column H) and the number of patients in the system is updated (Column G). Depending on the number of patients with *n*-1 periods remaining the optimal production decision is shown in Column I of Table 4-5. The results are dynamic in the sense that they utilize the patient demand process, represented here by the number of patients in the system, and the raw material inventory level to determine the optimal finished good production quantity. Also, the raw material inventory decreases from period to period as a result of production decision and expedited production.

As previously mentioned, total cost may have components which are known with certainty and others which have elements of uncertainty. In the following section we conduct sensitivity analysis on the internal and external expediting costs.

4.5.2 Impact of Expediting Costs

External expediting costs are composed of material costs and administrative costs. The expedited product cost includes an additional cost premium for the product (e.g., 10%). If a shortage is experienced there is an administrative or management cost to find a suitable source for fulfilling demand. Though a subset of the external expediting cost is known, the total cost may be difficult to quantify thus we vary the total cost of external expediting from 40 to 60. Recall that we assume that $b_1 < b_2$ in order for convexity of the total expected cost function (Vila-Parrish et.al, 2009) and therefore the cost of internal expediting ranges from 2 to 40. The following table summarizes the first period production quantity and raw material order quantity for each internal and external cost combination and the resulting total expected cost.

Table 4-6. Results of sensitivity analysis on the cost of internal and external expediting, the values noted are the production and raw material quantity (in grams).

Internal Expediting Costs	External Expediting Costs	R	x(0,R)	x(1,R)	x(2,R)	x(3,R)	x(4,R)	x(5,R)	Total Expected Cost
2	40	33		1	1.5	2.5	5.5	7.5	640.43
5	40	33.5		2	2.5	4	5.5	7.5	663.96
10	40	34		2	3	4.5	6	9	693.34
20	40	34		3	3	5	6	9	737.82
30	40	35.5		3.5	3.5	5	6	9.5	771.62
2	50	33		1	1.5	2.5	5.5	7.5	711.26
5	50	35		2	2.5	4	5.5	7.5	734.17
10	50	35.5		2	3	4.5	5.5	7.5	763.19
20	50	36		3	3	4.5	6	9	806.36
30	50	37.5		4	4	5	6	9.5	838.07
40	50	38		4	4	5	6.5	11	864.14
2	60	36		1	2	2.5	5.5	7.5	775.84
5	60	36.5		2	3	4	5.5	7.5	798.38
10	60	37		2	3	4	5.5	7.5	828.65
20	60	38		3.5	3.5	4.5	6	9	867.82
30	60	38.5		4	4	5	6	9.5	898.33
40	60	39		4	4	5	6.5	9.5	924.36

Table 4-6 shows that as the internal expediting cost increases it is optimal to produce more finished good in advance of demand. When the value of b_1 is 20 or more, the optimal first period production quantity increases, from the baseline decision, for each patient state. The optimal raw material order quantity increases as the production decisions increase. Lastly, the total expected cost increases as expediting costs (of either type) increase (see last column).

The optimal finished good production quantities are dependent on the state of the patient demand process and the raw material inventory level (i.e., the state of the system). The results of the sensitivity analysis show that the decisions x and R are dependent on the expediting cost assumptions. Therefore, understanding the true cost (or cost range) of expediting costs is imperative to making cost effective decisions. In the next section we present a simpler model which ignores fluctuations in the number of patients in the system. We are interested in the cost improvement achieved by the more complex model.

4.5.3 Cost Impact of Ignoring Patient Population Dynamics

In contrast to our dynamic model which incorporates the nonstationary patient demand process, we now study the impact of utilizing a stationary stochastic demand function (i.e., ignoring fluctuations in the number of patients). We empirically derive a discrete demand distribution function which aggregates all patients, shown in Figure 4-5, using the 2008 dataset.

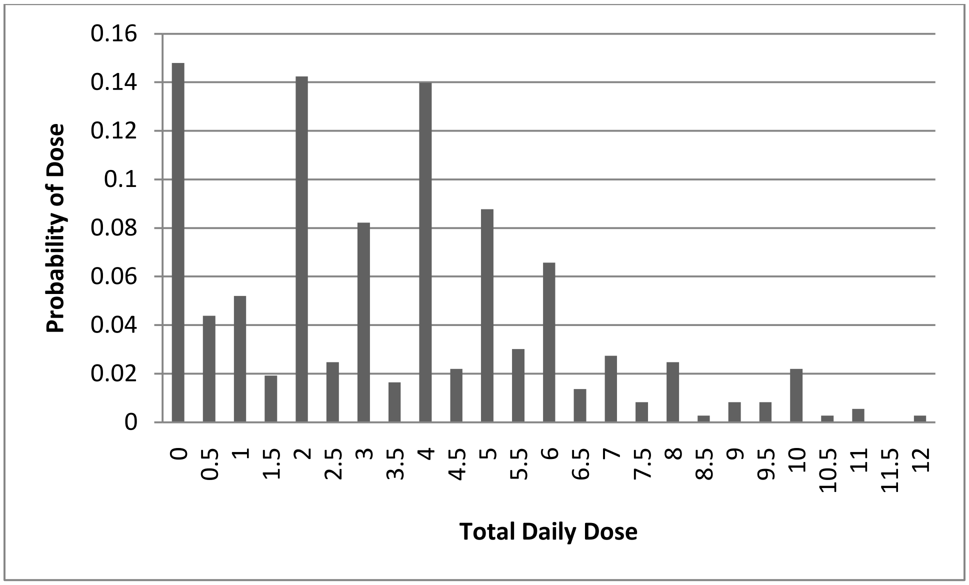


Figure 4-5: Probability mass function for total daily dose in grams, ignoring the number of actual patients in the system

The optimal production policy is independent of the number of patients in the system, and the advanced production decision is still constrained by the raw material inventory level in that period. We evaluate the performance of the production and raw material order policies derived from this stationary demand distribution when applied to the true nonstationary patient demand process. We measure the difference in cost assuming that the patient demand process is accurately reflected by our dynamic model. For the baseline cost parameters there was an increase of 7.08% from the dynamic model's total cost. Table 4-7 summarizes the percent change in cost between these two models.

Table 4-7. Percent change in total expected cost of the dynamic model evaluated with the policies derived by assuming that the demand model is stationary.

b_1	2	**5**	10	20	30	40
% Change	2.59%	**7.08%**	10.86%	15.08%	17.23%	18.12%

The results in Table 4-7 show that the optimal decisions derived from dynamic model outperform the stationary policies. The difference in performance increases as the cost of internal expediting increases. This is due to the fact that the production policies developed from the stationary demand process tend to underutilize the advanced production opportunity. Therefore expediting is needed more often and total expected costs increase. The variability of the total daily dose and high correlation between the number of patients in the system and total daily demand influences the superiority of the solutions from the dynamic model. The comparison of the cost performance between these two models further emphasizes the importance of cost parameter values (i.e., the cost difference varies dependent on the cost parameter values).

4.6 Discussion

Given historical patient information such as the number of patients in the system and Meropenem drug dose administered by patient, we create a patient demand process which is dependent on the number of patients in the system. We represent this process as a Markov chain and empirically derive transition rates which represent how the number of patients changes from day to day. The number of patients in the system corresponds directly to a discrete demand distribution. Given these model inputs we evaluate two demand processes.

In the dynamic model, we use a patient-based demand process, derived from historical data, to determine optimal production and raw material order quantities. The finished good production quantity is dependent on the system state: current state of the demand process and the raw material inventory level. By conducting sensitivity analysis on expediting costs which are the most uncertain, we find that an increase in internal expediting cost corresponds to an increase in advance dose production. We then compare the cost performance of a stationary aggregate demand model. We derive the optimal production and raw material order quantities using this model and then evaluate the impact of ignoring the Markovian process when the true demand process is nonstationary. While the total cost of optimal decisions of the dynamic model is less, the difference in total cost is dependent on the assumed expediting cost.

Understanding the true system costs are essential in making cost effective production and inventory decisions. Despite this, our model shows that there is an opportunity to produce some quantity of finished goods in advance of demand. The ability to do so may improve pharmacy efficiency and accuracy in preparing doses. The patient safety issues described in section 4.1 may also be addressed by this advanced preparation strategy. This modeling framework can be applied to other hospital units and the definition of the states of the patient demand process can be expanded to include more patient information (such as ICD-9 codes).

CHAPTER 5:

Research Contribution and Future Work

5.1 Research Contribution of Chapter 2

Our research extends the literature on two modes of supply by deriving inventory and ordering policies under nonstationary demand. We represent the demand process of our fluctuating environment by using a Markov modulated demand structure which is a more flexible model that allows for dependency in demand evolution (Song and Zipkin, 1993, 1996). Our research enriches the literature on Markovian demand processes by considering the entire product lifecycle (PLC) that is not only inclusive of the obsolescence process but also of the product ramp process. This end to end view of the lifecycle extends the knowledge of the inventory control policies which depend on the state of the product lifecycle.

We specify the demand process to correspond to the characteristics of demand during each phase of the PLC and consider two types of demand evolution. The problem is formulated as an MDP which allows us to model the dynamics of the inventory and demand model. Two cases are defined: Case 1 is the general model which models demand evolution as an embedded Markov chain, and second is a special case (i.e., Case 2) where demand evolution is time-dependent. By defining these two cases, we evaluate different product environments with different levels of PLC demand uncertainty and progression. Further, we prove structural properties of these policies that give insight into the interrelationship between shipment quantities per method and the (1) Demand characteristics and (2) Demand evolution. The structure of the optimal policy, defined by two base stock levels, is nonstationary (i.e., state- and time-dependent) and dependent on the type of demand evolution. While past research has focused on stationary demand, our results show that it is important to consider the type of demand evolution as well as the demand variability. The implications of these results are that different strategies (i.e., the split of total order quantities per mode of supply) should change throughout and according to the product lifecycle evolution.

This analysis can be extended in many ways. First, this analysis can be extended to include time-dependent costs. Specifically, holding and backorder costs may vary throughout the PLC. These nonstationary costs may further influence how the PLC impacts ordering and inventory decisions. Second, we can relax the assumption that the fast mode has zero lead time and find conditions that result in base stock policies. It would be interesting to characterize the proportion of order quantities via each method of shipment as a function of lead time differences between the modes. Lastly, a multi-product environment where there is dependency between the demand processes of each product would enrich the realism of the current model.

5.2 Research Contribution of Chapters 3 and 4

In Chapter 3, we extend the multi-echelon perishable inventory literature by formulating a two-stage perishable inventory system where production activities affect the remaining shelf life of the product. The production decision which converts the raw material (in Stage 1) to the finished good (in Stage 2) significantly increases the perishable nature of the product. We define a stochastic demand process that is a property of the system instead of an exogenously defined random variable. This demand function represents the link between the patient population and their demand. We formulate the problem as a Markov decision process (MDP) and define two shortage fulfillment scenarios relevant for pharmaceutical drug management.

We prove the existence of an optimal raw material order quantity and finished good production quantities. We extend this proof to consider discrete production quantities. The solutions to the optimal production and inventory decisions are dependent on the shortage fulfillment method. The difference in total expected cost for the two fulfillment options gives insights on the value of acquiring the capability to internally expedite. We present numerical examples that show the sensitivity of the solutions to various cost structures.

In Chapter 4 we apply the two-stage perishable inventory model developed in Chapter 3 to a real world example. Based upon analysis of data from a large public hospital we derive demand characteristics and distributions for total daily dose. We develop a Markovian demand process that correlates to the demand distributions. We conduct sensitivity analysis to gain insights on the impact of various cost structures to the optimal solutions.

Different ordering and production environments would extend the research completed to date. For example, future work could involve incorporating a positive lead time for the drug production. There are drugs which must undergo a quarantining period and thus production time is long (e.g., two weeks). Hospital pharmacies may also be interested in alternative ordering models where they can order raw material daily at a cost premium versus weekly. A fixed cost of ordering may be appropriate in that case. Given the patient dataset, our data analysis can be extended. We can expand the current analysis to more patient care units in order to learn the differences in usage between units. For some of the units, the use of the decisions from simpler model defined in Chapter 4 may yield acceptable results. We can enlarge the state definitions to include patient diagnosis information and quantify the value of this additional information. Lastly, we can explore the interaction of demand between units if patients move between them.

Bibliography

Alverson, C. (2003-11-01). Beyond purchasing--Managing hospital inventory. Managed Healthcare Executive. Retrieved on 2008-01-01. http://www.managedhealthcareexecutive.com/mhe/article/articleDetail.jsp?id=75802

Baumer, AM., Clark, AM., Witmer, DR., Geize, SB., Vermeulen, LC., and Deffenbaugh, JH. 2004. National survey of the impact of drug shortages in acute care hospitals. *American Journal of Health Systems Pharmacists* 61(19):2015-22.

Beyer, D., and Ward, J. 2001. Network server supply chain at HP: A case study. J. Song and D. Yao, eds. *Supply Chain Structures*: *Coordination, Information and Optimization*, Kluwer Academic Publishers, Boston, MA, Chapter 8.

Brodheim, E., Derman, C. and Prastacos, G.P. 1975. On the Evaluation of a Class of Inventory Policies for Perishable Products Such as Whole Blood. *Management Science,* 21(11): 1320-1325.

Chazan, D., and Gal, S. 1977. A markovian model for a perishable product inventory. *Management Science*, 23(5): 512-21.

Chiang, C., and Gutierrez, G.J. 1998. Optimal control policies for a periodic review inventory system with emergency orders. *Naval Research Logistics*, 45(2): 187-204.

Chiang, C. 2003. A periodic review inventory system with two supply modes. *European Journal of Operational Research*, 149(1): 527-547.

Cohen M.A. 1976. Analysis of single critical number ordering policies for perishable inventories. *Operations Research* 24(4):726-741.

Cohen, M.A., Pierskalla, W.P., and Yen, H. 1981. Analysis of ordering and allocation policies for multi-echelon, age-differentiated inventory systems. *TIMS Studies in the Management Sciences*, 16:353-378.

Cousins, D.H., Sabatier, B., Begue, D., Schmitt, C. and Hoppe-Tichy, T. 2005. Medication errors in intravenous drug preparation and administration: a multicentre audit in the UK, Germany, and France. *Quality and Safety in Health Care*. 14(3):190-195

Dave, U. 1979. On a discrete-in-time order-level inventory model for deteriorating items. *Journal of Operations Research Society* 30(4):349-54.

Feng, Q., Sethi, S., Yan, H., and Zhang, H. 2006. Are base stock policies optimal in inventory problems with multiple delivery modes? *Operations Research*, 54(4): 801-807.

Fox, E.R. and Tyler, L.S. 2003. Managing drug shortages: Seven years' experience at one health system. *American Journal of Health Systems Pharmacists* 60(3):245-53.

Fraind, D.B., Slagle, J.M., Tubbesing, V.A., Hughes, S.A., and Weinger, M.B. 2002. Reengineering Intravenous Drug and Fluid Administration Processes in the Operating Room: Step One: Task Analysis of Existing Processes. *Anesthesiology*: 97(1): 139-147.

Fujiwara,O., Soewandi, H., and Dayani, S. 1997. An optimal ordering and issuing policy for a two-stage inventory system for perishable products, *European Journal of Operational Research*, 99(2):412-424

Fukuda, Y. 1964. Optimal policies for the inventory problem with negotiable lead time. *Management Science*, 10(4): 690-708.

Gallego, G. and Hu, H. 2004. Optimal policies for production/inventory systems with finite capacity and Markov-modulated demand and supply processes. In Stochastic Models of Production-Inventory Systems, G. Liberopoulos et al. (eds.), *Annals of Operations Research*,126: 21-41.

Gillerman, R.G., and Browning, R.A. 2000. Drug use inefficiency: A hidden source of wasted health care dollars. *Anesthesia Analgesia*, 91(4):921-4.

Goh, C., Greenberg, B., and Matsuo, H. 1993. Two-stage Perishable Inventory Models. *Management Science*, 39(5): 633-649.

Goyal, S.K., and Giri, B.C. 2001. Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*. 134(1):1-6.

Graves, S.C. 1982. The application of queueing theory to continuous perishable inventory systems. Management Science 28(4):400-6. Goyal, S.K., and B.C. Giri. 2001. Recent trends in modeling of deteriorating inventory. *European Journal of Operational Researc*h. 134(1):1-16.

Haijema, R., vanderWal, J., and vanDijk, N. 2007. Blood platelet production: Optimization by dynamic programming and simulation. *Computers & Operations Research*, 34(3):760-79.

Handfield, R., Working Paper, Supply Chain Resource Cooperative, College of Management, North Carolina State University, 2007

Kapuscinski, R., Zhang, R., Carbenneau, P., Moore, R., and Reeves, B. 2004. Inventory decisions in Dell's supply chain. *Interfaces*. 34(3): 191-205.

Ketzenberg, M.E. and Ferguson, M. 2008. Managing Slow-Moving Perishables in the Grocery Industry. *Production and Operations Management*, 17(5): 513-521

Kurawarwala, A. A., and Matsuo, H. 1996. Forecasting and inventory management of short life-cycle products. *Operations Research*, 44(1): 131-150.

Lee, H.L. 2002. Aligning supply chain strategies with product uncertainties. *California Management Review*, 44(3): 105-119.

Lian, Z. and Liu, L. 1999b. A discrete-time model for perishable inventory systems. *Annals of Operations Research*, 87:103-16.

Liu, L. and Lian, Z. 1999a. (s, S) continuous review models for products with fixed lifetimes. *Operations Research*, 47(1):150-8.

Lystad, E., Ferguson, M., and Alexopoulos, C. 2006. Single stage heuristic for perishable inventory control in two-echelon supply chains. Unpublished Ph.D. dissertation, Georgia Institute of Technology, USA

Meropenem. DrugPoint Summary. Thomson Micromedex. Greenwood Village, Co. http://thomsonhc.com Accessed November 2, 2009.

Minner, S. 2003. Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics*, 81-82(1): 265-279.

Moizandeh, K. and Nahmias, S. 1988. A continuous review model for an inventory system with two supply modes. *Management Science*. 34(6):761-773.

Nahmias, S. 1975. Optimal ordering policies for perishable inventory--II. *Operations Research*, 23(4):735.

Nahmias, S. and Pierskalla W.P. 1976. A two-product Perishable/Nonperishable inventory problem. *SIAM Journal of Applied Math*, 30(3):483-500.

Nahmias, S. 1982. Perishable inventory theory: A review. *Operations Research*, 30(4):680-708.

Papachristos, S., and Katsaros, A. 2008. A periodic-review inventory model in a fluctuating environment. *IIE Transactions*, 40(3): 356-366.

Prastacos, G. 1981. Allocation of a Perishable Product Inventory. *Operations Research*, *29*(1): 95-107.

Puterman, M.L. 1994. *Markov Decision Processes.* ***.*** *Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Mathematical Statistics, New York.

Rosenshine, M. and Obee, D. 1976. Analysis of a standing order inventory system with emergency orders. *Operations Research*, 24(6): 1143-1155.

Schmidt C.P and Nahmias S. 1985. (S - 1,s) policies for perishable inventory. *Management Science*, 31(6):719-28.

Sethi, S.P., Yan, H. and Zhang, H. 2001. Peeling the layers of the onion: Inventory model with multiple delivery modes and forecast updates. *Journal of Optimization Theory and Applications*, 108(2): 253-281.

Sethi, S.P., Yan, H. and Zhang, H. 2003. Inventory models with fixed costs, forecast updates, and two delivery modes. *Operations Research*, 51(2): 321-328.

Song, J., and Zipkin, P. 1993. Inventory control in a fluctuating demand environment. *Operations Research*, 41(2): 351-370.

Song, J., and Zipkin, P. 1996. Managing inventory with the prospect of obsolescence. *Operations Research*, 44(1): 215-222.

Tagaras, G. and Vlachos, D. 2001. A periodic review inventory system with emergency replenishments. *Management Science*, 47(3): 415-429.

Telford, S.J. 2008. Discrete Event Simulation in C#. Master's Project Report, Integrated Manufacturing Systems, N.C. State University.

Treharne, J.T. and Sox C.R. 2002. Adaptive inventory control for nonstationary demand and partial information. *Management Science*, 48(5): 607-24.

Veeraraghavan, S., and Scheller-Wolf, A. 2008. Now or later: Dual index policies for capacitated dual sourcing systems. *Operations Research*, 56(4): 850-864

Veinott, A. F., Jr. 1960. Optimal Ordering , Issuing, and Disposal of Inventory with Known Demand, unpublished Ph.D. dissertation, Columbia University.

Vila-Parrish, A., Ivy, J. and King, R. 2008. A Simulation-Based Approach for Inventory Modeling of Perishable Pharmaceuticals. *Proceedings of the 2008 Winter Simulation Conference*.

Vila-Parrish, A., Ivy, J. and King, R. 2009. A Two-Stage Inventory and Production Model for Perishable Products with Markovian Demand and Various Demand Fulfillment Scenarios. *Working Paper*

Weinger, MB. 2001. Drug wastage contributes significantly to the cost of routine anesthesia care. J Clin Anesth 13(7):491-7.

Weiss, HJ. 1980. Optimal ordering policies for continuous review perishable inventory models. *Operations Research*, 28(2):365-74.

Whittemore, AS. and Saunders, SC. 1977. Optimal inventory under stochastic demand with two supply options. *SIAM Journal of Applied Math*, 32(2): 293-305.