ABSTRACT

LAVIN, JAMES ALLEN. Periodic Review Perturbed Demand Inventory Models with Stochastic Demand. (Under the direction of Russell E. King and Donald P. Warsing Jr.).

Perturbed demand models penalize retailers for stock-outs by altering future demand, opposed to the classic method of penalizing stock-outs by imposing a cost proportional to the number of stock-outs. In this paper we consider a periodic-review, perturbed-demand inventory model with stochastic demand. Current consumer willingness to purchase the product after stock-out occurs is captured by a state variable referred to as demand deflation. First we consider the infinite horizon case with instantaneous delivery where we show that the optimal policy converges in steady state and has a base-stock structure that is dependent on demand deflation. Next, we extend the model to include positive reorder lead times and show the optimal policy converges but does not take on an easy to characterize policy. Lastly we consider the finite horizon version of this problem. Due to the computational complexity of this problem we develop heuristics which are shown to have small deviations from optimal gain.

Each perturbed demand problem variation can be formulated and solved as a Markov decision process (MDP) to determine the optimal policy and expected gain. For perturbed demand models, the optimal policy takes on a more complex structure than typical penalty cost models because the order quantity is dependent on current inventory and demand deflation. In traditional penalty cost models there is an optimal policy corresponding to a specific penalty cost value, but the value of the penalty cost can be difficult to determine. Our perturbed demand model creates a common reality in which these penalty cost policies can be tested. The penalty cost policy resulting in the best gain within our perturbed demand reality is one way to estimate the value of the penalty cost which most accurately represents customer behavior.

Periodic Review Perturbed Demand Inventory Models with Stochastic Demand

by
James A Lavin

A dissertation 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

2012

APPROVED BY:

Dr. Russell E King

Co-Chair of Committee

Dr. Donald P Warsing Jr.

Co-Chair of Committee

Dr. S Sebnem Ahiska

Dr. Anita Vila-Parrish

DEDICATION

To my parents,

for their endless support and encouragement.

BIOGRAPHY

James A Lavin was born in Raleigh, North Carolina on March 26, 1986. After growing up in Cary and graduating from Apex High school in 2004 he attend North Carolina State University for both undergraduate and graduate school from 2004 – 2012. His research interests are in supply chain and inventory/production systems. James will transition to a career in industry.

ACKNOWLEDGMENTS

There are so many people that have helped me along the way. There were many points through the journey that I was close to quitting but I made it through with the help of many people that I would like to thank. I first want to acknowledge family for their never-ending support during my college career. While I was given a hard time about being a professional student and my “research” (always in quotes like it was not real), they always supported me. Thank you to my friends both in and out of school. Nils and Yingying you are the best research group members. My running group you guys gave me an outlet to let out my frustration and kept me somewhat sane. Of course none of this would have been possible without my committee, especially my co-chairs Dr. King and Warsing, who always pushed me to be better. They constantly guided me through my graduate school career, and shaped me into the researcher I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Determining the value of a stock-out is difficult, yet current inventory literature relies on an estimated stock-out cost value in order to make an ordering policy. Arrow et al. (1951) argue that the value of a stock-out is the amount a seller is willing to pay to satisfy the unmet demand. This is a vague quantity which is hard to determine without knowing a great deal of information concerning both the firm's cost and their consumers' behavior. Individual consumers' reactions to stock-outs vary based on the consumer's personality and product-specific characteristics (Zinn and Liu, 2001). High stock-out rates are commonly observed across the United States. It is estimated that 8.2% of grocery items and 35% of women's apparel are out of stock at any given time (Corsten and Gruen, 2003). These large percentages of out-of-stock products can lead to a significant number of customers leaving without their intended purchase.

Retailers can influence the number of stock-outs by altering their ordering decisions. The optimal order policy is a balance between having to purchase and hold extra inventory versus the cost of a stock-out. Overestimating the impact of a stock-out will increase the amount ordered causing larger than necessary inventories to be held. Shelf space is a limited and precious resource that cannot be wasted on unnecessary inventory. On the other hand, undervaluing the cost of a stock-out leads to small orders resulting in frequent stock-outs. Modeling stock-outs accurately will find the optimal balance between holding inventory and customer disappointment.

The most common method to represent the cost of a stock-out in an inventory system is to use a penalty cost. A penalty cost is an immediate one-time cost imposed on a retailer for each stock-out. The penalty cost value is comprised of both the immediate loss of revenue from missing a sale and the long-term impacts that occur as a consequence. Commonly, long-term costs are described as loss of consumer "goodwill," which includes losses of future demand (Liberopoulus et al., 2010).

However, it is difficult to put a singular monetary value on the loss of goodwill. This leads some decision makers to simply guess the stock-out cost (Hadley and Whiten, 1963). The penalty cost has a major impact on the order quantity; using an incorrect penalty cost can lead to poor ordering decisions.

One alternative to the traditional penalty cost model is the *perturbed demand* (PD) model. The perturbed demand model was first introduced by Schwartz (1966), where stock-outs alter future demand as a consequence of unsatisfied consumers not returning for future purchases (or convincing others to shop elsewhere). Perturbed demand models are founded on the idea that customers who are faced with a stock-out are less likely to return in the future because they will get their needs satisfied elsewhere. This claim is supported by consumer behavior literature (Verbeke et al., 1998; Campo et al. 2004). We assume loss of future demand is directly proportional to the number of stock-outs that occur. It is commonly assumed that customers will eventually forgive/forget that they were disappointed and return to their previous ordering behavior (Schwartz, 1966). This assumption enables the system to find a steady state balance between customers leaving due to stock-outs and customers returning. Without the ability to gain demand back over time, demand would monotonically decrease. Schwartz's model is only concerned with the long term state of the system after demand has reached a steady state rate. The assumptions made concerning the order policy and state space limits the applicability of the model. Our model relaxes some of his assumptions, thereby allowing demand loss to be updated periodically and enabling the model to be more realistic.

This dissertation extends the current perturbed demand literature to include periodic updating of demand loss and lead time. The model developed in this paper is a stochastic demand, periodic review perturbed demand model with lost sales and non-seasonal demand. A method is developed to modify demand as a function of the number of stock-outs that occur. Chapter 2 analyzes the instantaneous order arrival case in an infinite-horizon setting while in Chapter 3 the model is extended to include lead time. Chapters 3 and 4 relax Schwartz's assumption by allowing orders to be made based on the current state of the system in each period. The state of the system is defined as the current inventory level and a measure of current demand, which represents how much of the original demand will be retained as a result of stock-outs in previous periods. Chapter 4 assumes that the order policy must be specified before the start of a finite horizon selling season. Finally, Chapter 5 summarizes the dissertation results and proposes areas for future work.

Chapter 2: Infinite Horizon Stochastic Perturbed Demand Model with Instantaneous Order Arrival

Traditional penalty cost models have been used extensively to the study infinite horizon periodic-review models. They have well-defined optimal policy structures making them easy to implement. We extend this work into models that have so-called perturbed demand where stock-outs alter future demand. It is proven that these models have an optimal policy which is base-stock policy where the order up to level is dependent on the demand deflation factor that is calculated considering stock-outs in current as well as past periods. The policy parameters are used to provide insight into the value of the implied penalty cost.

2.1. Introduction

Stock-outs are more common than many retailers are comfortable with. On average, 8% of grocery items and 35% of women's apparel items are out of stock at any given time (Corsten and Gruen, 2003). The effects of the stock-outs are a loss of an immediate sale and a loss of future sales. The most common method for quantifying the impacts of a stock-out is to impose a penalty cost on a retailer at the time of the incident. The penalty cost can reasonably estimate the short-term cost of a stock-out, which is an opportunity cost imposed on a retailer for not being able to make a sale. Long-term effects of stock-outs however are hard to include when using this type of model. These issues include the potential loss of future demand and goodwill, the value of which is difficult to estimate in a single cost value.

The consumer behavior literature suggests consumers have a wide spectrum of reactions when faced with a stock-out. Schary and Becker (1978) suggest that consumers faced with a stock-out do not always return to their original ordering behavior. Consumer reactions are based on differences in their personalities as well as the product type (Camp et al. 2003). The most common actions undertaken by consumers in the face of a stock-out are switching brands, switching stores, and postponing the purchase (Verbeke et al., 1998). The immediate impact of a stock-out takes the form of consumers purchasing an alternative product or the loss of sales. In the long-term, consumers may not return to their original ordering behavior or may do so only after the memory of the stock-out experience fades. At worst, the switch to a different brand or store may be permanent, leaving a lasting impact on the demand distribution. Schwartz (1966) was the first to merge consumer behavior research with inventory control models by assuming that future demand is modified as a result of stock-outs under a known and fixed demand rate. This type of model is referred to in the literature as a *perturbed demand* model. This model incorporates how stock-outs alter consumers' long term purchasing behavior. Missed sales result in altered future demand instead of a penalty cost. Schwartz (1966) argued that this is a more direct way of modeling the effects of stock-outs and is a more realistic representation of the actual consumer behavior.

This paper extends the perturbed demand literature into the stochastic demand infinite horizon periodic-review setting. Our model assumes that the demand is non-seasonal; stock-outs result in lost sales; there is no fixed order cost; and orders arrive instantly before the start of demand in the period. Schwartz (1966) assumed a base stock policy in his model because there was no fixed order cost; the optimal policy for our model is proven to be base-stock, dependent on the level of "demand deflation", which determines the current demand distribution. Additionally, we prove that the optimal policy converges to a steady state policy. The remainder of this paper reviews the relevant literature, develops our model, and proves that the optimal policy structure is a variation of the traditional base-stock. Then, approximations are developed to improve the solution time and relate our model back to the penalty cost setting.

2.2. Literature Review

The consumer behavior literature suggests stock-outs alter consumer ordering behavior. Numerous studies show that when companies are faced with a stock-out, three common consumer reactions occur: they switch products, delay the purchase, or leave the store (Zinn and Liu, 2001; Verbeke et al., 1998). Surveys have been conducted to estimate the percentage of people who would pick each alternative for various products (Verbeke et al., 1998; Corsten and Gruen, 2003; Campo et al., 2004). These studies show numerically that a consumer's reaction to stock-outs is product dependent. Zinn and Liu (2001) studied which product and situational factors have the greatest influence on the consumer's out-of-stock decision. The most influential factor is how surprised consumers are when the product is unavailable. Zinn and Liu (2001) found that consumers of new products understand the difficulty of keeping these items in-stock; consequently, they are not surprised when the product is unavailable. The lack of surprise results in consumers being more forgiving for stock-outs of new products. There is a significant amount of research supporting the idea that consumers alter their behavior after stock-outs occur, and their behavior depends greatly on the product (Walter and Grabner, 1975; Emmelhainz and Stock, 1991; Verbeke et al., 1998; Corsten and Gruen, 2003; Campo et al., 2004; Van Woensel et al., 2007).

Traditional periodic-review inventory models that incorporate a penalty cost have a well-defined optimal policy structure. When there is no fixed order cost, a base-stock policy is optimal (Bellman et al., 1955). Base-stock polices order up to a set amount (S) at the start of each period. One of the first studies to address this topic was by Karlin (1952), who studied a single-period problem. The research evolved quickly to include fixed order costs, which changed the order polices into the form (s, S), which requires the total cost function to be K-convex (Scarf, 1960). Under this policy it is optimal to order up to S when the inventory is at or below s. The base-stock policy structure is a special case of the (s, S), with $s = S - 1$. The (s, S) policy structure has been studied by many other scholars e.g. Iglehart (1963), Veinott (1966), and Zheng (1991).

The limiting behavior of the base-stock and (s, S) policies was studied by Bellman et al. (1955), Iglehart (1963) and Johnson (1968). These papers show that the policy converges to a steady-state policy as the number of periods in the planning horizon increases. Steady state policies allow the problem to be extended to the infinite horizon case. It is shown that as the literature has found, our model has similar limiting behavior.

Calculating the optimal steady-state policy is computationally intensive for both penalty cost and perturbed demand models. Many different techniques have been developed to solve the penalty cost version of the problem. A bounded search method suggested by Veinot and Wagner (1965) finds the optimal policy but it is slow. Sethi and Cheng (1993) used a Markov chain to find optimal inventory polices. Roberts (1962) used renewal theory to calculate an approximate policy, which is fairly accurate. Ehrhardt (1979) and Ehrhardt and Mosier (1985) suggested a near-optimal and very fast power regression approximation. This work was later extended to include service level constraints by Schneider and Ringuest (1990). Additional optimal solutions or approximations are found in publications by Wagner et al. (1965), Tijms (1972), Johnson (1968), and Bell (1970). As they all found, exact solutions are difficult to obtain computationally. Approximations have been developed to find faster solutions and are generally accurate in the penalty cost setting.

Unlike the aforementioned papers, the model developed in this paper does not incorperate a traditional penalty cost. Instead our model assumes perturbed demand, first formulated by Schwartz (1966, 1970), in which future demand is altered when customers are faced with a stock-out. Other models where demand is affected by ordering decisions include those by Hill (1976), Ernst and Powell (1995), and Diana and Petruzzi (2001). All of these models assume that demand is altered as a function of the ordering decision and not the actual stock-out events. Only Robinson (1990) uses the actual number of disappointed customers in each period to modify demand periodically. In this model, the current number of satisfied and unsatisfied customers along with the time dependent changes to the underlying demand distribution change the demand seen by a retailer. Our underlying demand distribution is stationary, with underlying demand being modified by a scalar multiple whose value is based on recent stock-outs and two consumer behavior parameters to yield the realized demand. This paper takes Schwartz's ideas on how stock-outs affect demand and apply them to an infinite-horizon periodic-review model with stochastic demand.

2.3. Model formulation

We develop a stochastic demand, periodic-review model with perturbed demand as an alternative to imposing a penalty cost for lost sales. We assume that in the absence of stock-outs, the demand each period follows a stationary cumulative distribution $G(\xi)$, which we will refer to as the *underlying demand distribution*. However, the consumer behavior literature has demonstrated empirically that current stock-outs affect the immediate future demand (some customers abandon this product and/or retailer while others will return to purchase at a later date). Thus, consistent with other modeling work with perturbed demand, we employ a *demand deflation factor* in each period *n*, denoted by α_n. This factor is computed based on the level of demand lost due to stock-outs in periods prior to period *n*. We state our model in terms of periods-to-go, such that period *n*–1 follows period *n*. Thus, the demand deflation factor with $n-1$ periods remaining, α_{n-1}, is based on actual demand fulfillment performance in period *n*, and is computed at the start of period $n-1$, before demand occurs. The actual realized demand with $n-1$ periods remaining is given by $\alpha_{n-1}\xi_{n-1}$ where ξ_{n-1} represents the realized sample from the underlying demand distribution $G(\xi)$. Thus, if the expected underlying demand is $\mu = \int \left(1 - G(\xi)\right) d\xi$, the expected realized demand, given demand deflation factor α_{n-1}, is $\alpha_{n-1}\mu$.

The timeline of events is similar to traditional periodic-review inventory models. At the start of the current period, say period n, we evaluate the state of the system, which consists of the current on-hand inventory (x_n) and current demand deflation factor (α_n). Based on the state values, the base-stock quantity S_n is determined, and the purchase cost is imposed at the time of the order. The order arrives before demand is realized. Holding cost and revenue are calculated based on the inventory remaining on-hand after fulfilling realized demand.

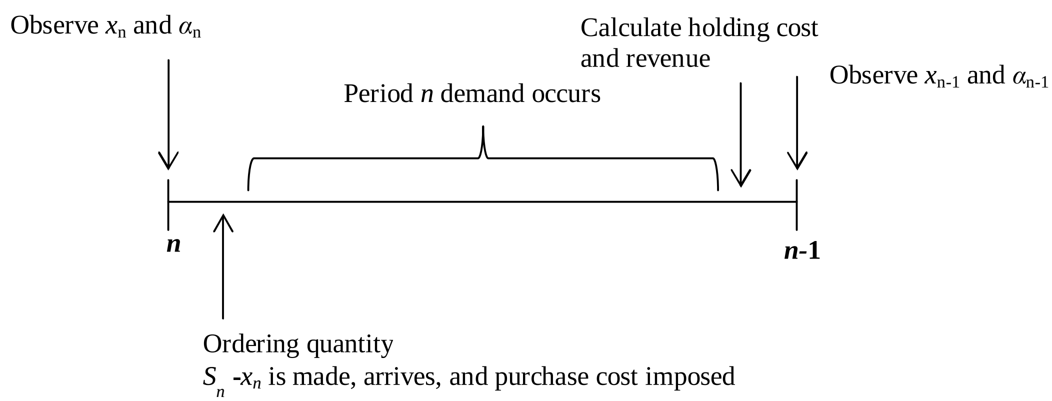


Figure 2-1 Timeline of events for zero lead time

A unit holding cost, h, is incurred for each item of inventory on hand at the end of period after demand occurs. For each unit sold, revenue r is received. Revenue is modeled as a negative cost since the problem is formulated as a standard cost minimization problem. Then, the immediate (one period) net cost, $L(y|\alpha)$, for y units on-hand given a demand deflation factor α, is calculated as follows:

$$L(y\mid\alpha)=\int_0^{\frac{y}{\alpha}} h(y-\alpha\xi)g(\xi)d\xi-\int_0^{\frac{y}{\alpha}} r\alpha\xi g(\xi)d\xi-ry\left(1-G\left(\frac{y}{\alpha}\right)\right). \tag{1}$$

Assuming unit purchase cost c ($c < r$), and given x_n units in inventory with a demand deflation factor of α_n, the total cost over the next n periods is written as the recursive dynamic program

$$f_n(x_n,\alpha_n)=\begin{cases}\min\limits_{S_n\ge x_n}\left\{\begin{array}{l}c(S_n-x_n)+L(S_n\mid\alpha_n)\\ +\omega\int_0^{\infty} f_{n-1}\left((S_n-\alpha_n\xi)^+,\alpha_{n-1}\right)g(\xi)d\xi\end{array}\right\} & n>0\\ 0 & n=0\end{cases} \tag{2}$$

$$f_0(x_0,\alpha_0)=0$$

where $g(\xi)$ is the underlying demand density function, ω is a discount factor ($0\le\omega\le1$), and $(a)^+=\max\{a,0\}$. The discount factor is limited between zero and one because future periods are assumed to be discounted by the time value of money.

Our claim is that, while the demand deflation factor α_n is influenced by prior lost sales, this effect may be attenuated by other aspects of consumer behavior, namely the persistence of their memory of past stock-outs and the possibility that some customers are willing to effectively "shrug off" a stock-out such that it does not affect their likelihood of a future purchase in any way. While the specific effect on demand is not known (we found no studies of consumer behavior that directly address this issue), we can attempt to capture presumed consumer behavior in two model parameters: β, which we call the *deflation intensity factor,* reflecting the strength of the effect of stock-outs on future demand; and λ, which we call the *deflation persistence factor*, reflecting how long past stock-outs continue to affect future demand.

Schwartz (1966) suggested that the effect of past disappointment will decline exponentially over time. Our model incorporates the consumer memory by exponentially smoothing the demand deflation factor across periods. The value of the demand deflation factor with *n*–1 periods remaining is dependent on the incoming value of the demand deflation factor (α_n); the realized sample from the demand distribution in period *n* (ξ_n); and the available inventory on-hand after the order arrives in period *n* (S_n), as follows:

$$\alpha_{n-1} = \begin{cases} \lambda(1) + (1-\lambda)\alpha_n & S_n \geq \alpha_n \xi_n \\ \lambda\left(1 - \dfrac{(\alpha_n \xi_n - S_n)\beta}{\alpha_n \xi_n}\right) + (1-\lambda)\alpha_n & S_n < \alpha_n \xi_n \end{cases} \tag{3}$$

From expression (3), we can see that, for a prior period (*n*, preceding period *n*–1) in which no demand is lost (in other words, available inventory was sufficient to meet the actual realized demand), α_{n-1} is computed by smoothing it with α_n, reflecting the "deflation persistence" from previous periods. Thus, if λ(the persistence factor) is close to 1, the effect of previous losses will decline rapidly; conversely, if λ is close to 0, previous demand fulfillment performance will dominate the current performance. From expression (3), we see that for a prior period in which there is underage $U_n = (\alpha_n \xi_n - S_n)^+$, the deflation intensity factor (β) lessens the impact of lost sales from the previous period on future sales through an increase in α_{n-1}.

Note that the demand deflation factor, α_n, is bounded between zero and one. This implies that demand cannot be negative, nor can it exceed the upper limit of the underlying demand distribution. It is also assumed that the intensity (β) and persistence (λ) factors are both constrained to be within the range of zero to one, inclusively. The persistence factor does not make sense outside of this range because it is an exponential smoothing factor, which is limited within the same bounds. The intensity factor is limited to constrain the model by not allowing stock-outs to increase demand (negative β) or allowing word of mouth effects (β>1). Implementing these two complications would require changes in some equations to account for the possibility that the demand deflation factor could be negative or above one. Based on consumer behavior studies, the probability of permanently switching stores is low (Walter and Grabner, 1975; Emmelhainz and Stock, 1991), hence the cap on β. It is assumed that the demand deflation factor is constant throughout each period, meaning that stock-outs within the period will not further deflate demand during that period.

This paper is concerned with the infinite horizon case. Below we show that the optimal policy in this setting converges to a steady-state base-stock policy in the long run.

Proposition 1: A demand deflation factor dependent, base-stock policy is optimal for the infinite horizon problem under perturbed demand.

Proof: Inventory models are base-stock when three conditions hold (Karlin, 1958):

i. $f(0)$ is not a relative minimum
ii. $\lim_{S_n \to \infty} f\left(S_n\right) = \infty$
iii. $f^{'}\left(S_n\right) = 0$ has a unique solution

There is no penalty cost in our model. Condition (i) is true when $c < r$: it is always better to order at least one unit. Condition (ii) states that as the order quantity goes to infinity, the cost approaches infinity, thereby deterring the purchase of an infinite number of units. In our model, as the order quantity approaches infinity, the materials and holding cost also approach infinity, but the number of units sold approaches the expected value of the demand function. Therefore, the second condition (ii) holds and is shown mathematically as follows:

$$\lim_{S\to\infty} f_n\left(x_n,\alpha_n \mid S\right)=\left\{c\left(\infty-x_n\right)+L\left(\infty \mid \alpha_n\right)+\omega\int_0^\infty f_{n-1}\left(\infty-\alpha_n\xi,\lambda+\left(1-\lambda\right)\alpha_n\right)g\left(\xi\right)d\xi\right\}$$

$$\begin{aligned} L\left(\infty \mid \alpha\right) &=\int_0^\infty h\left(\infty-\alpha\xi\right)g\left(\xi\right)d\xi-\int_0^\infty r\alpha\xi g\left(\xi\right)d\xi-r\infty\left(1-G\left(\infty\right)\right) \\ &= h\infty-r\alpha E\left[\xi\right]\to\infty \end{aligned}$$

Condition (iii) holds because the second derivative of the total cost function, given an initial demand deflation factor, is always positive, and the first derivative transitions from negative to positive as we show below.

$$L\left(S_n \mid \alpha\right)=\int_0^{\frac{S_n}{\alpha}} h\left(S_n-\alpha\xi\right)g\left(\xi\right)d\xi-\int_0^{\frac{S_n}{\alpha}} r\alpha\xi g\left(\xi\right)d\xi-rS_n\left(1-G\left(\frac{S_n}{\alpha}\right)\right)$$

$$\frac{dL\left(S_n \mid \alpha\right)}{dS_n}=\left(h+r\right)G\left(\frac{S_n}{\alpha}\right)-r$$

$$\lim_{S_n\to 0}\frac{dL\left(S_n \mid \alpha\right)}{dS_n}=-r$$

$$\lim_{S_n\to\infty}\frac{dL\left(S_n \mid \alpha\right)}{dS_n}=h$$

Below we show that the immediate reward function is convex since its second derivation is positive, i.e.

$$\frac{d^2L\left(S_n \mid \alpha\right)}{d^2S_n}=\frac{1}{\alpha}\left(h+r\right)g\left(\frac{S_n}{\alpha}\right)\geq 0$$

Thus, the single period cost function $f_1\left(x_1,\alpha_1\right)=c\left(S_n-x_n\right)+L\left(S_n \mid \alpha_n\right)$ is convex since it is the sum of two convex functions. Similarly, the two-period problem is convex because the immediate reward is convex and the single period problem is convex. Through induction, it can be shown that the *N*-period problem is convex (see appendix).

Therefore, all three of Karlin's (1958) conditions hold; and the optimal policy is a base-stock policy dependent on the value of the demand deflation factor. □

While Schwartz (1966, 1970) assumes the optimal policy takes on a base-stock structure, we have proven that the structure of the optimal policy of our model is base-stock and whose value is dependent on the demand deflation value. The dependence on the demand deflation value makes implementing the policy more complex than in the penalty cost setting. Still, knowing the policy's structure aids in determining the optimal policy.

Schwartz (1966), Ernst and Powell (1995) assume that the optimal policy converges in steady state, but they never prove the policy converges. Below we provide a proof of convergence for our problem.

Proposition 2: The demand deflation factor dependent, base-stock levels converge to a steady state policy as the number of periods remaining approaches infinity.

Proof: Bellman et al. (1955) give four conditions that when true result in the optimal solution being unique and bounded. Iglehart (1963) states that if the solution is unique and bounded, the order policy will converge to a single policy. The Bellman et al. (1955) conditions for having a unique and bounded optimal solution are:

i. $g(\xi)$ is a probability density function
ii. The penalty cost is continuous and monotonically increasing
iii. The order cost is continuous and ordering nothing has a cost of zero
iv. $0 \le \omega \le 1$

Conditions (i) and (iv) hold by definition of the demand density function $g(\xi)$ and discount factor ω. For condition (ii), our problem does not have a direct penalty cost, but it does have an indirect penalty cost consisting of loss of revenue and future demand. It is monotonically increasing cost as the number of missed sales increases. An increase in the number of stock-outs yields a larger loss of demand in future periods: thus additional potential revenue is lost. This is shown in more detail in the appendix. Condition (iii) holds because the order cost is linear with an intercept of zero. Conditions ii and iii are specific to continuous distributions; Veinott and Wagner (1965) proved that theses conditions hold in the discrete case. Therefore, all the conditions hold, and our problem has a policy that converges to a single steady state policy as *n* approaches infinity. The policy is fixed where the base-stock level is dependent on the current level of demand deflation. □

2.4. Solution Procedure

The optimal solution can be determined through the use of the value iteration procedure that solves the Markov Decision Process (MDP) formulation of the problem (Howard, 1960). The value iteration procedure was originally designed to calculate both the optimal policy and the optimal gain for finite horizon problems but it can also be used for the infinite horizon problems, because as the number of periods considered increases, the policy converges (Zhang and Zhang, 2001). The Markovian assumptions hold in this case because realized demand and, consequently, the transition probabilities are only dependent on the current state of the system, i.e. inventory and demand deflation value at the beginning of period. Although, the demand deflation value can, in reality, assume any of an infinite number of continuous values we discretize it into uniform segments defined by a selected grid size in order to have a finite-state MDP. Chow and Tsitsiklis (1988) developed a method for discretizing a continuous variable into uniform segments for use in a MDP. The demand deflation value is rounded to the nearest discretized level, transforming it from a continuous to discrete variable.

The MDP-based procedure is based on a backward formulated dynamic program (Howard, 1960). The procedure starts with one period remaining then increases the number of periods considered until the policy and gains have converged. The total number of states is N, where N equals the number of inventory levels times the number of discrete demand deflation levels. If the system is in state i and order quantity k is selected, then the system will transition to state j with probability p_{ij}^k based on the demand distribution. The cost for making the transition from i to j under order quantity k is noted as R_{ij}^k, which is a function of the cost parameters $\left(R_{ij}^k = ck + hx_j - r\left(x_i + k - x_j\right) \quad x_j \leq x_i + k\right)$, where x_i and x_j refer to the inventory levels for state i and j. The expected one period cost associated with being in state i and ordering k units is noted as q_i^k where $q_i^k = \sum_{j=1}^{N} p_{ij}^k R_{ij}^k$. The total expected cost when starting in state i with n periods remaining is denoted as $v_i\left(n\right)$. The goal of the problem is to choose for each state the value of k that minimizes the expected total cost.

$$v_i\left(0\right) = 0$$

$$v_i\left(n\right) = \min_k \left[q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j\left(n-1\right) \right]$$

As n increases, the policy converges to a steady state policy; specifically, the gain $\left(v_i\left(n\right) - v_i\left(n-1\right)\right)$ of each state converges (see Proposition 2). The policy that minimizes the gain is the optimal policy. Calculating the optimal policy is computationally intensive using this procedure because constructing p_{ij}^k, R_{ij}^k is time consuming; additionally, it can take a large number of iterations for the gain to converge when there are a large number of states. The convergence of the gain is slower in a perturbed demand model than in traditional penalty cost models because of the addition of the demand deflation factor dimension of the state space.

An alternative MDP solution procedure is policy iteration, which is often used to evaluate steady state policies (Howard, 1960). However, policy iteration assumes that the state space is ergodic, meaning that the recurrent set of states are independent of the starting conditions. As we demonstrate in our experimentation, this assumption does not hold for all parameter combinations in the perturbed demand model; thus, it is not used to solve our model. This phenomenon is explained further in section 2.4.2.

2.4.1. Computational experiments

All of the numerical experiments conducted in this paper use problems generated by sampling from a uniform distribution using the upper and lower bounds given in Table 2-1.

Table 2-1 Experiment parameter ranges

Parameter		Range
Unit purchase cost	c	1
Unit holding cost	h	[0, 1]
Unit revenue	r	$[c+h, c+h+2]$
Deflation intensity	β	[0, 1]
Deflation persistence	λ	[0, 1]
Demand distribution shape parameter	p	[0, 1]
Demand distribution scale parameter	m	$\frac{20p}{1-p}$

Revenue is a function of the holding and the purchase costs to ensure that each problem is profitable. The demand distribution is assumed to be a negative binomial with the scale parameter (*m*) that is a function of the shape parameter p in order to keep the demand mean fixed at twenty units ($\mu = 20$). A constant mean is used to keep the problems from being too large in order to keep the solution time at a reasonable level. Additionally, a constant mean allows us to test the effects of demand variability on the optimal solution. We believe that these bounds result in a wide range of problems, enabling the evaluation of the robustness of the model.

2.4.2. Issues with discretization of the demand deflation factor

Most MDPs used for inventory control purposes are ergodic, meaning the recurrent states are independent of the starting state; however, for our problem, ergodicity is not always guaranteed due to the discretization of the demand deflation factor (α). A grid size that is too fine results in a very large state space and ultimately causes computational intractability; however, a grid size that is too coarse may ultimately cause issues with non-ergodicity. We focus on transitions to states with higher demand deflation values because transitions to states with lower demand deflation values were not observed to cause issues with ergodicity.

We denote the grid size associated with the demand deflation factor as Δ, which represents the difference between the levels of the demand deflation factor, i.e. $\alpha_n \in \{0, \Delta, 2\Delta, 3\Delta..., 1\}$. Depending on the problem parameters, it may not be possible to transition from a given system state to a state with a higher valued demand deflation value. This is due to the rounding required to discretize the demand deflation factor. The amount the demand deflation value can decrease or increase in a single period depends on the demand persistence value, λ. From Equation (3), it can be shown that largest increase in the demand deflation value over one period occurs when all the demand is met (top portion of Equation (3)). In this case

$$\alpha_{n-1} = \alpha_n + \lambda\left(1 - \alpha_n\right).$$

However, based on the rounding process, the amount of the increase from the current value must be at least $\Delta/2$ to allow α_{n-1} to transition to the next highest level, i.e. a value of $\alpha_n + \Delta$.

Mathematically,

$$\lambda\left(1-\alpha_n\right) \geq \frac{\Delta}{2}.$$

Rearranging terms yields

$$\alpha_n \leq 1-\frac{\Delta}{2\lambda}. \tag{4}$$

For a given level of the demand persistence factor (λ), if the demand deflation value at the start of a period does not satisfy the inequality described in (4), then it is not possible to transition to the next higher demand deflation level (i.e. $\alpha_{n-1} = \alpha_n + \Delta$) regardless of the order quantity. Clearly, the higher the value of α_n the less likely it is that the inequality holds for any given values of λ and Δ. Therefore, if it is possible to transition from a state with α_n = 1-Δ (the second highest value of the demand deflation factor) to a state with α_{n-1} = 1 (the highest value of the demand deflation factor) then all other transitions are also possible. Substituting α_n = 1-Δ into equation (4) yields

$$1-\Delta \leq 1-\frac{\Delta}{2\lambda}$$
$$\left(2\lambda-1\right)\Delta \geq 0$$

This inequality holds and the process is ergodic if either Δ=0 (meaning no discretization) or if $\lambda \geq 0.5$ (meaning customers put an importance weight that is more than 50% on the most recent stock out experience). If neither condition holds, then it is still possible that the process is ergodic under optimal control but there is no guarantee of ergodicity.

In practice, the MDP solution procedure requires $\Delta > 0$ and it is possible for products to have λ< 0.5. The smaller the value of Δ the more likely the process will be ergodic when $\Delta > 0$. Fifty test problems with λ< 0.5 were randomly generated, and the percentage of ergodic problems was determined for different values of Δ in the range [0.01, 0.25]. As Δ decreased, the percentage of ergodic problems increased non-linearly as shown in Figure 2-2.

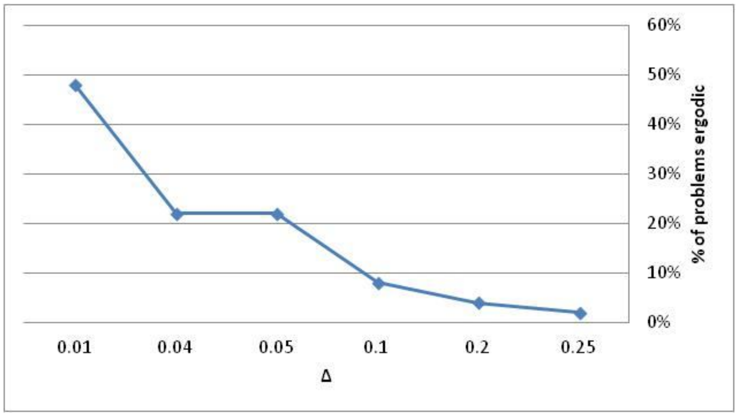


Figure 2-2 The percentage of ergodic problems across different Δ values

When Δ=0.02, nearly half of the problems were ergodic. In general, these were the ones with higher λ values. The trade-off is a realistic grid size vs. speed: decreasing Δ increases the size of the state space, resulting in a longer solution time.

2.4.3. Unknown number of stock-outs

An additional complexity results when the number of stock-outs is not known. Many retailers are faced with this dilemma—they know when they have units on the shelf and when they don't, but they are unaware of how many sales are missed when their observed inventory falls to zero, i.e. when the shelf is empty. It is impractical to ask every customer who leaves the store whether they intended to buy the unavailable product, i.e. whether they encountered an empty shelf when seeking to make a purchase. Consequently, a retailer's actual demand deflation value may not be known with certainty.

The unknown number of stock-outs situation can be formulated as a partially observable Markov decision process (POMDP). The POMDP belief state is the probability that you believe the system is in a particular state, which extends the state space to include a belief aspect. However, even with a relatively large Δ , the state space of the POMDP belief state becomes unmanageably large making this approach impractical for our problem.

An alternative to the POMDP formulation is to use a point estimator for the demand deflation value based on whether or not the product stocked out and the available demand data. The case where we assume the number of stockouts is known every period (and thus, the demand deflation value is known) can be solved to optimality as shown in Section 2.4 using the MDP model. The optimal policy obtained under this assumption can be implemented using an estimated value for the demand deflation value as a surrogate for the actual value. When the estimated demand deflation value turns out to be the same as the actual demand deflation value, the order quantity is optimal. As the difference between the estimated and actual demand deflation values increases, the order quantity may diverge from optimal.

To test the cost, in terms of overall system gain, of not knowing the actual number of stockouts (and therefore not knowing the actual demand deflation value), the following procedure is used. First, the optimal policy is determined by solving the MDP formulation for the case where the value of the demand deflation factor is assumed known. Then this policy is implemented in a Markov Process (MP) setting where the order quantity each period is based on the current inventory and current estimated demand deflation value. The transition probabilities are based upon the true demand deflation value. This means that the state space has to be extended to include an additional state variable, the estimated demand deflation value. The resulting state space consists of the inventory level, the true demand deflation value and the estimated demand deflation value. This procedure provides an upper bound on the POMDP performance.

A conditional expectation can be used to estimate the number of stock-outs. When there is inventory left on the shelf at the end of the period, then there is no need to estimate the number of stock-outs. Otherwise, the expected number of stock-outs is estimated as the expected demand above the base stock level S_n. Earlier, we define underage when there is *n* periods to go as $U_n = \left(\alpha_n \xi_n - S_n\right)^+$; thus we can define the estimated underage as $\hat{U}_n = E\left[U_n \mid \hat{\alpha}_n \xi_n > S_n\right]$, which is calculated as shown below. Within the MP framework when the realized demand exceeds the inventory we estimate underage using

$$\hat{U}_n = \frac{\int_{\frac{S_n}{\hat{\alpha}_n}}^{\infty} \left(\hat{\alpha}_n \xi_n - S_n\right) g\left(\xi_n\right) d\xi}{\int_{\frac{S_n}{\hat{\alpha}_n}}^{\infty} g\left(\xi_n\right) d\xi}.$$

Then the demand deflation can be updated using the estimated underage as follows

$$\hat{\alpha}_{n-1} = \lambda\left(1 - \frac{\hat{U}_n \beta}{\hat{\alpha}_n \xi_n}\right) + \left(1 - \lambda\right)\hat{\alpha}_n.$$

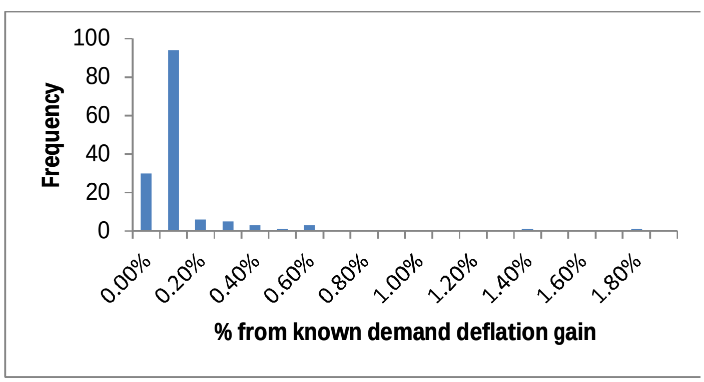
The order quantity is then based on the estimate demand deflation and current inventory.

It is assumed that the demand deflation value is 1 (i.e. no stock outs from previous periods) at the beginning of the planning horizon, $\hat{\alpha}_\infty = 1$ since it is sensible to the start of the horizon with no demand lost yet.. To estimate the demand deflation value with *n*-1 periods remaining, we use the most recent estimation $\hat{\alpha}_n$. Demand is known to be at least *S* units when there is a stock-out; therefore, only the part of the demand distribution above *S* matters when calculating the expected number of stock-outs. The estimated value of α will be used in the expected stock-outs formula because the true demand deflation value is unknown. The expected number of stock-outs is used to estimate the demand deflation value. Like the actual demand deflation value, the estimated demand deflation value changes over time as a function of stock-outs. The difference is that the exact number of stock-outs is not known, hence an estimated value is used when a stock-out occurs at the end of a period. This procedure to develop an order policy is used with an MP framework to evaluate the gain of the approximation.

The impact on the system gain associated with using an estimated demand deflation value was tested through 150 randomly generated problems using the bounds described in Table 2-1. The optimal gain for the case where the demand deflation value is known was compared to the gain obtained when the same policy is implemented considering the estimated demand deflation values instead of the true demand deflation values. The addition of the estimated demand deflation factor to the state space increases the number of states significantly, which makes evaluating the MP infeasible for very fine grid sizes. For that reason, a grid of $\Delta = 0.02$ was used. The results, displayed in Table 2-2, indicate that the value of having a perfect stock-out information is relativity low, i.e. not knowing the actual stock-out amounts in every period does not result in significant revenue loss. The deviation from optimal gain when an estimated value for demand deflation factor is used was only 0.06% on average among the test problems used.

Table 2-2 Percentage deviation from optimal gain when estimated demand deflation value is used

% deviation from optimal gain	
Minimum	0.00%
Mean	0.06%
Max	1.73%



Using an estimated demand deflation value to implement the optimal policy works well in the long run. The period immediately following a stock-out has the largest disparity between actual and estimated demand deflation values. For most problems, stock-outs are rare under optimal control, meaning that large disparities are uncommon. When no stock-outs occur both the estimated and actual demand deflation values tend towards one. Both values asymptotically approach one thus the difference between the two estimates shrinks between stock-outs occurances. Using the policy based on an estimated demand deflation value yields an overall gain close to optimal because over the long run the difference between actual and estimated demand deflation values is small. Therefore, for the rest of this paper, we assume that the number of stock-outs is known with certainty and thus the demand deflation value is known as well.

2.5. Policy approximations

The optimal policy can be determined by solving the MDP model using the value iteration method, but that is computationally intensive. As an alternative, approximations are developed to reduce the computational effort. In the optimal policy, for each α level there is a potentially different base-stock level, which makes the implementation of the policy not practical given a large number of states. A lookup table is needed to implement the policy. Approximations are developed to decrease the computational time and provide a straightforward and implementable policy.

2.5.1. Derivative approximations

The use of a derivative to determine the optimal order policy is intising because of the convex nature of the cost function. Taking the derivative and setting it equal to zero will give an optimal order quantity analytically if it is possible to take the derivative. An approximation is developed based on the derivative of the cost function, by considering only up to three periods in the planning horizon in order to keep the equations manageable. Exponential smoothing of the demand deflation factor causes the decision made with n periods remaining to have some effect on every subsequent period, but at an exponentially decreasing rate of importance. The interdependence between periods makes the actual derivative too complicated to use when looking at problems with large n values. The infinite horizon case is the limiting behavior as n tends towards infinity. Our approximation only uses a few periods because the most recent stock-out or fulfilled demand events influence the demand deflation factor the most.

When n=1, the optimal order up to level can be determined by taking the first derivative of the cost function, i.e., equation (2), with respect to order up to level and setting it equal to zero, as shown below.

$$\frac{df_1(x_1,\alpha_1)}{dS_1} = c - r + (r+h)G\left(\frac{S_1}{\alpha_1}\right) = 0$$
$$\Rightarrow S_1^* = G^{-1}\left(\frac{r-c}{r+h}\right)\alpha_1$$

Thus, a one-period problem has a critical fractile solution. Fractile based solutions are often seen in penalty cost models. The one period solution is easy to implement but since it considers only one period, it cannot adequately model the effects of stockouts on future demand. The demand deflation factor cannot modify future demand because only one period is considered.

The derivative of the cost function for the two-period problem (n=2) has a similar fractile but it also takes into consideration the altered demand in the second period resulting from stockouts at the end of first period. Lost sales are considered for only one period, making the two-period derivative short sighted and thus undervaluing the loss of future demand.

$$\frac{df_2\left(x_2,\alpha_2\right)}{dS_2}=\left\{\begin{array}{l} c-r+\left(r+h-\omega c\right)G\left(\dfrac{S_2}{\alpha_2}\right) \\ -\omega\left(h+r\right)\left(\dfrac{\beta\lambda}{\alpha_2}\right)\displaystyle\int_{\frac{S_2}{\alpha_2}}^{\infty}\frac{1}{\xi_2}\left[\int_0^{\frac{S_1^*}{\left(\lambda-\lambda\beta+\frac{S_2\beta\lambda}{\alpha_2\xi_2}+(1-\lambda)\alpha_2\right)}}\xi_1 g\left(\xi_1\right)d\xi_1\right]g\left(\xi_2\right)d\xi_2 \end{array}\right\}$$

When three periods (*n*=3) are considered, the effects of lost sales are larger but the cost function has a much more complicated derivative as shown below. Additionally, setting the derivative equal to zero in the two and three period derivatives requires solving nested smaller problems. The *n* = 2 solution depends on the optimal solution for one-period problem, i.e. $S_1^*\left(\alpha_1\right)$. Similarly, the decision with 3 periods remaining depends on $S_2^*\left(\alpha_2\right)$ and $S_1^*\left(\alpha_1\right)$. Solving all the nested problems makes solving the large *n* period derivatives difficult to impliment.

$$
\frac{df_3\left(x_3,\alpha_3\right)}{dS_3}=\left\{
\begin{array}{l}
c-r+\left(r+h-\omega c\right)G\left(\frac{S_3}{\alpha_3}\right) \\
-\omega\left(h+r-c\omega\right)\left(\frac{\beta\lambda}{\alpha_3}\right)\int\limits_{\frac{S_3}{\alpha_3}}^{\infty}\frac{1}{\xi_3}\left[\int\limits_0^{\frac{S_2^*}{\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_3\xi_3}+(1-\lambda)\alpha_3\right)}}\xi_2 g\left(\xi_2\right)d\xi_2\right]g\left(\xi_2\right)d\xi_2 \\
-\omega^2\int\limits_{\frac{S_3}{\alpha_3}}^{\infty}\left[
\begin{array}{l}
\int\limits_0^{\frac{S_2^*}{\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_2\xi_2}+(1-\lambda)\alpha_2\right)}}\left[\int\limits_0^{\frac{S_1^*}{\left(\lambda+(1-\lambda)\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_3\xi_3}+(1-\lambda)\alpha_3\right)\right)}}\left(h+r\right)\left(1-\lambda\right)\frac{\beta\lambda}{\alpha_3\xi_3}\xi_1 g\left(\xi_1\right)d\xi_1\right]g\left(\xi_2\right)d\xi_2 \\
+\int\limits_{\frac{S_2^*}{\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_2\xi_2}+(1-\lambda)\alpha_2\right)}}^{\infty}\left[\int\limits_0^{\frac{S_1^*}{\left(\lambda-\lambda\beta+\frac{S_2^*\beta\lambda}{\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_3\xi_3}+(1-\lambda)\alpha_3\right)\xi_2}+(1-\lambda)\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_3\xi_3}+(1-\lambda)\alpha_3\right)\right)}}\left(r+h\right)\left(\left(
\begin{array}{l}
\frac{S_2^*\frac{\left(\beta\lambda\right)^2}{\alpha_3\xi_3}\xi_2}{\left(\left(\lambda-\lambda\beta+\frac{S_3\beta\lambda}{\alpha_3\xi_3}+\left(1-\lambda\right)\alpha_3\right)\xi_2\right)^2} \\
+\left(1-\lambda\right)\left(\frac{\beta\lambda}{\alpha_3\xi_3}\right)
\end{array}\right)\xi_1\right)g\left(\xi_1\right)d\xi_1\right]g\left(\xi_2\right)d\xi_2
\end{array}\right]g\left(\xi_3\right)d\xi
\end{array}\right\}
$$

Solving the infinite horizon problem is the main concern of this paper. For the infinite horizon steady state policy, the base-stock decision changes as a function of the demand deflation value and not the number of periods remaining. We see in the optimal policy that small changes in the demand deflation value yield small changes in the order policy. To make the implementation of the two and three period derivatives easier, we assume that the same order up-to-level *S* is used for every period, i.e. $S_3 = S_2 = S_1 = S$, for a given α value. We know that the infinite horizon policy converges (proposition 2), so we can approximate that behavior by producing a single base-stock level when using the *n*=2 or 3 approximations.

Approximations that focus on a few periods can work well because the demand deflation factor is exponentially smoothed placing the highest weight on the most recent event. This hybrid approximation modifies the two-period derivative to take into consideration more future events, resulting in a balance between the two and three period derivatives. More weight is put on loss of future demand by assuming that λ=1. We call the resulting equation the hybrid derivative approximation, which is shown below.

$$c - r + (r + h - c)G\left(\frac{S}{\alpha}\right) - (h + r - c)\left(\frac{\beta}{\alpha}\right)\left[\int_0^S \xi_1 g(\xi_1)\,d\xi_1\right]\left[\int_{\frac{S}{\alpha}}^{\infty}\left(\frac{1}{\xi_2}\right)g(\xi_2)\,d\xi_2\right] = 0 \qquad (5)$$

The above equation looks like the two-period derivative but also includes some aspects of the three-period derivative. The third term includes purchase cost associated with the second order decision like the three-period derivative does. The λ term is dropped from the fourth term to give more weight to the loss of future demand term. This was an attempt to include some of the later effects of altered demand into one term, instead of the three terms in the three-period derivative.

The decision variable (*S*) is assumed to be an integer. The derivative will transition from negative to positive across consecutive integers. We set the approximate order quantity equal to the value which has the smalled absolute value from the derivative, for a given demand deflation value. It is possible to arrive at an S value using this method for any demand deflation value, thus the demand deflation factor does not need to be discretized.

2.5.2. Single *S* approximation

A traditional penalty cost model under the same model conditions (no fixed cost and instant inventory replenishment) has a base-stock optimal order policy. The optimal policy structure of the perturbed demand model is dependent on the demand deflation factor. The use of a single base-stock level is easier to implement because the order quantity only depends on the inventory level. Additionally, not being dependent on the demand deflation factor means that it does not have to be tracked by a retailer. When the expected change in the demand deflation value across periods has low variance, using a single base-stock level, which does not allow the order to change based on demand deflation, performed well in our experiment set. Infrequent stock-outs and low demand sensitivity will lead to the demand deflation value staying relativity constant over a long horizon, allowing the use of a single base-stock level to perform well.

The cost function is convex in S, which makes solving for the optimal single base-stock value easier. The cost function is convex under the optimal policy, as shown by Proposition 1. The single base-stock (S) is approximately a weighted average of the base-stock values for each demand deflation value where the weights are the probability of being in each deflation level. The sum of convex functions is convex; thus, the total cost function is convex over a single S policy. The convex nature of the cost function allows a gradient search procedure to be used to find the base-stock value which results in the minimum cost assuming a single base stock level policy structure. The gradient searches use the hybrid derivative approximation's α=1 base-stock value as its initial solution. The base-stock levels one above and one below the initial base stock level from the hybrid approximation are then compared to the initial solution, which then moves in the direction with the highest improvement in gain. The search continues until the gain becomes worse. The step size of the gradient search is held constant at one ordering unit. Larger step sizes can be used, but the hybrid approximation gives a good initial solution, lowering the number of gradient steps to find the best base-stock level. The procedure to find the value of the single base-stock policy is explained below.

Step 0: Generate an initial S policy by using the hybrid approximation developed in Section 2.5.1 assuming $\alpha = 1$

Step 1: Calculate the gain for S using a fixed policy MP procedure

Step 2: Calculate the gains for $S+1$ and S-1 using a fixed policy MP procedure

Step 3: Set S equal to the order up to level with the minimum gain.

Step 4: If S is unchanged then stop, a local minimum is found, otherwise go to Step 2.

The single S policy is easier to implement than the optimal policy because the order quantity is not dependent on the current demand deflation value. However it can take longer to find the best single base-stock approximation value than the optimal policy because multiple policies must evaluated in order to find the best S solution. The advantage is that the policy is easy to implement and takes on the same structure as penalty cost models. If the single S policy performs well, then penalty costs models are a good approximation to the perturbed demand reality, when the appropriate penalty cost is used.

2.5.3. Experimental results

The accuracy of each approximation is tested through 200 randomly generated test problems sampled from Table 2-1. One hundred problems were generated for each of the two λ cases: λ>0.5 and λ<0.5. Each problem is then solved optimally and also using each approximation method. The MDP value iteration procedure gives the optimal policy and the optimal gain to which the gain of all approximations is compared. All approximate policy gains were calculated using a Markov process (MP) with a fixed policy. For each state of the MP, the approximations generate an order quantity, and then the respective policies are implemented to calculate the gain. Four approximations are tested against the optimal gain: the two, three and hybrid derivative approximations along with the single base-stock approximation. A summary of the performance of each approximation is shown in Table 2-3 and a comparison of solution times is displayed in Table 2-4.

Table 2-3 Percentage cost difference for various approximations

% deviation from optimal gain	Min	Mean	Max	95% CI on Mean
Two-period derivative	0.00%	0.18%	2.78%	(0.121%, 0.239%)
Three-period derivative	0.00%	0.21%	2.78%	(0.142%, 0.266%)
Hybrid derivative	0.00%	0.05%	1.48%	(0.025%, 0.072%)
Single base-stock	0.00%	0.12%	3.62%	(0.073%, 0.159%)

Table 2-4 Solution time for various solution methods

Processing Time (seconds)	Min	Mean	Max
MDP	712.10	1,653.21	8,159.63
Two-period derivative	0.04	0.08	0.80
Three-period derivative	0.09	2.66	200.70
Hybrid derivative	0.04	0.07	0.38
Single base-stock	99.59	181.85	1,196.97

All of the derivative approximations work well for the tested problems. However, the hybrid derivative approximation is statistically better than the other three approximations according to the 95% confidence interval. All three derivative approximations performed poorly for the same set of problems. A common characteristic among these problems is that they have $\lambda < 0.1$. Problems with low values of λ are slow converging or likely non-ergodic even at the fine grid size (Δ=0.005). With a low λ, more periods need to be considered in the approximation to account for the dampening effect of λ on the demand deflation value, i.e. the change in demand, up or down, persists over a longer horizon. We believe most products will not have a $\lambda<0.1$ because it seems reasonable for customers to put a significant amount of weight on the most recent events. All three derivative approximations are much faster than the optimal solution technique without sacrificing much profit. Moreover, there is no need to discretize α when using derivative approximations.

The single base-stock level is used to replicate the penalty cost optimal solution. Using the steady state probabilities, the distribution of the demand deflation values of the recurrent states can be used to predict the performance of a single base-stock policy. Using the steady state probabilities from the MP we can construct a probability mass function for the demand deflation factor. The variance (as measure by its coefficient of variation, CV) of this distribution seems to have an effect on the performance of the single base-stock policy. The CV of the steady state distribution of the demand deflation value is influenced by the specific problem parameters. Figure 2-3 shows a plot of deviation of the cost of the single base-stock approximation from optimal cost as a function of the CV of the demand deflation factor for the test problems. The CV has a non-linear impact on the deviation from optimal. The single base-stock approximation method performs better generally for the problems where the steady state demand deflation distributions have lower coefficient of variation (CV), i.e. the demand deflation factor stays relativity constant. High CV problems indicate that the demand deflation factor can change dramatically across periods which results in poorer performance of a single base-stock approximation. This is intuitive since the optimal policy (base stock) is dependent on the demand deflation value.

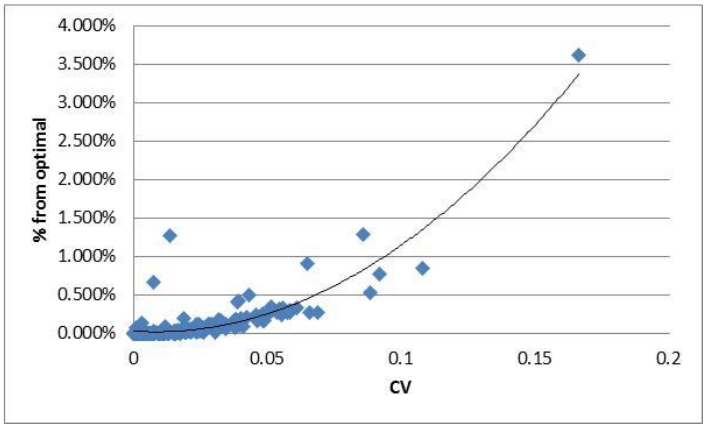


Figure 2-3 Base-stock approximation error as a function of demand deflation CV

A linear regression was used to fit the CV as a function of parameter individually to determine how the parameter affects the CV. For each term, a linear fit with significant positive slope indicates the error increases as the parameter increases while significant negative slopes show an inverse relationship.

Analysis of the regression leads to the following observations, summarized in Table 2-5. Increasing *r* lowers the CV of the distribution of the demand deflation value. Higher *r* values raise the cost of a stock-out; thus, the order policy will be adjusted to make stock-outs occur less frequently. Less frequent stock-outs means that demand deflation changes less frequently and consequently lowers the CV. High *p* values in the negative binomial distribution lower the CV of distribution of demand deflation values by shifting the underlying demand distribution. The other variables with statistically significant impact on CV of the demand deflation distribution are *h*, β, and λ which all raise the CV as they increase. Increasing *h* has the opposite impact on the order policy that increasing *r* has, increasing the likelihood of stock-outs. Frequently occurring stock-outs increases the variability in steady state demand deflation. Increasing β and λ create larger swings in α as a result of stock-outs, thus increasing the CV of demand deflation distribution.

Table 2-5 Parameter effect on CV

Parameter	Slope value
r	-0.002
p	-0.031
h	0.026
β	0.039
λ	0.033

Of the four approximation methods considered in this paper, the hybrid derivative approximation works the best. It is easy and fast to calculate and it eliminates the need for a lookup table or rounding of demand deflation because the order quantity can be quickly calculated, using current inventory and demand deflation values. The single *S* policy does have the advantage of having the same structure as a traditional penalty cost model. The single *S* policy approximation allows comparisons of multiple traditional penalty cost policies to be performed in the same cost structure which can be used to estimate the appropriate penalty cost which is explained further in Section 2.7.

2.6. Sensitivity to misestimating β, λ

Determining an appropriate value for penalty cost is difficult for a manager. Similarly, it may be difficult to estimate the demand sensitivity parameters (β and λ) that are required in the perturbed demand model. The intensity factor (β) might be estimated through a series of consumer surveys, similar to what has been done with common consumer products (Walter and Grabner, 1975; Emmelhainz and Stock, 1991; Verbeke et al., 1998; Corsten et al., 2003; Campo et al., 2004; Van Woensel et al., 2007). The persistence factor (λ) can be estimated in terms of number of periods that must mass before a consumer forgets the disappointment of the stock-out experience. If we denote this number of periods by M, an approximate λ can be calculated using an adjusted formulation of Ma and Lu (2010) work to deal with small period quantities via

$$\lambda = \frac{2}{M+1}.$$

The MDP solution gain is not very sensitive to misestimating β or λ which means they do not need to be estimated with precisely. To measure the cost of misestimating β and λ, sixty-four of the randomly generated problems used in Section 2.5.3 are re-solved for a large number of β and λ combinations. Each problem was solved using $\beta = \{0, 0.1, 0.3,\ldots 1\}$.The persistence values are the values associated with a having a memory of one through fifteen days (M={1, 2, 3,…15}), and having all the weight on the initial demand deflation value (λ= 0). Setting λ= 0 is the same as not having demand deflation at all, this allows us to compare to a penalty cost model where π=r.

The results of this experiment set indicate that misestimating the demand sensitivity parameters is not that costly. The large variation of β and λ combinations resulted in different base-stock levels within a set of cost and demand parameters. For a given problem set varing β and λ may yield different base stock levels across the experiment set. The number of base stock levels varied from two to five, with the most common number of base-stock levels being three. For each set of cost parameters, we took the maximum percent change in gain from miss ordering. We assumed that a single base-stock policy is used, allowing us to see the difference in over- or under-ordering by a specific number of units. The measurement of over- and under-ordering is shown in

Table 2-6, i.e. what the manager should have ordered minus what was actually ordered. “Should” refers to the best single base-stock level which would result in the minimum gain had the true problem parameters been known exactly. This gain is compared to the policy that was implemented based on the estimated sensitivity parameters. Based on the estimated parameters the best base stock level possible will be refered to S_d, this policy will yield the smallest gain assuming the estimated parameters are accurate. The S_d policy is based on the estimated $\hat{\beta}$ and $\hat{\lambda}$ parameters, but the true β and λ parameters can take on alternative values. The error associated with miss estimating will result in a difference between the implanted policy S_d and the best policy possible if all parameters were known denoted as S_s. The difference in order policies results in a difference in gain which we term the error as given below.

$$Error = \frac{f(S_s) - f(S_d)}{f(S_s)}$$

In Table 2-5 over-ordering $(S_d > S_s)$ will have a negative value, and under-ordering will have a positive value. The maximum error for each parameter and deviation in order policy is generated at a specific β and λ combination. The shortest Euclidian distance from that point to a β and λ combination resulting in $S_s = S_d$ is displayed by the last two columns ($\Delta\beta = \beta - \hat{\beta}$ and $\Delta\lambda = \lambda - \hat{\lambda}$).

Table 2-6 Sensitivity to estimating β and λ

Should - Did	Count	Mean of Max Error %	Max Error %	Mean $\Delta\beta$	Mean $\Delta\lambda$
4	1	0.15%	0.15%	0.40	0.88
3	8	0.62%	0.92%	0.81	0.19
2	55	0.53%	1.14%	0.68	0.14
1	119	0.23%	0.84%	0.47	0.03
0	183	0.00%	0.00%	-	-
-1	119	0.22%	0.70%	(0.35)	(0.11)
-2	55	0.52%	1.40%	(0.51)	(0.16)
-3	8	0.55%	0.94%	(0.21)	(0.20)
-4	1	0.15%	0.15%	(0.90)	(0.78)

Table 2-6 shows that the error for misestimating β and λ is never greater than 1.4%, with most of the errors being less than 0.5%. These values are similar to over/under ordering by the same quantities in the lost sales penalty cost setting. This is even when managers misestimate β and λ by a significant quantity. The sensitivity parameters can only influence the order policy by small amounts, the fine tuning parameters. In the problem we studied, the amount (both in gain and order policy) was small across a range of sensitivity parameters. The cost parameters have a much larger impact on the order policy than β and λ.

2.7. Implied penalty cost

A single base-stock policy works well for most of the tested perturbed demand model problems. This policy structure matches the optimal policy for the lost sales penalty cost solution with instantaneous order arrival. In the penalty cost setting, the base-stock value is calculated using a simple critical fractile. This fractile solution is derived in Morton (1971) and is shown in the equation (5) below. By solving this equation for the penalty cost (π), it is possible to obtain the implied penalty cost for the perturbed demand model given the best single base-stock policy is known (equation 6). Here S^* represents the best single base-stock policy for the perturbed demand reality. The value of π resulting from equation (6) gives the closest approximation for the implied penalty cost in the perturbed demand reality for a given set of consumer response and cost parameters.

$$S^* = G^{-1}\left(\frac{\pi - c}{\pi + h - c}\right) \tag{6}$$

$$\pi = \frac{c + G\left(S^*\right)\left(h - c\right)}{1 - G\left(S^*\right)} \tag{7}$$

Due to the discrete nature of the demand distribution and the requirement of integer orders, a range of penalty cost values will yield in the same order policy. Equation (6) gives the lower bound penalty cost for a base-stock quantity of S^*. The $G\left(S^*\right)$ term forces the penalty cost to equal the lower bound because of the stair step structure of a discrete distribution's CDF. Using the value S^*+1 in place of

S^* in equation (6) will give the upper bound on π, non-inclusively. Thus, the range of penalty cost values that lead to the same base-stock level S^* are

$$\pi^* = \left[\frac{c + G(S^*)(h - c)}{1 - G(S^*)}, \frac{c + G(S^* + 1)(h - c)}{1 - G(S^* + 1)} \right)$$

The value of S^* is a function of the problem parameters and can be computationally intensive to determine. Determining S^* is difficult because the interactions between periods makes using analytical solution difficult. As an alternative a regression is developed to estimate S^* based on the problem parameters: c, r, h, β, and λ. The dependent variable in the regression is the average of $G(S^*)$ and $G(S^* + 1)$. Regressing on the fractile allows the solution to be independent of the demand distribution's shape. Using the average of the two bounds allows the regression to slightly over- or under-estimate and still result in the same order quantity. In this way, the same base-stock policy results for any fractile in the range $\left[G(S^*), G(S^* + 1) \right)$. The problems used to build the regression model are the same problems as shown in Section 2.5.3. The resulting regression equation is

$$G(\hat{S}) = 0.8193 + 0.1624 \ln(r) - 0.6144h - 0.0016\beta - 0.0221\lambda + 0.0856\beta\lambda + 0.2927h \ln(r)$$

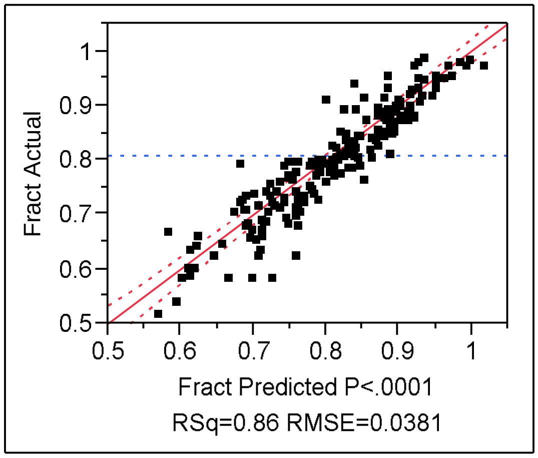


Figure 2-4 Fractile linear regression

The fit of the distribution for the test problems is high: R^2=0.86. Additional terms were considered but dropped because they did not have statistically significant coefficients. The regression's performance was tested using a hundred new randomly generated problems, constructed using the bounds in Table 2-1. The regression was used to determine an *S* value which is then evaluated using an MP to determine the gain. This gain was compared to the best single *S*'s corresponding gain for the one hundred test problems. The regression performance is measured as a percentage difference from the best single base-stock policy gain. The results are shown in Figure 2-5. The regression performed well for most of the sample problems because the estimated fractile resulted in a base stock near the best possible base-stock policy. Of the 100 problems, 86 resulted in a gain less than 0.5% from the best single base-stock policy gain. The max error is 4.43% which occurred for a problem with a relatively low margin (c=1, r=1.351, h= 0.297) and a short memory demand persistence value (λ=0.8234). Other problems for which the regression model did not perform had a high demand intensity factor (β) and a low demand persistence factor (λ).

	% from best single S
Min	0.00%
Mean	0.25%
Median	0.05%
Max	4.43%

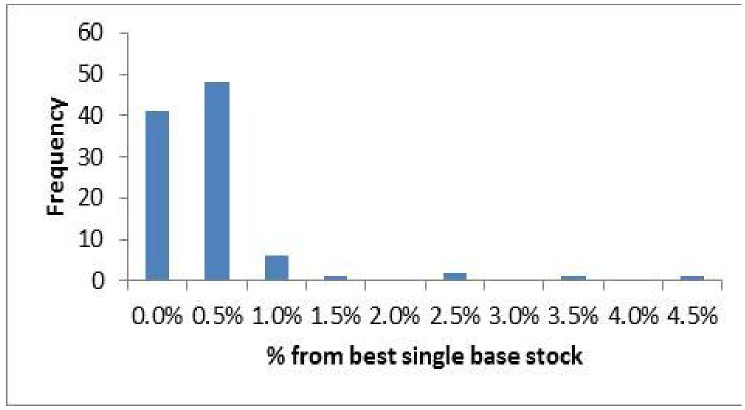


Figure 2-5 Regession performance

Looking at the implied penalty cost from the best single S solution, we see that the penalty cost is heavily influenced by the sales price (r). The value of goodwill is defined as the difference between the penalty cost and the sales price. The largest inflence of the goodwill cost is β. A factorial experiment design was used to test which parameters had the greatest influence on the value of goodwill. The expected value of the demand distribution was kept constant at 20 units but the p value ranged {0.2, 0.5, 0.8}. The cost parameters had levels r = {1.5, 2}, h = {0.05, 0.2, 0.4} and c held constant. The sensitivity factors related to stockouts was varied as well (β = {0.25, 0.5, 1} and λ = {0.35, 0.5, 0.75, 1}). The value of goodwill was calculated for each of these 216 problems. The value ranged from 0.07 to 1.56 with an average value of 0.33. The cost of goodwill was statistically different across levels of p, r and β at the 95% confidence level. The cost of goodwill increases as each of these parameters increased. Changing the demand distribution to have a higher p value pushes the demand's mode higher. Thus there are larger steps between each fractile, meaning that a larger goodwill cost is needed to increase to the next highest order quantity. Increasing the sales price (r) causes the numerator and denominator of the fractile to become larger, resulting in the goodwill to be larger to make the same size shift in the fractile compared to the lower cost problems. The parameter β determines the intensity of the lost sales; thus, increasing β increases the size of the shift in demand deflation. The larger shifts resulting from increasing β causes a larger drop in future demand, making missing sales more costly. The demand persistence factor (λ) had no significant impact on the goodwill cost. The total cost of a missed sale was not altered by changine λ. Altering λ changes how the demand loss is spread not on the magnitude, thus it has little effect on the goodwill cost. For a given β value, every non-zero λ will result in the same number of total lost sales. The infinite horizon and a constant value of money makes the timing of lower demand have no effect on the cost of a stock-out.

Assuming the perturbed demand model is reality, lost sales penalty cost models perform the best when the implied penalty cost is used. Misestimating π can lead to a poor approximation of the perturbed demand reality. In Chapter 4 we will show that for the finite horizon case, the implied penalty cost is not zero or infinity which was often the case in Schwartz's model shown by Liberopoulus et al. (2010).

In Section 2.6 we saw that there was a very limited range of base-stock policies that performed best for a given set of cost parameters, across a range of sensitivity parameters. Each base-stock level correlates to a range of penalty cost parameters. Based on the range of best performing base stock level across the combinations of β and λ, an implied penalty cost for each level can be obtained. The largest and smallest implied penalty costs can be used to bound the estimation of the penalty cost. Within this range the penalty cost model will perfrom well, the further from this range the estimated penalty cost is, the worst the policy will perform. On average, the range of penalty cost is between 1.06r and 1.75r. This is a relatively small range on average considering traditional penalty costs can take on any positive value. If a penalty cost outside of the range for a particular problem is used, the consumer behavior cannot be explained by our model. The use of penalty costs above the range implies that consumers are behaving too dramatically to model using our current formulation.

If we allow β>1 the model would be able to mimic the behavior of more dramatically reacting consumers.

2.8. Conclusion

The perturbed demand model developed and tested in this paper yields some interesting findings. Our model supports some assumptions made by Schwartz (1966), such as the policy converging to a base-stock policy. The optimal policy is a base-stock policy that is dependent on the demand deflation value which is more complicated than Schwartz assumed because of its dependence on the demand deflation.

This paper proved the structure of the optimal policy and that the policy will converge in the infinite horizon case. Three derivative approximations were developed which performed well under all the generated test problems. The approximation methods have the advantage of not requiring discretization of the demand deflation value and they are much faster than solving for the optimal policy. Additionally, the problem was solved forcing the same policy structure as traditional penalty cost models. For most tested problems this policy structure worked well compared to the optimal policy. Forcing the single base-stock policy structure allowed us to gain insight into the implied penalty cost. We saw that the revenue parameter had the greatest impact on the goodwill cost.

Perturbed demand research can be expanded to account for more realistic situations which are not considered in our model. Extension such as longer or stochastic lead times have been considered by penalty cost models, but not perturbed demand models. Penalty cost models have made numerous additional extensions that perturbed demand models have not considered, such as having backorders instead of lost sales. Further investigation into long-term consumer behavior should be a priority so that a more accurate model of how demand should be altered can be developed. As long as demand deflation monotonically decreases as a function of stock-outs, the optimal policy structure will likely remain constant.

In the infinite horizon setting, the perturbed demand model acts similarly to the traditional penalty cost model. The perturbed demand setting allows for an estimated penalty cost to be calculated. The cost of goodwill can be determined as a function of consumer behavior. The current literature on consumer behavior could be improved to give better estimates of how consumers behave in the face of a stock-out. In reality, stock-outs alter the future demand. Inventory models need to reflect this reality.

Chapter 3: Infinite Horizon Perturbed Demand Model with Known and Constant Lead Time

In this paper we consider a perturbed demand model where stock-outs reduce future demand. The reduction in future demand is driven by a demand deflation factor. This factor is updated at the end of each period based on parameters related to how consumers react to stock-outs, specifically the magnitude and persistence of their reactions. Our model extends the literature to include a positive reorder lead time. A Markov decision process model is formulated and the optimal solution is found using the value iteration procedure. In addition, a limited policy search approach is developed to reduce the solution time. In the infinite horizon, the policy converges to a steady-state solution. Due to the dependence on the demand deflation factor, however, the optimal policy does not have a simple policy structure. Much like the traditional lost sales penalty cost model, the optimal ordering policy is based on the expected state of the system at order arrivals. Due to the complexity of the policy structure, we develop an easily-implemented approximate solution based on a fractile of the demand distribution computed directly from the problem parameters.

3.1. Introduction

The United Sates' retail industry is becoming increasingly dependent on manufacturing from other counties. While the labor is cheaper, the downside is an increased lead time between retailers ordering a product and its arrival. Instantaneous order arrival is often assumed in inventory control models, but this does not always reflect real life. This paper is concerned with extending the perturbed demand models, a class of inventory models where stock-outs result in a loss of future demand. The inclusion of lead time is common in traditional penalty cost models, but perturbed demand models have not made this extension.

Penalty cost models factor in lead time by tracking outstanding orders as an added dimension in the state space, and this modeling approach is no different with perturbed demand models. The added dimension to the state space leads to increased solution time, often referred to as "the curse of dimensionality" (Bellman, 1957). Lost sales models with penalty costs do not have simple optimal order policy structures like their backorder equivalents (Arrow et al., 1958). The resulting complexity of the order policy does not allow an easy characterization of the policy. Perturbed demand models have an even more complex ordering policy due to the dependency of the order level on the current level of demand deflation.

This paper extends infinite horizon, non-seasonal, periodic review perturbed demand models to include non-instantaneous lead times. Cases with lead times of one period are the main focus of the paper, but problems with lead times of two periods are also solved in Section 3.6. The optimal policy looks forward in time and orders based on the expected state of the system when the order arrives, where the state of the system is defined by both the inventory and current level of demand deflation.

3.2. Literature Review

A traditional lost sales penalty cost model with positive resupply lead time was first introduced by Arrow, Karlin and Scarf (1958). They found that the optimal order policy for the lost sales case was not like its backorder counterpart; in other words, the optimal policy did not take on a traditional base stock policy structure. The optimal order policy has a constant order level for low on hand inventory levels then as on hand inventory increase the order quantity start to decrease. While the structure may not be the same, the policy still converges to a steady state policy in the long run (Iglehard, 1963).

Our model assumes perturbed demand, first formulated by Schwartz (1966, 1970), in which future demand is altered when customers are faced with a stock-out. Other models where demand is affected by ordering decisions include those by Hill (1976), Ernst and Powell (1995), and Diana and Petruzzi (2001). All of these models assume that demand is altered as a function of the ordering decision and not the actual stock-out events. Only Robinson (1990) uses the actual number of disappointed customers in each period to modify demand periodically. In this model the current number of satisfied and unsatisfied customers alters the underlying demand distribution seen by a retailer. In contrast, our underlying demand distribution is stationary, with underlying demand being modified by a scalar multiple whose value is based on the stock-out history and on two consumer behavior parameters.

The previous models dealt in the instantaneous order arrival case, the following literature deals with models with lead time. Due to the complex nature of the single period lead time optimal policy, many approximations have been developed to solve the lost sales penalty cost model. Morton (1971) created a myopic critical fractile procedure that works well for short lead times. Zipkin (2008b) provides a comprehensive overview of many approximations that have been developed to solve this problem. One common assumption is that the policy takes on a base stock policy structure because it is easy to implement. The total cost function for the penalty cost setting is convex in the base stock level *S*, making the search for the base stock level relatively simple (Janakiraman and Roundy 2004). The literature suggests that lost sales models with penalty costs have complex ordering policies that can easily be approximated by simpler base stock policies without a large increase in cost. For example, Huh et al. (2009) proved that as the penalty cost grows large in relation to other costs, a base stock policy asymptotically approaches the optimal solution. In the perturbed demand case, an approximation is developed to speed up the solution procedure. Additionally, as this paper will show, even if we force the instantaneous order arrival structure on the lead time model, it is not easy to solve for the best policy, nor does this always provide a close approximation. The advantage, however, is that is the policy is easier to implement once determined.

In Chapter 2, it was shown that the optimal order policy for a model assuming perturbed demand with instantaneous arrivals has a base stock policy with a different base stock level for each demand deflation level. This optimal order policy converges to a steady state solution. Other perturbed demand models have made the same assumption about instantaneous order arrival (Schwartz 1966, 1970, Robinson 1990 and Ernst, Powell 1995). In this chapter, we add a lead time to the order, develop a solution approach to this problem, and compare the results to those of traditional penalty cost models.

3.3. Model formulation

We develop a stochastic-demand, periodic-review model with perturbed demand—i.e., demand deflation as a result of previous lost sales due to supply shortfalls—as an alternative to imposing a penalty cost for lost sales. We assume that in the absence of stock-outs, the demand each period follows a stationary cumulative distribution $G(\xi)$, which we will refer to this as the *underlying demand distribution.* However, the consumer behavior literature has demonstrated empirically (Campo et al. 2004; Verbeke et al. 1998) that stock-outs affect the immediate future demand—meaning that some customers "abandon" on this product and/or retailer while others will return to purchase at a later date.

Although this analysis assumes an infinite horizon, for the purposes of notation, it will be helpful to consider use n to denote "periods-to-go" as we would in a finite-horizon setting. Consistent with other modeling work that deals with perturbed demand, we employ a *demand deflation factor* α_n, where n denotes periods-to-go This factor is computed based on the level of demand lost due to stock-outs in periods prior to period n. The timeline of events shown in Figure 3-1 is similar to traditional periodic review inventory models. At the start of the period, we evaluate the current state of the system, which consists of the current on-hand inventory (x_n) and the demand deflation factor from the previous period (α_n). Based on the state values, an order is made for z_n units, and the materials cost is imposed at the time the order is placed. Next, demand occurs for the period. Then holding cost and revenue are calculated based on the inventory remaining on-hand after fulfilling demand. Last, the order made at the start of the period arrives. Similar to Bellman's (1955) formulation, orders are assumed to arrive at the end of the period after holding cost is assessed—and in our case, after demand deflation is calculated as well—in order to eliminate the need for an extra state dimension. Our model makes the assumption that the lead time is only allowed to be an integer multiple of the period length to ensure that orders do not arrive mid-period.

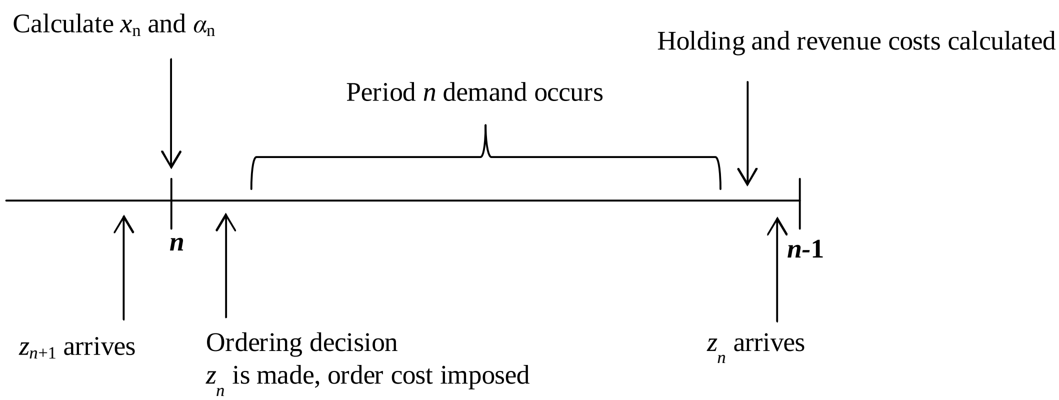


Figure 3-1: Timeline of events for lead time one case, starting with *n* periods to go

Thus, demand deflation in with $n-1$ periods remaining is based on actual demand fulfillment performance with *n* periods remaining; therefore, α_{n-1} is computed at the start of $n-1$, before demand occurs. The actual realized demand in with $n-1$ periods remaining is given by $\alpha_{n-1}\xi_{n-1}$ where ξ_{n-1} represents the realized sample from the demand distribution $G(\xi)$. Thus, if the expected underlying demand is $\mu = \int \left(1 - G(\xi)\right) d\xi$, the expected realized demand, given demand deflation factor is α_{n-1}, is $\alpha_{n-1}\mu$.

Our claim is that, while the demand deflation factor α_n is influenced by prior lost sales, this effect may be attenuated by other aspects of the behavior of the consumers in question. These aspects are the persistence of their memory of past stock-outs and the possibility that some customers are willing to effectively "shrug off" a stock-out such that it does not affect their likelihood of future purchase in any way. While the specific effect on demand has not been directly studied in the literature, we capture presumed consumer behavior in two model parameters: β, which we call the *deflation intensity*, reflecting the strength of the effect of stock-outs on future demand; and λ, which we call the *deflation persistence*, reflecting how long past stock-outs continue to affect future demand. Schwartz (1966) suggested that the effect of past disappointment will decline exponentially over time. Similarly, our demand deflation function with $n-1$ periods remaining depends on the incoming demand deflation factor (α_n); the realized sample from the demand distribution in period n (ξ_n); and the available inventory on-hand at the start of period n (x_n), as follows:

$$\alpha_{n-1} = T_\alpha(\alpha_n, \xi_n, x_n) = \begin{cases} \lambda + (1-\lambda)\alpha_n & x_n \geq \alpha_n \xi_n \\ \lambda\left(1 - \dfrac{(\alpha_n \xi_n - x_n)\beta}{\alpha_n \xi_n}\right) + (1-\lambda)\alpha_n & x_n < \alpha_n \xi_n. \end{cases} \quad (1)$$

Expression 1 shows that, for a prior period (n, preceding period n–1) in which no demand is lost (top portion of the bracket in (1)), α_{n-1} is computed by smoothing it with α_n, reflecting the "deflation persistence" from previous periods. Thus, if the persistence factor λ, is close to 1, the effect of previous losses will dissipate rapidly, whereas if it is close to 0, previous demand fulfillment performance will dominate the current performance. Also, from expression (1), we see that for a prior period in which there is a demand loss (bottom portion of the bracket), the first term includes the deflation intensity (β) effect. The impact of lost sales from the previous period ($\alpha_n \xi_n - x_n$) on future sales through α_{n-1} is attenuated by a multiple of β.

Note that the demand deflation factor is bounded between zero and one. This implies that demand cannot be negative, nor can it exceed the upper limit of the underlying demand distribution. It is assumed that the demand deflation is constant throughout a given period, meaning that stock-outs within the period will not further deflate demand in that period.

In this model, we assume that lead time is constant and equal to one; orders are made at the start of the period and arrive at the end of the period. Underlying demand is assumed to be stationary over time, but demand deflation, and consequently the realized demand, is dependent on past events.

A unit holding cost, *h*, is incurred for each item on-hand after demand occurs. For each unit sold, revenue (r) is received. Revenue is modeled as a negative cost since the problem is modeled as a standard cost minimization problem. The demand distribution has an underlying cdf $G(\xi)$, which is scaled by current demand deflation. Then, the immediate (one period) net cost given y units on-hand and demand deflation factor α, $L(y|\alpha)$, is

$$L(y|\alpha)=\int_0^{\frac{y}{\alpha}} h(y-\alpha\xi)g(\xi)d\xi-\int_0^{\frac{y}{\alpha}} r\alpha\xi g(\xi)d\xi-ry\left(1-G\left(\frac{y}{\alpha}\right)\right).$$

Assuming unit materials cost c ($c < r$) and given x_n units in inventory with a demand deflation factor α_n, the total cost over the next n periods is a recursive expression,

$$f_n(x_n,\alpha_n)=\min_{z_n>0}\left\{cz_n+L(x_n|\alpha_n)+\omega\int_0^{\infty} f_{n-1}\left((x_n-\alpha_n\xi)^+ +z_n,\alpha_{n+1}\right)g(\xi)d\xi\right\}, \tag{2}$$

where $g(\xi)$ is the underlying demand density function and ω is the one-period discount factor ($0 \le \omega \le 1$).

Additional lead time can be included in the cost function with a few changes being made. Specifically, outstanding orders must be tracked as part of the state definition and as an additional parameter in the cost function. If the lead time is two units, then the cost function is $f_n(x_n, z_{n+1}, \alpha_n)$, where z_{n+1} is the amount on order, to be received at the end of period n.

3.4. Calculation of the optimal policy

Similar to Chapter 2, the model can be formulated as a Markov decision process (MDP) and the optimal policy determined using the value iteration method (Howard, 1960). Since the demand deflation factor α appears in the discrete-valued state space of the MDP formulation, this continuous value must be expressed as a discrete value, which requires us to round its value using a specified rule. Using a uniform gird to approximate continuous variables is commonly done because of the simplicity this affords in solving stochastic control problems (Chow and Tsitsiklis 1988). Increasing the number of discrete intervals used to approximate α results in a more realistic representation of the problem. The realism, however, comes at the cost of a larger state space, thereby resulting in a longer solution time. Another argument for discretizing the demand deflation factor is that the retailer may actually not know the demand deflation factor with exact precision and would need to estimate it anyway.

The discretization of the demand deflation factor has an undesirable consequence, which is that it may contribute to the stochastic process no longer being ergodic. While the process is guaranteed to be ergodic if $\lambda > 0.5$ (as shown in Chapter 2 for the zero lead time case), for problems with smaller λ values (those for which the effects of previous demand losses are more persistent), it is possible that, once a significant reduction in the value of α_n occurs, realized demand may get "stuck" at a level persistently below its desired value. Larger gaps between the discrete values of α can exacerbate this effect; therefore, a more finely specified α lessens the likelihood of the non-ergodicity causing a problem. We assume that the demand deflation factor has an initial value of one, because we feel it is a sensible starting point for the start of the horizon, no demand is lost yet. This assumption ensures that the problem can obtain near one demand deflation if desired. The non-ergodicity occurs when trying to transition to states with demand deflation values close to one; thus, starting at a demand deflation value of one will ensure it is possible stay in these demand deflation state if you are willing to order enough to satisfy a high percentage of demand.

The optimal solution is computationally difficult to calculate because of the large state space, and the optimal policy is not easily characterized. The optimal policy is dependent on both dimensions of the state space with a one-period lead time. Computationally the value iteration method must be applied until the gain for each state converges. We stop value iterations when the following conditions are met: the sum of the absolute change in gain for all periods is less than some specified value, and none of the decisions have changed across periods. These conditions may take many iterations to converge, depending on the value of λ, the weight put on the most recent period. High λ values solve quickly, but when λ is low, the demand deflation factor tends to change slowly from period to period, meaning that a larger number of iterations are needed for the state probabilities to converge.

Recall from our earlier discussion that penalty cost models with lost sales do not always have nicely defined optimal order policies. While the backorder case results in a base-stock policy, the optimal policy of a lost-sales model does not take on an easily characterized policy structure (Nahmias, 1979). Similarly, when demand is perturbed as a result of stock-outs, the optimal order quantity is a function of the state the system expects to be in when the order arrives. The order made with n periods remaining is used to fulfill customer demand with $n-1$ periods remaining. We demonstrate below that the optimal order quantity may increase, level off, then decrease as a function of the beginning-of-period inventory as seen in Figure 3-2. The figure was generated from solving the optimal policy for eight problems. The cost and demand distribution were held constant at c = 1, r = 1.5, h = 0.2 and the demand was negative binomial (20, 0.5). The demand sensitivity parameters were varied with four levels of λ = {0.35, 0.5, 0.75, 1} and two levels of β = {0.25, 1}.When on-hand inventory is low, then stock-outs will likely occur, lowering demand in the future. This reduces the amount ordered because when the order arrives, demand will have already been lowered. The optimal order quantity increases as the on-hand inventory increases because fewer stock-outs are expected, resulting in less demand deflation. The order quantity continues to increase until ordering more units yields little or no improvement in expected demand deflation to start the next period. When on-hand inventory is large, inventory will be carried over to the next period; thus the order quantity can decrease because some units will already be available when the order arrives. As the inventory on hand increases, the policy tends take on a base stock policy structure. The base-stock order policy structure can be seen in Figure 3-2, when on hand inventory is above 35 units.

The steepness of the optimal order quantity increase, for the low inventory section of the policy, is dependent on the demand sensitivity parameters. Figure 3-2 shows the optimal policy for eight different combinations of β and λ. The high inventory side of the policy seems to be largely unchanged as a function of the sensitivity parameters. The low inventory side is highly dependent on the sensitivity parameters. As β increases, there is a sharp drop in the low inventory order quantity. High β cases will have a dramatic shift in demand deflation if stock-outs occur, and stock-outs will likely occur when inventory on-hand is low, reducing future demand.

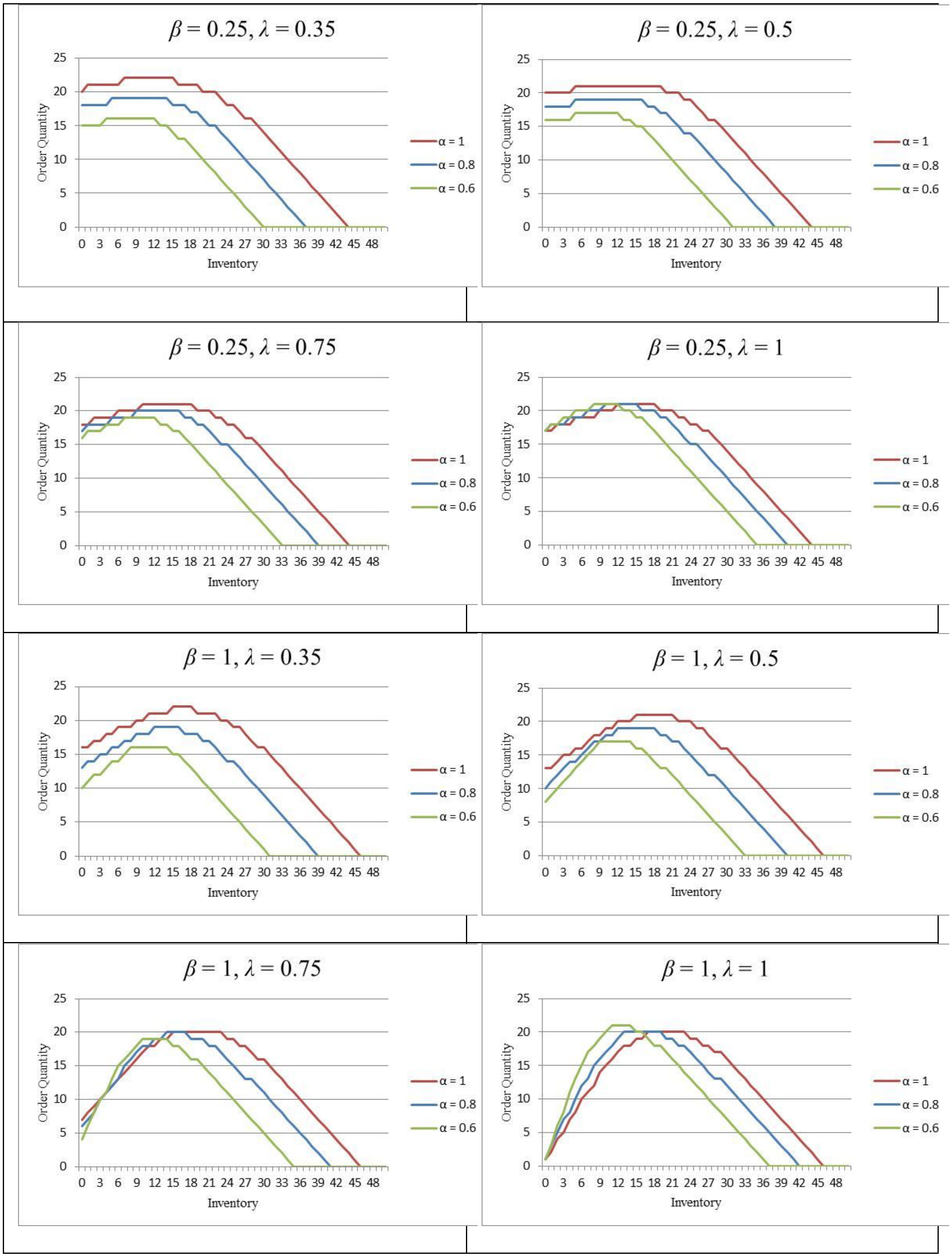


Figure 3-2 Optimal policy example Neg. Binomial(20, 0.5), c = 1, r = 1.5, h = 0.2

3.4.1. Unknown number of stock-outs

If stock-out occurs at a retail location for a given item, it is quite likely that the retailer will not have a way to determine exactly how many units of demand were lost. Many retailers only know the number of units of inventory on-hand at any given time. Once they have zero units on the shelf, the number of sales that are missed remains uncertain. It is impractical to interview all customers to determine if they were able to find what they wanted. Thus, in reality, the demand deflation will not be known with certainty. As discussed in Chapter 2, a point estimate can be used to take the place of knowing the actual demand deflation.

A conditional expectation can be used to estimate the number of stock-outs. When there is inventory left on the shelf at the end of the period, then there is no need to estimate the number of stock-outs. Otherwise, the expected number of stock-outs is estimated based on the conditional demand distribution. In Chapter 2, we define underage with n periods remaining as $U_n = \left(\alpha_n \xi_n - S_n\right)^+$. Thus, we can define $\hat{U}_n = E\left[U_n \mid \hat{\alpha}_n \xi_n > S_n\right]$ and compute it as follows

$$\hat{U}_n = \frac{\int_{\frac{S_n}{\hat{\alpha}_n}}^{\infty} \left(\hat{\alpha}_n \xi_n - S_n\right) g\left(\xi_n\right) d\xi}{\int_{\frac{S_n}{\hat{\alpha}_n}}^{\infty} g\left(\xi_n\right) d\xi}$$

Then the demand deflation can be updated using the estimated underage as follows

$$\hat{\alpha}_{n-1} = \lambda \left(1 - \frac{\hat{U}_n \beta}{\hat{\alpha}_n \xi_n}\right) + \left(1-\lambda\right) \hat{\alpha}_n .$$

The order quantity is then based on the estimate demand deflation and current inventory.

It is assumed that demand deflation is known at the beginning of the horizon, $\hat{\alpha}_{\infty} = 1$. Thus to estimate the demand deflation with *n*-1 periods remaining we can use most recent estimation $\hat{\alpha}_n$. Demand is known to be at least *S* units when there is a stock-out; therefore, only the part of the demand distribution above *S* matters when calculating the expected number of stock-outs. The estimation of α will be used in the expected stock-outs formula because the true demand deflation is unknown. The expected number of stock-outs is used to estimate demand deflation. Just like actual demand deflation, the estimated demand deflation will change over time as a function of stock-outs. The difference is that the exact number of stock-outs is not known.

A full factorial experimental design is used to test how the cost of ordering based on the point estimate of the demand deflation value. A finely discretized demand deflation factor is used to test the fractile versus the MDP value iteration procedure.

Table 3-1: Experiment design

Parameter	Levels
c	1
r	1.5, 2, 10, 40
h	0.05, 0.2, 0.4
β	0.25, 0.5, 1
λ	0.35, 0.5, 0.75, 1
Underlying Demand Distribution, $G(\xi)$	Neg Binomial (20, 0.5)

To test both procedures in a common setting, we generate the approximation-based policy and then use as a fixed policy in a Markov process to determine the gain of the approximation policy. For the factorial design described above, use of the estimated demand deflation value can be compared to the actual demand deflation. The gain for the known demand deflation MDP is compared against the gain achieved when using the estimated demand deflation. In the MDP setting the demand deflation factor must take on discrete values where the difference between successive values is denoted as Δ. Small Δ values more accurately represent a continuous variable, but increase the size of the state space. The added state dimension, estimated demand deflation, makes using a very fine grid infeasible; therefore, for all tested problems, Δ= 0.02 is used in place of the smaller Δ=0.005 used later. Using a point estimate of demand deflation performed well in our test set. The results summarized in

Table 3-2 indicate that the loss in value from having imperfect stock-out information is relativity low. Specifically, the maximum difference in gain across all problems as a result of using an estimated demand deflation compared to the actual is less than 0.35%.

Table 3-2 Error in gain when using estimated demand deflation

% difference	Min	Mean	Max
Error from using estimated number of stock-outs	0.000%	0.023%	0.334%

3.4.1.1. Simulated unknown number of stock-outs

The downside of using the MDP approach is the large state space required to track both real and estimated demand deflation, which does not allow the finest Δ to be used. Simulation can be used as an alternative to the MDP approach to deal with large state spaces. In a simulation, there is no need to round demand deflation between periods for the purpose of estimating realized demand. However, we only know the optimal policy to Δ = 0.005; thus, demand deflation must be rounded in order to determine the order quantity. This is a much finer grid than could be used when both real and estimated demand deflation need to be tracked.

The simulation is implemented over 200 periods and 100 replications for each case described in Table 3-1. A 200-period horizon is used because we observed through experimentation that the average cost per period has steadied out, and 100 replications are used to get a more accurate estimate of the mean cost. Observing 100 replications brings the standard deviation down to 3% of the total cost on average. Both the known and estimated demand deflation simulations are run in parallel under the same randomly sampled demand sequence. The difference in total cost is the metric we use to measure the accuracy of the estimated demand deflation. A small absolute cost difference indicates that estimating demand deflation does not significantly impact the overall cost of the system. The same test cases used in section 3.4.1 are used in this section. The results are summarized by the blue bars in Figure 3-3, where the percentage difference is used to measures the difference between the known and unknown demand deflation values. Negative values indicate that the estimated demand deflation simulation outperformed the known demand deflation simulation on average. Small negative values were seen in the simulation but the confidence intervals around all the instances contained zero at the 95% confidence level. None of the cases results in large average differences in cost, and none of them are statistically different from zero at the 95% confidence level. This data supports the idea that estimating the demand deflation works well for the infinite horizon setting.

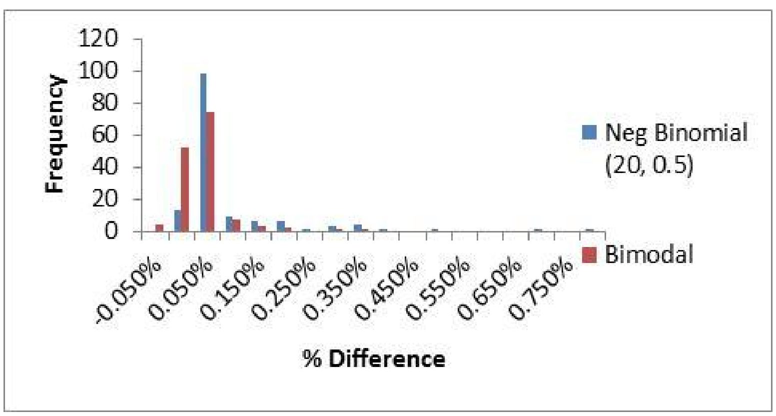


Figure 3-3 Simulation percent difference

The negative binomial distribution (20, 0.5) is approximately normally distributed, which may account for why the errors are small. To test if the distribution's shape has an effect on the error when estimating the demand deflation value, the simulations are redone using a bimodal underlying demand distribution. The results of the new simulation are displayed in Figure 3-3 by the red bars. The errors for the new simulations with the bimodal distribution are distributed similarly to the original distribution as seen in Figure 3-3. Bimodal demand does not change the fact that none of the differences in costs between the estimated and known demand deflation cases were statistically significant.

Estimating the demand deflation works well over the long run because stock-outs are not very common. The largest differences between the estimated and actual demand deflation value occurs right after a stock-out occurs. The periods between stock-outs allow both actual and estimated values to approach one before another stock-out occurs. For each period *n*, in the simulation, the differences between the actual (α_n) and estimated $(\hat{\alpha}_n)$ demand deflation values were tracked as γ_n.

$$\gamma_n = \left|\alpha_n - \hat{\alpha}_n\right|$$

The number of times that the two factors jump from having a difference less than 0.02 to having a difference of greater than 0.02 (n_γ) was tracked. A threshold of 0.02 was chosen because it provides a balance between signaling when there is a difference in demand deflation value and signaling when the change in demand deflation actually results in a change in the order quantity. We define $\overline{m}_\gamma$ as the average number of periods after signaling a difference that it took to close this difference in demand deflation values to below 0.005. This value was tracked because it gives insight into how long the deviation lasted. From the experiment set we saw that large differences occurred as a result of three variables: r, h, and β. In low profit cases (high h and low r), the retailer was willing to stock-out more often due to the low profit margin. For cases with a high value of β, n_γ was larger because each stock-out caused a larger change in deflation, the estimation procedure is more important because we are multiplying by a larger number (β). The variable $\overline{m}_{\gamma}$ was influenced by λ. For low values of λ it took longer for the two demand deflation terms to converge. Putting more weight on the most recent events allows the system to correct itself faster when no stock-outs occur. Thus, the cost of not knowing the exact demand deflation is low when β is low or λ is high. Even for the worst cases, the simulated cost differences were low; we therefore assume that demand deflation is known with certainty for the remainder of the paper.

3.4.2. State reduction method

The optimal policy is time consuming to determine using standard MDP solution procedures because of the large number of states, thus we used alternative methods that sped up the calculation of the optimal policy. The state space reduction method developed by Ding et al. (1988) is implemented to deal with large state space problems. This method relies on the limiting behavior of the Markov process having only a few recurrent states that are dependent on the order policy. Most of the states of a MDP are transient, with only a handful of recurrent states that matter. By eliminating transient states, the state space is reduced making the problem manageable. Ding et al. (1988) also show that the recurrent set changes with every policy update, but only a small number of states need to be considered to evaluate the policy's gain. The state space is reduced to union of three sets: *R* the recurrent set; *A* the neighboring states; and *C*, the states reachable from *A*. With the policy improvement procedure, the reward function of the MDP becomes

$$\min_{k \in K_i^{'}} \left[q_i^k + \sum_j p_{ij}^k v_j(n) \right] \quad i, j \in R \cup A \cup C .$$

$K_i^{'}$ denotes the set of alternatives for state *i* that can transition to states in $R \cup A \cup C$. For a two period lead time problem with $\Delta = 0.05$, the union of *R*, *A*, and *C* make up about 5% of the total state space. The reduction of the state space makes solving very large state space problems possible.

The reduction in states makes calculating the gain for a given policy much faster, but this state reduction method also reduces the ability of the system to transition to all the states, potentially causing the procedure to converge to a local optimum and not to the globally optimal solution. The algorithm is said to converge when only a small number of decisions change across iterations: change in gain is below a set threshold, and the change in $v_i(n)$ for all i is below a threshold. However, rounding issues may cause two decisions to result in the same gain. Issues with rounding cause states to change decisions across periods but have no noticeable effect on the gain. For all the problems tested with lead time one, a policy with the same gain was found using both the traditional method and this state reduction method. The advantage comes in terms of the speed increase, which is illustrated in Figure 3-4. The example problem used to create the figure had an underlying demand distribution of negative binomial (20, 0.5), with cost parameters c = 1, r = 1.5, h = 0.05, and demand deflation parameters β = 0.5, and λ= 0.75. A smaller grid size was used to create a larger state space while using the same parameters. The restricted decision method described in the previous section made a significant reduction in solution time, but the state space reduction technique outperforms it in this example.

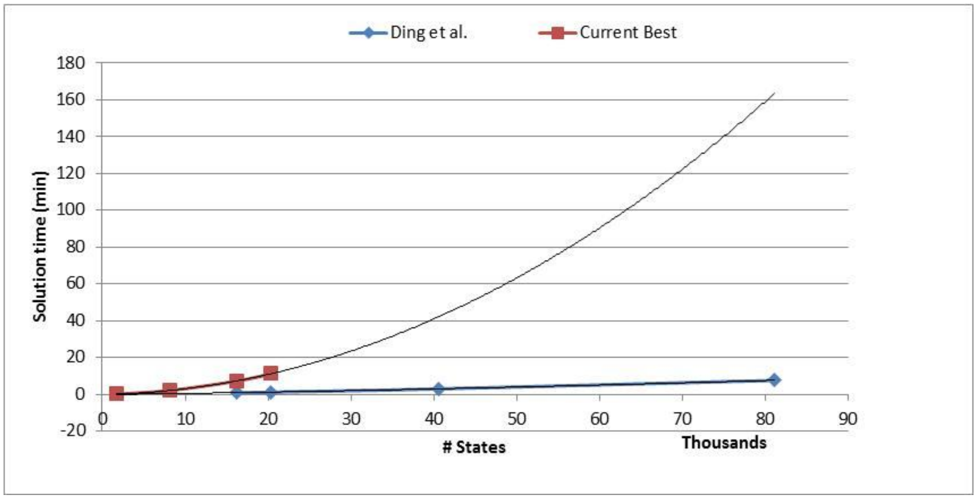


Figure 3-4 Speed comparison restricted decision versus Ding et al. algorithm

The improvement becomes more significant as the number of states grows. This is important because the state space grows by one dimension for each additional period of lead time beyond a lead time of one period. The curse of dimensionality causes the state space to grow exponentially with lead time making the state reduction procedure the only way to solve problems with lead time greater than one period.

3.5. Policy heuristics

Due to the computational difficulties involved in calculating the optimal policy, heuristics were developed which enable us to solve problems which are computationally difficult. The optimal policy is not only difficult to determine, but it is also difficult to implement because it does not take on a nice structure. The first heuristic is based on determining a fast heuristic by implementing a critical fractile approach. The second and third approaches make implementing the policy easier by restricting the policy structure to take on a simpler form.

3.5.1. Fractile heuristic solution

The optimal policy is difficult to implement in practice because it does not have an easy-to-characterize policy structure. It requires the use of lookup tables, which can become cumbersome as the number of states increases. The lookup table is a matrix which uses the demand deflation value and inventory levels as keys to find the appropriate order quantity. Thus, we propose an alternative solution based on a modification of the simple newsvendor approach. Like a newsvendor solution, our heuristic bases the current-period order on a critical fractile that balances overage and underage costs. The procedure is described below.

For a given state (for example, an inventory and demand deflation factor pair (x,α)),the order quantity is determined as follows:

Step 1) Estimate the value of the demand deflation factor $\hat{\alpha}$ for the following period by assuming that the current period demand equals the mean demand adjusted by the current demand deflation factor α (for example, assume demand is $\mu_\alpha = \alpha\mu$). Thus, the update formula yields

$$\hat{\alpha}_n = \begin{cases} \lambda + (1-\lambda)\hat{\alpha}_n & x \geq \mu_\alpha \\ \lambda\left(1 - \dfrac{(\mu_\alpha - x)\beta}{\mu_\alpha}\right) + (1-\lambda)\hat{\alpha}_n & x < \mu_\alpha \end{cases}$$

Step 2) Calculate the order-up-to quantity, *S*, using a critical fractile based on the optimal solution of the zero-lead-time lost sales model (Morton 1971), but with revenue replacing the penalty cost. Then multiply by the estimated demand deflation factor, specifically

$S = G^{-1}\left(\dfrac{r-c}{r-c+h}\right)\hat{\alpha}$. Chapter 2 showed that revenue is a close approximation to the penalty cost.

Step 3) Order an amount z based on the expected on-hand inventory at the end of the period:

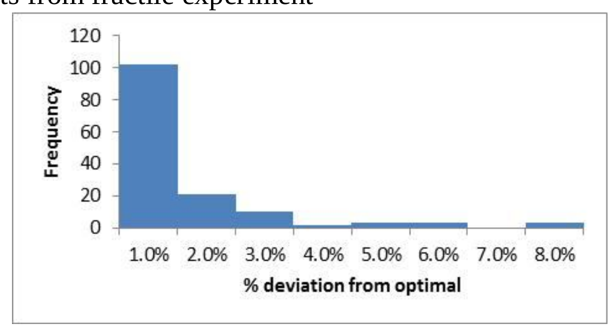
$$z = S - \left(x - \mu_{\alpha}\right)^{+}.$$

The procedure outlined above has a few advantages over the MDP solution. First, the MDP table lookup process is replaced by the simple calculations above. Second, there is no longer a need to discretize the demand deflation factor, thereby eliminating the issues resulting from rounding. Lastly, the solution procedure is straightforward and intuitive.

Table 3-3 describes the performance of the fractile solution using the experiment set described in Table 3-1. On average, it performed well with a median deviation from optimal of 0.52%. However, in a few cases, the deviation exceeded 1% with a maximum of 7.84%. The worst-performing cases were those with the lowest revenue values coupled with the highest holding cost. In these cases there is a very low margin; thus, even orders that are a few units from optimal result in a large change in performance.

Table 3-3: Results from fractile experiment

	% deviation from optimal
Min	0.12%
Mean	1.05%
Median	0.52%
Max	7.84%



3.5.2. Multiple base stock policy

Traditional lost sales penalty cost models with order lead time, are commonly approximated using their instantaneous order policy structure (Hue et al., 2009). The optimal policy for lead-time greater than one models does not have an easily implemented policy structure. Implementing the optimal policy would require that the order quantity be looked up from a table of state values, which can become large. The demand deflation dependent base-stock level policy (multi-*S* base-stock) has an advantage of being much easier to implement compared to the optimal policy.

Finding the best demand deflation dependent base stock level policy is not easily done. The total number of possible base stock policy combinations equals the number demand deflation levels raised to the number of possible decision. This quantity grows quickly when there are many possible order decisions. For our examples, there would be ~11^{45} (7.29 x 10^{46}) possible order combinations, assuming Δ=0.1. This number can be reduced by assuming that the order quantity will decrease as the demand deflation decreases; however, the reduction is not enough to allow for enumeration of all policies.

A modified two phase Howard (1960) approach is used to find the best policy with the policy structure constraint. All the states with the same demand deflation are grouped together and a common base stock decision is made for the group instead of individual states. Davis et al. (2009) did similar groupings of states and made the decision based on what is best for the entire group of states—not just the individual state. Ordering based on restricted information changes the problem from an MDP to a restricted observation Markov decision process (ROMDP). The importance of a state within the group is weighted based on its likelihood.

The ROMDP solution procedure needs both starting and ending conditions specified before implementation can begin. The state probabilities need initial values, so for this case, we assume that all states with a demand deflation of one have an equal probability whereas all other states have a probability of zero. This assumes that the demand deflation to start the system is one, meaning that no customers are disappointed yet. The initial policy is generated from the fractile solution procedure. The initial solution does not have the desired structure, but it will allow steady state probabilities and

v_i values to be calculated. The procedure will terminate when the policy is the same for three cycles in a row and the sum of absolute difference in gains falls below a predetermined tolerance value. If these conditions are not satisfied after one hundred iterations, then the solution procedure is terminated.

The traditional two phase Howard (1960) method uses repeated cycles of value determination and policy improvement routines. Figure 3-5 demonstrates how the procedure works. The value determination step must be modified to account for problem instances where the gains may not always converge to the same value. The value determination operation does three things: it calculates new v_i values, calculates the gain, and tracks the state probabilities. Fixed policy value iterations will be used in place of value determination. Value iterations are done implementing the policy provided by the policy improvement routine; and the transition probabilities are tracked in order to update the state probabilities. Value iterations continue until the probabilities and gains have converged or two hundred iterations have been completed, whichever comes first. With the new probabilities and cost values calculated, the procedure can move to policy improvement techniques. Using methods developed in Davis et al (2009), all states with the same demand deflation level are grouped (state space is segmented into sets S_j) together and one base stock decision will be made for that whole group. Within the group states are weighted based on state probabilities. For each level of demand deflation a best base-stock level for the group is determined. If there is no change in the policy from the last iteration, then the policy has likely converged; however, to ensure that nothing changes, one additional cycle is completed to ensure convergence.

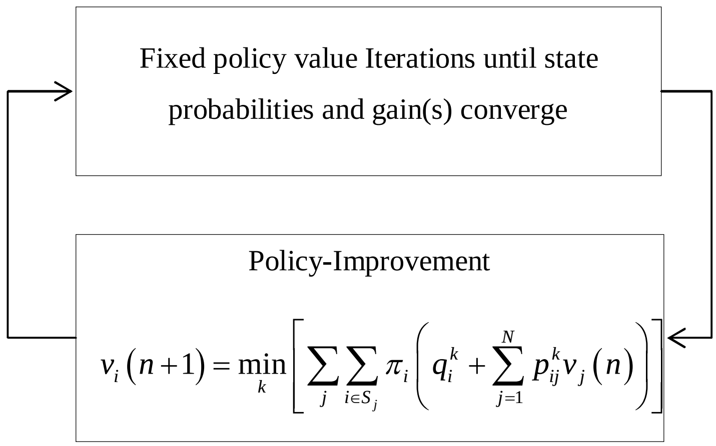


Figure 3-5 Modified two phase Howard solution to demand deflation dependent base stock policy

Grouping the states does not guarantee that a global minimum has been found (Davis et al. 2009). After a multi-S base stock policy is found, a π perturbation can be used to slightly alter the state probabilities, which will then alter the policy (Wei, 2005). The state probabilities will be modified in an attempt to find an alternative minimum. The π perturbation procedure adds a random value [0, 1] to each state. The state probabilities are then rescaled, state probability divided by sum of state probabilities, to ensure the state probabilities sum to one. Each state has a random value between zero and one added to it and then the probabilities will be rescaled. This procedure worked well for Davis et al. (2005) to find the global optimum. The procedure will recalculate a new policy based on the altered state probabilities. State probability perturbations and new policies will be generated until a worse gain is found (Wei, 2005) in an attempt to avoid a non-optimal local minimum; a good initial policy decreases the number of non-optimal solutions found.

The Figure 3-6 summarizes the results from the factorial design with a demand distribution of negative binomial (20, 0.5). The same factorial experiment set was used in this section as was used earlier and described in Table 3-1. As revenue increases relative to the other costs, forcing a multi base stock policy becomes a better approximation to the optimal policy. This finding is similar to that of Huh et al. (2009) in that as penalty cost increases relative to other holding costs, order-up-to policies become closer to optimal. There is a large drop in deviation from optimal when the revenue increases from 1.5 to 2. Anything with revenue of two or greater has a max deviation from optimal of less than 2%. The solution time is not faster than the optimal solution procedure; the advantage of this policy procedure is its simplicity to implement.

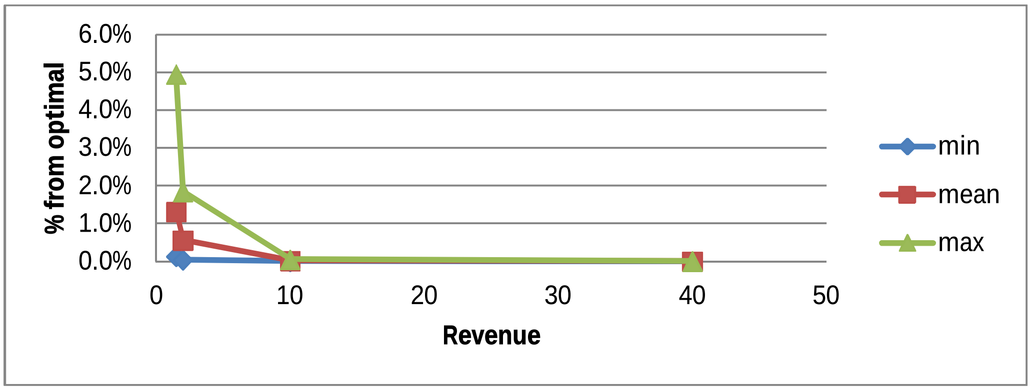


Figure 3-6 Multi-S base stock accuracy increase as revenue increases (Δ=0.1)

The demand deflation dependent base stock policy was tested using various values of Δ, all evaluated in a common reality. Regardless of the Δ size used to determine the policy each case is evaluated in the smallest Δ MDP to ensure a fair way to compare polices. Using a common Δ across policies we can demonstrate the value of using more demand deflation values. The policy was constructed using different numbers of demand deflation levels, each with its own unique base stock value. Each base case was solved using three different demand deflation granularity sizes. As more base stock levels are used there is a reduction in the maximum error. The mean error was largely unchanged as a function of the number of base stock levels. As noted earlier, this procedure does not always find the best demand deflation dependent base stock policy. As the number of demand deflation levels increases, it is more likely a local minimum base stock policy was generated because of the additional combinations of solutions. Some of the error associated with the 21 base stock levels could be due to the best solution not being found. The results of the factorial design are summarized in Table 3-4.

Table 3-4 Comparing various multi *S* base stock policies

	5 base stock values	11 base stock values	21 base stock values
Min	0.000%	0.000%	0.000%
Mean	1.036%	1.028%	1.145%
Median	0.077%	0.163%	0.533%
Max	14.076%	6.874%	4.619%

The demand deflation dependent base stock policy works better as a finer grid is used; however, problems using finer grids take longer to solve. Additionally at the smallest grid used the max error was still high 4.619%. The demand deflation base stock level policy works well in general but the largest error resulted from instances with low *r* values.

3.5.3. Morton's penalty cost heuristic

Morton (1971) developed a myopic heuristic for the lost sales inventory model with order lead time for the penalty cost setting. The heuristic was developed because the optimal policy structure does not take on base stock policy structure, yet there is still a need for an easy-to-calculate and easy-to-implement policy. Morton's (1971) heuristic uses a critical fractile approach, balancing overage and underage costs. In the experiment set designed by Morton, the heuristic had a max deviation from optimal of less than 1% for the lead time one test problems. Morton's heuristic requires a penalty cost, which perturbed demand models do not have. Earlier work from the instantaneous order case suggests that using the best penalty cost policy as a heuristic to the perturbed demand model works well (Chapter 2). A penalty cost (π) estimation method has been developed earlier in Section 2.7 for the instantaneous order perturbed demand setting. Applying this estimated penalty cost in Morton's heuristic allows a fast and easy to implement policy.

Morton's heuristic makes use of the structure of the optimal policy for the lead time one problem assuming lost sales.. The optimal order policy is level and then starts to decrease, similar to a base-stock policy, as the inventory on-hand increases. Consequently, Morton suggested that the heuristic should take on the following order quantity $z_n = \left\{\min\left(\tilde{z}, \tilde{S} - x_n\right)\right\}^+$ where $\tilde{z}$ and $\tilde{S}$ are based on the fractile below. The term G_2^{-1} denotes the convolution of two periods of demand from the underlying demand distribution. The formulation for $\tilde{z}$ and $\tilde{S}$ are shown below.

$$\tilde{z} = G^{-1}\left(\frac{\pi - c}{\pi - c + h}\right)$$

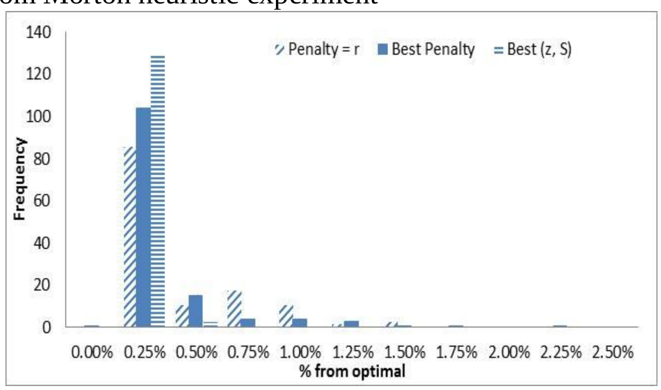
$$\tilde{S} = G_2^{-1}\left(\frac{\pi - c}{\pi - c + h}\right)$$

Using the policy structure provided above, the policy can be compared to the optimal demand deflation-based policy. Three different approaches were used to determine policies that follow this structure. The first is a simple heuristic in which the revenue (r) is used as the penalty cost (π). The second enumerates all the policies that can be generated using any π value ranging from 1 to 400. All the approaches are evaluated in the demand deflation setting. This method allows us to determine which penalty cost gives the policy which performs the best in the perturbed demand reality. The third method uses a gradient searches, with the $\pi = r$ policy as the starting point and a neighborhood size of one. This allows the policy to take on ($\tilde{z}$, S) combinations which are not possible using any penalty cost approach. Using a penalty cost forces the same fractile to be used for both $\tilde{z}$ and $\tilde{S}$, but as shown, this may not lead to the best policy.

Each policy was tested using the factorial design outlined in Table 3-1. All three policies perform well with the test problems; specifically, all of them have lower max deviations than the fractile solution discussed earlier. Using the revenue as the penalty cost is a fast and easy procedure to implement, but it does have the highest mean error at 0.31%. The best penalty cost works well for most problems with an average deviation of 0.20%, but it takes much longer to determine due to the need to search every possible policy which can be generated using a penalty cost. Allowing the policy to take on any ($\tilde{z}$, S) form performs the best because it is the least restrictive, resulting in a mean error of 0.06%. A summary of the experiment set can be found in Table 3-5. The Morton heuristic structure performs poorly on the same problems as does the fractile solution; problems with low profit margins and highly sensitive consumers are difficult to approximate because of the dramatic changes in demand across periods.

Table 3-5 Results from Morton heuristic experiment

% from optimal	$\pi = r$ policy	Best π policy	Best ($\tilde{z}, \tilde{S}$) policy
Min	0.00%	0.00%	0.00%
Mean	0.31%	0.20%	0.06%
Median	0.08%	0.04%	0.02%
Max	2.19%	2.19%	0.62%



The performance of Morton's heuristic suggests that penalty cost models can be a good approximation to a perturbed demand reality for lead time problems. Better estimates of penalty cost will result in a better performing gain as seen in the improvement of the best π policy compared to the π= r policy . For some problem sets, specifically low margin and highly sensitive demand cases (high β and λ), even the best penalty cost policy has a large error (>2%). In these cases, the demand deflation changes rapidly enough to make implementing an order policy perform poorly unless it is allows to change the order quantity based on demand deflation.

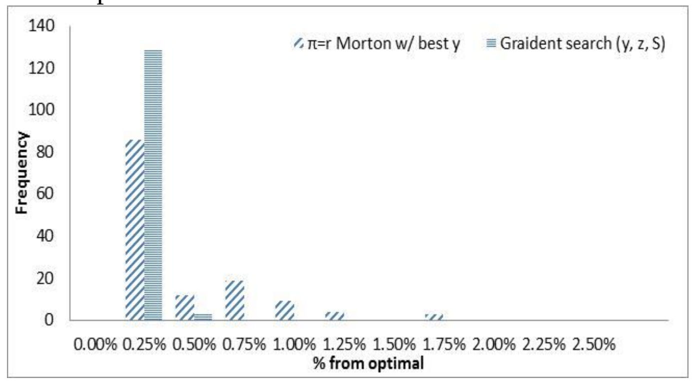
While the optimal policy for the lost sales case looks like the Morton (1971) heuristic, the perturbed demand optimal solution has an initial upward trend in the order policy when the on-hand inventory is low. An initial upward sloping line added to the Morton structure makes the heuristic take on a shape that is closer to optimal. We assume that the upward slope of the line is positive one because it lies near the middle of the range of possible slopes which occur across different values of β and λ as seen in Figure 3-2. Additionally, low inventory states occur infrequently which results in less importance being placed on the accuracy of the order quantity. The intercept of the line is noted as y, where $0 \leq y \leq \tilde{z}$.

$$z_n = \left\{ \min\left(y + x_n, \tilde{z}, \tilde{S} - x_n \right) \right\}^+$$

Two heuristics will be considered with this extension. The first takes the $\pi = r$ policy then finds the best y by enumerating all integers between zero and $\tilde{z}$. The second starts with the $\pi = r$ Morton policy with $y = \tilde{z}$ then gradient searches from there to find a local minimum. Each policy parameter (y, $\tilde{z}, \tilde{S}$) is allowed to be altered by plus or minus one for each iteration of the gradient search. The experiment set is resolved so that the performance improvement can be measured by including the third policy parameter. A summary of the performance from optimal is displayed in Table 3-6.

Table 3-6 Percentage from optimal for three term heuristic

% from optimal	$\pi = r$ policy with best y	Best $(y,\ \tilde{z}, \tilde{S})$ policy
Min	0.00%	0.00%
Mean	0.29%	0.06%
Median	0.08%	0.02%
Max	1.59%	0.62%



Adding in the best y term in the $\pi = r$ policy results in a significant improvement in the maximum error by very small changes to the mean or median. The y term helps the most in problems where stock-outs are more frequent—e.g., problems where revenue is low relative to the other costs. For problems where stock-outs are rare, then there is little benefit from adding in the y term because most of the states that use it are unlikely to occur. While there is some improvement for using the three term gradient search compared to the two term gradient search, the improvement is small and our study revealed that it did not improve the maximum deviation from optimal to within two significant digits. A frequent occurrence in problems with high r terms is that $y = \tilde{z}$, which indicates that there was no improvement for lowing the y value; also, it means that the two term policy worked just as well as the three term policy. Additional improvement may be seen by allowing the slope of the initial positive-sloped section to vary across problem parameters, but this would involve searching in four dimensions, dramatically increasing the search time.

The Morton heuristic method can be applied to problems with longer lead time by using a formula with only slight modifications. The calculation of $\tilde{z}$ remains the same, but $\tilde{S}$ changes as a function of lead time. The equation for $\tilde{S}$ becomes $\tilde{S} = G_{(v+1)}^{-1}\left(\dfrac{\pi - c}{\pi - c + h}\right)$ where v is the lead time in periods. Additionally, the order quantity calculation needs to account for units on order, such that $z_n = \left\{ \min\left(\tilde{z}, \tilde{S} - x_n - orders \right) \right\}^+$. Morton (1971) showed for the penalty cost case, the error increased as a function of lead time but was still fairly low for lead time of two periods. Zipkin (2008) showed that the error for this type of policy keeps increasing through four periods of lead time, with a maximum error of 7.1% for a Poisson distribution.

3.5.4. The cost of ignoring β, λ

In the context of our demand deflation model, a naïve Morton policy of $\pi = r$ assumes $\beta = 0$, which means that there is no future demand loss as a result of stock-outs. Managers unable to estimate β or λ may fall back on this policy because it relies only on available information. This policy was tested earlier for problems with $\beta \le 1$, and we will now relax this assumption to allow $\beta > 1$. Values of $\beta > 1$ assume that more consumers than simply those who are faced with a stock-out will take their business elsewhere. This is a word of mouth effect; consumers faced with a stock-out not only do not return in the immediate future, but they also convince additional people to change their ordering behavior. Table 3-7 outlines the experiment set used to determine the effects of ignoring the demand sensitivity parameters.

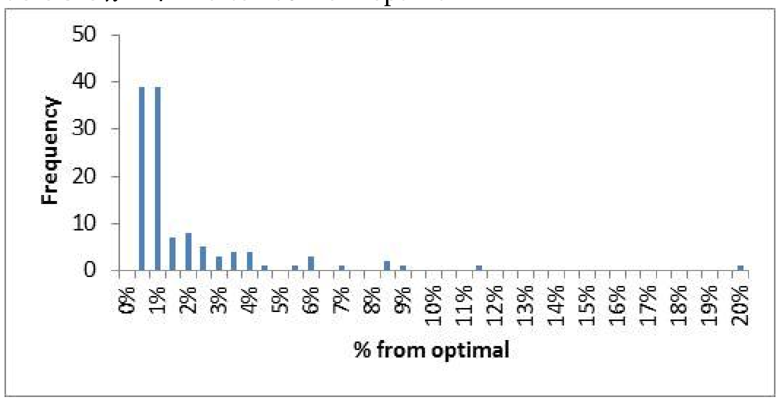
Table 3-7 Experiment design high β

Parameter	Levels
c	1
r	1.5, 2
h	0.05, 0.2, 0.4
β	0.25, 0.5, 1, 2, 4
λ	0.35, 0.5, 0.75, 1
Underlying Demand Distribution, $G(\xi)$	Neg Binomial (20, 0.5)

For each problem in the experiment set, we compute the optimal gain that results from a Morton approximation with $\pi = r$. Most of the time, the percentage deviation from optimal is low with a median value of 0.75%. The extremely poor-performing cases pull the average up to 1.64%. Table 3-8 summarizes the results of the experiment set. The extreme points are all cases where $\beta \geq 2$.

Table 3-8 $\pi = r$ Morton % from optimal

$\pi = r$ Morton % from optimal	
Min	0.03%
Mean	1.64%
Median	0.75%
Max	19.73%

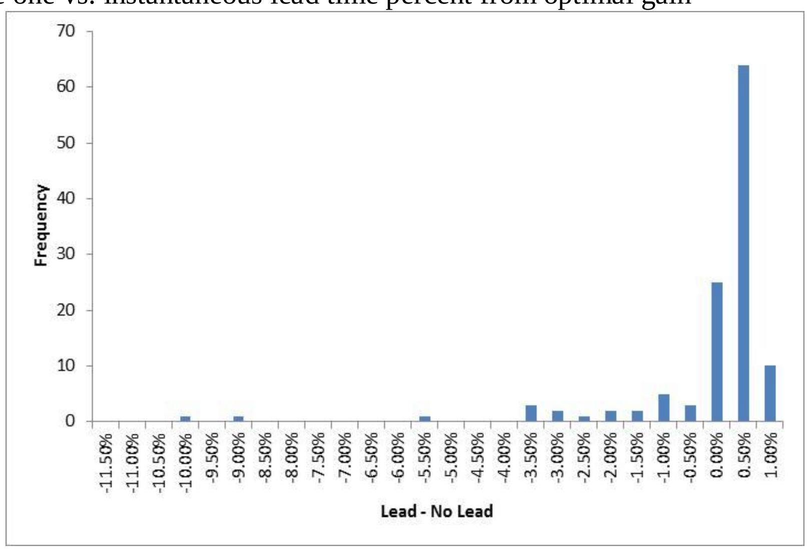


Increasing β increases the long term impact of a stock-out. As β increases, the error associated with the $\pi = r$ solution also increases. The mean error for $\beta = 4$ cases is statistically higher than any other level of β at 95% significance level. Additionally, $\beta = 2$ cases have a statistically higher mean error than cases with $\beta \leq 1$. In terms of mis-estimating λ, $\lambda = 1$ is statistically worse than $\lambda = 0.35$, but all other cases are not statistically different. Generally the errors are small when β and λ are small. This suggests that mis-estimating the parameters is not too costly if the estimates are close to the actual values.

Using this same experiment set for both lead time and instantaneous order arrivals, we can measure the sensitivity change across lead time. Using the $\pi = r$ solution to both problems will give a naïve solution gain to both lead-time and no-lead-time problems. Each of these gains is then compared to the optimal gain for its respective problem. We measure the sensitivity change as the difference between these two values as the percent deviation from optimal to the lead time case minus the deviation from optima for the no lead time case. If lead time is less sensitive to β and λ, difference between them will be negative, while a positive value means that the no lead time case is less sensitive than the lead time case. Table 3-9 summarizes the results of this experiment set. While most of the problems had no lead time closer to optimal than the lead time case (97 of 120 instances), when it was worse, it was much worse. The problem sets where the lead time case was closer to optimal occurred in the high β cases. The lead time did not see as an extreme difference from optimal as in the high β cases, but in most cases, no lead time was a little better than the no lead time.

Table 3-9 Lead time one vs. instantaneous lead time percent from optimal gain

% from optimal lead – no lead time	
Min	-10.32%
Mean	-0.38%
Median	0.08%
Max	0.63%



3.6. Implied penalty change as a function of lead time

In this section we use Morton's (1971) approximation to estimate a penalty cost for various lead times. Morton uses a pair of values z and S to define the heuristic policy as described in Section 3.5.3. The same set of problems will be used to determine the best penalty cost policy for lead times of zero through three periods. We notice that the order policy changes as a function of lead time. In section

3.5.3 we showed how to use the Morton approximation to calculate an implied penalty cost. These penalty costs will be compared for the same problem but different lead times. This procedure will help to determine if lead time influences the penalty cost.

The Morton heuristic structure allows an implied penalty cost to be determined. Enumerating all the penalty cost policies, it is possible to determine the best penalty cost to use. The Morton (1971) heuristic has been shown to perform increasingly poorly for longer lead times (Zipkin, 2008), but it is one of the few ways to relate the perturbed demand model to a lost sales penalty cost model. One alternative would be to use a base stock approximation, however it has been shown to be a bad approximation for the lost sales model when lead time is long and penalty cost is low compared to holding cost (Zipkin, 2008; Hue et al. 2009).

Thirty six problems are solved for the best Morton penalty cost policy for three cases: instantaneous, one-period, and two-period lead times. The gain for each policy and problem instance is calculated within an MP setting. To keep the state space manageable for the lead time two case, the mean demand was reduced to 5 from 20 and Δ was increased to 0.05 from 0.005. The test set consists of the following values: $c = 1$; $r = \{1.5, 2\}$; $h = \{0.05, 0.2, 0.4\}$; $\beta = \{0.25, 0.5, 1\}$; and $\lambda = \{0.75, 1\}$, with demand being negative binomial (5, 0.5). The average penalty cost across the 36 test cases was taken for each lead time instance. Figure 3-7 demonstrates that how average best penalty cost decreases as lead time increases. There is a significant drop from lead time one to lead time two, and it even dips below the revenue term value. As the lead time increases from one period to two the z value in the policy consistently drops (29 of 36 problems) while the S value consistently increases (36 of 36 problems), from best policy. The increase in S even with a decrease in π is due to the convolution of an additional period of demand. The decrease in z and increase in S make the policy have a longer flat section, meaning that the order quantity is a fixed level longer instead of having a base-stock structure.

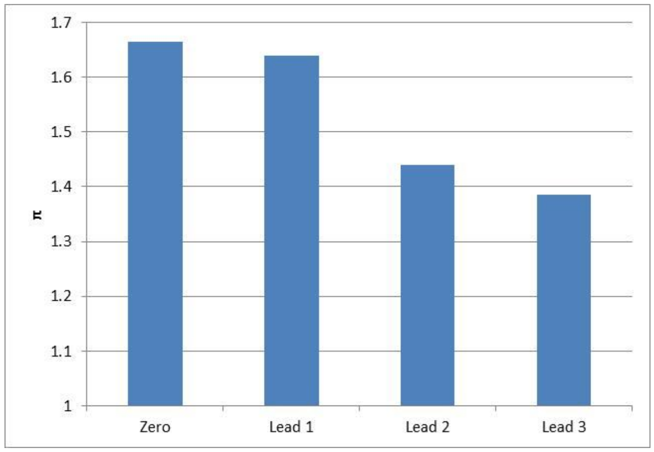


Figure 3-7 MP setting penalty cost as a function of lead time

Using a Markov Process to evaluate policies becomes impractical for problems with lead times of greater than three periods. The trend is to have a lower penalty cost as the lead time increases. The decreasing trend holds for all problem parameter subgroups. There are several possible explanations of this occurrence. One is that longer lead times result in a worse service level, thereby lowering the demand deflation so fewer orders are needed. There is a drop in demand deflation, but it is not enough to account for the significant drop in penalty cost. The second is that we are pooling risk as more periods are convoluted in the calculation of $\tilde{S}$. The two fractiles are the same, so lowering the needed fractile in S also lowers the $\tilde{z}$ fractile, resulting in a lower implied penalty cost as lead time increases.

3.7. Conclusion

The model developed in this paper extends periodic review perturbed demand models studied in Chapter 2 to include lead time. The optimal policy has a non-standard policy structure which makes it difficult to implement. The policy orders are based on where the system expects to be when the order arrives, factoring in both inventory and demand deflation values. As lead time grows, there is more uncertainty in the future state of the system when the order arrives. A heuristic was developed that found a good starting policy. A method to dynamically search the state space was developed, which decreased solution time and computer memory usage in solving the problem. For some problems, it made sense to implement Ding et al.'s (1988) solution method to decrease the solution time. Both methods use the idea that some sections of the transition and reward matrixes are never used and thus do not need to be calculated. The optimal solution to every problem was found using both methods to reduce the solution time. The Ding et al. (1988) procedure can be used to solve larger problems than standard value iteration approaches because less computer memory is used, because the entire probability transition and reward matrices do not need to be constructed.

Much like the lost sales penalty cost heuristics that assume a base stock policy structure, our model was solved by forcing the policy structure to match the zero lead time case. The zero lead time optimal policy structure has a different base stock level for each demand deflation. A method was developed that would systematically search for the best multiple base stock policy structure. This procedure is based on Howard's (1960) policy iteration method, but we alter the procedure to force groups of states to have the same decision and handle non-ergodic problems. Problem instances with low marginal profit values that force this policy structure lead to large deviations between the approximation and the optimal policy. As the profit margin grows, forcing this policy structure leads to much smaller deviations. This result is similar to that found by Huh et al. (2009), which showed that the base stock policy assumption works increasingly well for problems with large penalty costs relative to holding costs. The optimal policy is near base stock for large on-hand inventory levels. As it becomes more costly to miss a sale, retailers are more willing to hold large enough inventories, resulting in a policy that behaves like base stock.

A policy structure closer to the optimal policy structure was the Morton (1971) structure. This heuristic works well for the lead time one test problems. Because it performed well as a heuristic, it was possible to use this policy to find an implied penalty cost. The Morton structure was used to find the implied penalty cost for one through three period lead times. With the increase in lead time we saw a decrease in the average implied penalty cost. While most penalty cost models assume that the penalty cost is not dependent on lead time, we show that lead time does, indeed, have some effect on the penalty cost. Longer lead times do not allow the decision maker to adapt to changes in demand deflation as fast as the instantaneous order arrival case because units are not available to satisfy demand as quickly.

The difference in optimal policy structure between penalty cost and perturbed demand models with lead times is a significant finding. The penalty cost models assume that stock-outs do not have a direct correlation to loss of demand; consequently, the optimal policy structure is flat and then decreases. Perturbed demand models, on the other hand, assume that stock-outs alter demand: if stock-outs occur, then demand will decrease in the next period, and the order policy reflects this occurrence. Lead time can be handled in a similar fashion as penalty cost setting, but the optimal policy differences are more noticeable with lead time.

Chapter 4: Models to Assess the Impact of Demand Loss on Ordering Behavior in Multi-Period Inventory Systems

This paper considers an inventory control model in which some of the future demand is lost as a result of stock-outs. Traditional models apply a penalty cost to every stock-out that occurs, the value of which can be difficult to quantify. Stock-outs result in loss of immediate revenue and also additional factors such as a loss of customer goodwill that are harder to quantify. In our multi-period inventory model, realized demand in later periods is affected by fulfillment performance in earlier periods. Our model explicitly represents the effects of stock-outs on future demand, which then translates to a financial impact through the loss of future revenue. The order policy is assumed to be base stock and set before the start of the horizon, to mimic a penalty cost model. Our demand alteration method is difficult to solve for optimality, because the interaction between periods causes the cost function to be non-convex under some parameter sets. An approximation to the optimal solution is developed to reduce computation time while keeping near optimal performance.

4.1. Introduction

Stock-outs are a major concern for retailers because they are costly: United States' retailers lose an estimated $7- $12 billion dollars per year in gross margin due to stock-outs (Anderson Consulting, 1996). Corsten and Gruen (2003) attribute 51% of stock-outs in the United States and 31% in Europe to ordering practices. Difficulties in ordering stem from both quantity and timing of an order. Additionally, they estimate a 3.9% drop in future sales due to items being out of stock. Traditionally a penalty cost is used to quantify the value of a lost sale by assigning the event a cost proportional to the number of stock-outs, but this approach has certain shortcomings. Determining the value of a lost sale is not obvious because many components make up the cost. Indirect costs are assets to a retailer due to stock-outs; these costs include loss of future demand and customer goodwill (Liberopoulus et al., 2010). Additionally, companies may incur additional costs associated with backorders such as rushed deliveries fees. The problem arises when modelers attempt to factor in the loss of consumer goodwill due to stock-outs. Loss of consumer goodwill is difficult to monetize into a one-time cost. Schwartz (1966) introduced an alternative method to penalize stock-outs in which future demand is reduced as a result of stock-outs. This type of model is referred to in the literature as a *perturbed demand* model. In a perturbed demand model, there is no penalty cost for lost sales, only a loss of future demand and potential revenue.

The majority of inventory models currently implemented utilize a penalty cost to capture loss of goodwill and revenue with a single cost. Some modelers avoid estimating a penalty cost by setting a desired service level instead when finding the inventory policy. This method allows decision makers to circumvent estimating penalty cost directly, because the penalty cost is purely a function of the service level and cost parameters. However, the service level dependent penalty does not take into consideration how people react when faced with a stock-out. Two products with identical cost parameters and target service level, therefore, have the same penalty cost. Alternatively, the demand altering method can use empirical evidence of consumer behavior to adjust the model in order to fit specific products. Schwartz (1966, 1970) argues that altering demand is a more direct way of penalizing retailers for stock-outs than traditional penalty cost methods.

4.2. Literature Review

This research combines two streams of literature: the consumer behavior/response in the face of stock-outs and the inventory models that accurately reflect this response. In this chapter, we develop a model that more directly reflects the behavior of the consumers by integrating these two areas of research. In the following subsections, the current literature on consumer behavior and inventory models is discussed.

4.2.1. Consumer Behavior

Surveys show that when faced with a stock-out situation, consumers behave in a variety of ways. Verbeke et al. (1998) found three common consumer responses to stock-outs for grocery goods: (1) switch to competing brand; (2) switch stores to get the preferred brand; or (3) postpone the purchase. Other surveys include the additional option of switching to another size of the same brand (Emmelhainz and Stock, 1991). Table 4-1 summarizes the findings of six consumer response studies done on a variety of products. The percentages for each product do not necessarily add up to 100% because some consider additional consumer responses which are not shown on the table.

Table 4-1: Consumer behavior study summary

Product	Switch brand	Switch store	Postpone	Reference
Margarine	72%	2.9%	3.1%	Campo et al (2004)
Cereals	86%	6.5%	7.5%	
Coca Cola (soda)	65%	14%	21%	Verbeke et al. (1998)
Croma	47%	34%	19%	
Frieshe Vlag	62%	20%	18%	
Lassie (rice)	50%	28%	22%	
OMO (detergent)	31%	23%	46%	
Consumer goods	32%	13.7%	26.8%	Emmelhainz and Stock (1991)
Liquor	64.1%	14.1%	2.5%	Walter and Grabner (1975)
Bread	84%	10%	6%	Van Woensel et al. (2007)
Diapers	20%	39%	10%	Corsten et al (2003)
Toothpaste	24%	37%	7%	
Paper Towels	37%	13%	12%	

Table 4-1 demonstrates that the percentage of people willing to take a particular action is dependent on the type of product. The cost of a stock-out is dependent on the product type because consumer reactions are a function of the product (Corsten et al., 2003). Croma cooking oil, for example, has a strong customer loyalty. Customers are more likely to make a trip to a second store to get the brand they want (Verbeke et al. 1998).

Campo et al. (2000) classified how consumer stock-out behavior changes due to various parameters. They categorized parameters into three groups: product, consumer, and situational. Product parameters change based on the product's characteristics. Consumer parameters classify consumers by their traits (e.g. household income). Situational parameters change with respect to time or outside factors. One such factor could be a consumer's need for a product to satisfy an immediate need. Current research supports the idea that not all products have the same consumer reaction to stock-out situations; it is product and situation specific. Our demand altering function includes sensitivity factors to account for consumer behavioral changes.

The previous studies focused on customer responses when faced with the first stock-out; only a few papers have studied the long term effects of stock-outs. Customers faced with repeated stock-outs change their reaction to the stock-out situation. Walter and Grabner (1975) studied the actions of customers who had faced two stock-outs in a row. The percentage of customers that switched stores nearly tripled from 14.1% to 39.9%. Another study on mail order goods showed a significant drop in sales after an initial stock-out occurred (Anderson et al. 2006). Campo at al. (2004) estimated that the unavailability of one brand of margarine caused a 2% drop in total margarine sales. Multiple stock-outs decrease the perceived service level, thereby increasing the chances consumers switch to a different store for their future needs. Once remedied, the drop in sales decreased, but not necessarily reaching pre-stock-out levels (Walter and Grabner, 1975).

4.2.2. Inventory Models

Research that explicitly calculates the cost of stock-out is limited. Walter and Grabner (1975) calculated the penalty cost by estimating the cost and likelihood of each consumer response probability. Customers facing repeated stock-outs impose a higher cost on a retailer because a larger percentage of customers choose more costly alternatives such as switching stores. Chang and Niland (1967) and Oral et al. (1972) also estimated the penalty cost using similar methods. Most models estimate the true cost of a stock-out with little justification of its value (Hadley and Whitin, 1963).

The method used to model loss of consumer goodwill has been traditionally a penalty cost, but that is not the only approach used in the literature. Arrow et al. (1951) suggested multiple methods for how to penalize stock-outs including both loss of future demand and penalty cost. Bursk (1966) considers goodwill as an asset that the company can gain or lose base on its actions. Goodwill can be gained through good service or purchased in the form of advertisements. Schwartz (1966) and Hanssmann (1959) argue that loss of goodwill should be represented as a loss of future demand. They suggest that this represents consumer behavior more accurately. Lost sales have a much greater impact on future periods than the current period (Schwartz, 1966).

Schwartz (1966) developed a new type of model that used the number of lost sales to alter future demand, which he referred to as a *perturbed demand* model. He estimated the change in demand due to expected lost sales, instead of simply imposing a fixed cost when a lost sale occurs. Schwartz assumed that demand has a constant known rate across an infinite horizon. The demand rate is a function of the order policy. He assumed that when the order policy is implemented over a long horizon, the demand stabilizes to an altered but constant demand rate. These assumptions allow Schwartz's model to have identical periods. Our model extends Schwartz's work by allowing demand to be altered as a function of recent stock-out events and applying it to a finite horizon problem.

Liberopoulos et al. (2010) connected Schwartz's original model to more traditional models to find an implied penalty cost by setting the optimal decisions in both models to the same values and solving for the penalty cost. There were two significant results from this paper: (1) implied penalty cost is almost always zero or infinity, and (2) there are cases where the two models do not give the same optimal order policies. By comparing traditional penalty cost policy performance in our perturbed demand reality the best penalty cost value can be found. Taking the penalty cost which performs the best in our perturbed demand reality is a way to estimate the true value of a stock-out in terms of a penalty cost.

Schwartz (1966, 1970) laid the foundation for perturbed demand models, however his model focuses on constant demand cases and the long run state of the system. For short-lifecycle products, steady state policies do not always perform well. We assume that the demand is not reduced based on the order policy; rather, it is reduced based on actual stock-out events. This is similar to the model presented by Robinson (1990), which is the only periodic review perturbed demand model found in the literature.

More recently, Aksen (2005) developed a mixed integer programming solution to a lot-sizing problem where demand was altered as a result of missing sales. Their model assumes that the most recent period holds almost all the weight for predicting how customers react. They theorized that customers forget about the previous missed sales if the most recent interaction is favorable. It is also assumed that demand for each period is known. Our model is characterized by stochastic demand in each period and is more flexible in regards to the amount of weight given to past stock-out events.

4.3. Model Formulation

In this section we develop a periodic review perturbed demand model under stochastic demand. This paper is concerned with the finite horizon setting where N is the total number of periods in the horizon. We assume that in the absence of stock-outs, the demand in each period follows a stationary cumulative distribution $G(\xi)$, which we will refer to as the *underlying demand distribution*. However, the consumer behavior literature has demonstrated empirically that current stock-outs do affect immediate future demand (some customers abandon this product and/or retailer while others will return to purchase at a later date). Thus, consistent with other work that deals with perturbed demand, we employ a *demand deflation factor* in each period *n*, denoted by α_n where *n* is the number of periods remaining. This factor is computed based on the percentage of demand that was lost due to stock-outs in prior periods (*N* to $n+1$ periods remaining). For the remainder of the paper, our model is described in terms of periods-to-go, such that period $n-1$ follows period *n*. Thus, the demand deflation value with $n-1$ periods remaining is based on actual demand fulfillment performance in period *n*, and therefore, α_{n-1} is computed at the start of period $n-1$, before demand occurs. The actual realized demand with $n-1$ periods remaining is given by $\alpha_{n-1}\xi_{n-1}$ where ξ_{n-1} represents the realized sample from the underlying demand distribution $G(\xi)$. Thus, if the expected underlying demand is $\mu = \int\left(1-G(\xi)\right)d\xi$, the expected realized demand, given demand deflation factor α_{n-1}, is $\alpha_{n-1}\mu$.

The timeline of events is shown in Figure 4-1. Before the start of a horizon consisting of *N* periods, the inventory policy is set for the entire horizon. The policy is assumed to be base stock, but the level, denoted by S_n, may vary across periods. This assumption is made to ensure that the order policy structure matched the optimal policy of the lost sales penalty cost model. Orders are made and received before the start of demand in each period (zero lead time). The underlying demand distribution is assumed to be stationary over time, but the demand deflation value, and consequently the mean realized demand, is dependent on past events. When demand occurs, the inventory is lowered and revenue is generated. At the end of the period, revenue and holding costs are calculated. Between periods a new demand deflation α_{n-1} value is calculated. The cycle repeats until N demand periods have occurred where N is finite. Holding cost in the final period has the same value as other periods but changes its meaning to a salvage cost (the cost associated with having extra units at the end of the horizon). Our model assumes that ending the horizon in a state with a high demand deflation value holds the same value as ending in a state with a low demand deflation value. This means that it does not monetarily benefit the retailer to have a high demand deflation value at the end of the horizon.

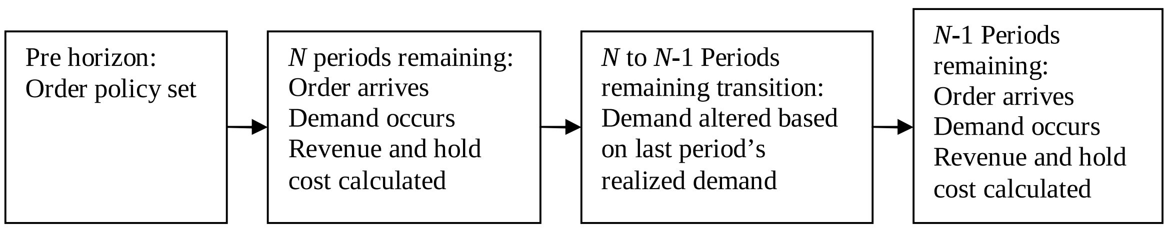


Figure 4-1 Timeline of events

A unit holding cost, h, is incurred for each item of inventory on hand at the end of period after demand occurs and revenue of r is generated for each sale. Revenue is modeled as a negative cost to keep the problem as a standard minimization problem. The immediate (one period) net cost, $L(y|\alpha)$, given y units on-hand after the order arrives and demand deflation factor α, is calculated as follows.

$$L(y\,|\,\alpha)=\int_0^{\frac{y}{\alpha}} h(y-\alpha\xi)\,g(\xi)\,d\xi-\int_0^{\frac{y}{\alpha}} r\alpha\xi g(\xi)\,d\xi-ry\left(1-G\left(\frac{y}{\alpha}\right)\right) \tag{8}$$

Given that the base-stock level is set to S_n, unit ordering cost is c ($c < r$) and x_n units are held in inventory at the beginning of the period with a demand deflation factor of α_n, the total net cost over the next n periods are calculated recursively using the following equations:

$$f_n\left(S_n\middle|x_n,\alpha_n\right)=c\left(S_n-x_n\right)+L\left(S_n\mid\alpha_n\right)+\omega\int_0^\infty f_{n-1}\left(S_{n-1}\middle|\left(S_n-\alpha_n\xi\right)^+,\alpha_{n-1}\right)g\left(\xi\right)d\xi$$
$$f_0\left(0\middle|x_0,\alpha_0\right)=0\quad\forall x_0,\alpha_0$$

where $g(\xi)$ is the underlying demand density function and ω is a discount factor ($0\le\omega\le1$). For discrete demand case, the total cost of the system can be calculated by enumerating all possible demand paths for N periods remaining. Problems with small number of states can be solved quickly using this procedure, but the time needed to calculate the total cost increases exponentially as the number of states increases.

Our claim is that while the demand deflation factor α_n is influenced by prior lost sales, this effect is attenuated by other aspects of the behavior of the consumers in question, namely the persistence of their memory of past stock-outs and the possibility that some customers are willing to effectively "shrug off" a stock-out such that it does not affect their likelihood of future purchase in any way. While the specific effect on demand is not known (we found no studies of consumer behavior that directly address this issue), we can attempt to capture the presumed consumer behavior using two model parameters: β, which we call the *deflation intensity*, reflecting the strength of the effect of stock-outs on future demand; and λ, which we call the *deflation persistence*, reflecting how long past stock-outs continue to influence future demand. β is limited to positive values and λ is bounded between zero and one. Schwartz (1966) suggested that the effect of past disappointment declines exponentially over time. Similarly, our demand deflation function with $n-1$ periods remaining depends on the incoming demand deflation factor (α_n); the realized sample from the demand distribution with n periods remaining (ξ_n); and the available inventory on-hand after the order arrives with n periods remaining (S_n), as follows:

$$\alpha_{n-1}=\begin{cases}\lambda+(1-\lambda)\alpha_n & S_n\geq\alpha_n\xi_n\\ \lambda\left(1-\frac{(\alpha_n\xi_n-S_n)\beta}{\alpha_n\xi_n}\right)^{+}+(1-\lambda)\alpha_n & S_n<\alpha_n\xi_n.\end{cases} \tag{9}$$

Expression (2) shows that, for a prior period (n, preceding period n–1) in which no demand is lost (top portion of the bracket in (2) where available inventory was sufficient to meet the actual realized demand), α_{n-1} is computed by smoothing it with α_n, reflecting the “deflation persistence” from previous periods. Thus, if λ (the persistence factor) is close to 1 (which means that almost only the most recent shopping experience matters for the customer), the effect of previous losses will decline rapidly, whereas if it is close to 0 (which means that the customer gives almost all importance to the earlier shopping experience, the most recent experience is the least important), previous demand fulfillment performance will dominate the current performance. The persistence factor can be thought of as a consumer’s memory with values close to one representing short memory and smaller values representing longer memory. Also, from expression (2), we see that for a prior period in which there is a demand loss (bottom portion of the bracket), the first term of the demand deflation formula reflects the deflation intensity (β) effect. The deflation intensity effect shows the impact of underage $(\alpha_n\xi_n-S_n)$ on future sales through α_{n-1}. The demand deflation value in future periods is altered by a multiple of β. The effect of a single instance of underage is exponentially smoothed across multiple periods by λ. At the start of the horizon we assume the demand deflation value equals one, meaning that there was no prior stock-out situation that would negatively affect current demand.

Note that the demand deflation factor is bounded between zero and one. This implies that demand cannot be negative, nor can it exceed the upper limit of the underlying demand distribution. It is assumed that the demand deflation value is constant throughout each period, meaning any stock-out that occurs within a period does not alter the demand until the next period.

4.4. Single-period

The simplest form of the problem is the single period problem (N=1), where a single order quantity needs to be determined. The horizon length is too short to have future demand lowered, and the value of having a positive demand deflation value at the end of the horizon is zero; consequently, the optimal policy looks similar to the optimal policy of the one-period penalty cost model. The single period total cost function and its first and second derivatives with respect to base-stock level S_1 are shown below.

$$f_1\left(S_1 \middle| x,\alpha\right) = c\left(S_1 - x\right) + \int_0^{\frac{S_1}{\alpha}} h\left(S_1 - \alpha\xi\right) g\left(\xi\right) d\xi - \int_0^{\frac{S_1}{\alpha}} r\alpha\xi g\left(\xi\right) d\xi - rS_1\left(1 - G\left(\frac{S_1}{\alpha}\right)\right)$$

$$\frac{df_1\left(S_1 \middle| x,\alpha\right)}{dS_1} = c - r + \left(h + r\right) G\left(\frac{S_1}{\alpha}\right)$$

$$\frac{d^2 f_1\left(S_1 \middle| x,\alpha\right)}{d^2 S_1} = \frac{1}{\alpha}\left(h + r\right) g\left(\frac{S_1}{\alpha}\right) > 0$$

$$\lim_{S_1 \to \infty} \frac{df_1\left(S_1 \middle| x,\alpha\right)}{dS_1} = c - r + \left(h + r\right) G\left(\infty\right) = c + h$$

$$S_1^* = G^{-1}\left(\frac{r - c}{h + r}\right)\alpha$$

The one period cost function is convex because of its positive second derivative. The optimal order quantity can be determined by setting the first derivative equal to zero. The optimal order quantity is a critical fractile multiplied by the current demand deflation factor α. The fractile solution takes on a standard newsvendor format like the single period lost sales penalty cost models do. The formula for the optimal base-stock level for the single-period perturbed demand model has two key differences compared to the one for the single period penalty cost model: (1) the penalty cost is replaced by revenue, and (2) the base stock level is scaled by the demand deflation factor. The revenue acts similar to a penalty cost term by imposing an opportunity cost in case of a lost sale. Scaling the fractile by the demand deflation value makes the optimal base stock level dependent on the current demand deflation value, reflecting the demand-lowering effect of past stock-out experience.

4.5. Multi-period problem

The two-period problem is the smallest problem where demand-lowering effects of past stock-out experience can be felt. We assume that retailers must specify their policy before the start of the horizon, thus for a two-period planning horizon, two base-stock values need to be specified. The optimal policy cannot be implemented under the current assumption of setting a single base stock value prior to start of the horizon because the optimal policy is dependent on current the demand deflation value. We forced the retailer to fix the policy upfront so that it would mimic a penalty cost model policy. The partial derivatives for the N=2 case which support this claim can be seen in the Appendix. The complicated structure of the partial derivatives makes difficult the determination of the optimal order policy analytically for N>1 problems. The interaction between periods is much more significant than in traditional penalty cost models, which results in difficult-to-calculate derivatives. The interdependence of periods is caused because the demand deflation value is calculated through exponential smoothing.

In addition to the interactions between periods, the cost function may not be convex for multi-period problems which can be demonstrated through numerical examples. Table 4-2 shows examples where N = 3 in which multiple local minimums were found using enumeration to evaluate all possible policies. In all of the examples in Table 4-2, one local minimum occurs with a small S_3, likely resulting in stock-outs and significantly lowering future demand. The second local minimum occurs when stock-outs are less likely and future ordering decisions remain relatively high. The most important decision in each problem is S_N because it has the largest impact on future demand; the effects of this decision can be felt for the most periods. All of the problems found to have multiple local minimums are extremely sensitive to stock-outs; demand changes dramatically across periods (high β, λ) as a result of stock-outs and the low margin (high h, low r), thereby making smaller order quantities attractive resulting in more stock-outs. The occurrence of multiple local minimums was not common among problems with more desirable cost and demand sensitivity parameters.

Table 4-2 N=3 non-convex examples

Distribution	c	r	h	β	λ	Local min 1 $[S_3, S_2, S_1]$	Local min 2 $[S_3, S_2, S_1]$
Neg. Binomial (20, 0.5)	1	1.2	0.55	2	1	[4, 0, 12]	[21, 17, 4]
Neg. Binomial (20, 0.5)	1	1.15	0.7	2	1	[3, 0, 12]	[21, 17, 4]
Neg. Binomial (20, 0.5)	1	1.25	0.7	2	1	[4, 0, 13]	[18, 17, 5]
Neg. Binomial (20, 0.5)	1	1.3	0.8	2	1	[4, 0, 13]	[21, 18, 7]

Non-convexity makes determining the optimal decision more difficult. It is required to solve a system of N equations with N variables in order to calculate the optimal policy analytically. Based on the analytical complexity demonstrated by the N=2 and 3 problems, more advanced solution techniques are needed to find the optimal policy.

4.6. Heuristic solution methodologies

Determining the optimal policy analytically is difficult for N>1 problems because of the interdependence between the orders; we propose two heuristic solution methodologies for determining an order policy. Due to the cost function's intricacy, we assume a discrete distribution of demand instead of a continuous distribution for our numerical evaluations. This is a realistic assumption since the demand the retailers face usually has discrete integer values. The negative binomial distribution is used to model demand because of its flexibility in taking on many different shapes. As an alternative to the analytical solution or total enumeration, two heuristic policy search procedures are considered in this paper, which are a gradient search method and a genetic algorithm.

4.6.1. Gradient Search

The gradient search method is a simple procedure that will guarantee that at least a local optimal order policy is found. Problems with multiple local minimums were rarely seen in two and three period problems through our experimentation. When the cost function is convex, this procedure will find the optimal solution. At each iteration the gradient search procedure used in this paper evaluates policies that are created by changing at least one base-stock variable in the current policy by +1 or -1. For a two period (N=2) problem, there are eight possible new neighbor policies to consider: (S_1+1,S_2), (S_1-1,S_2), (S_1+1,S_2+1), (S_1+1,S_2-1), (S_1,S_2+1), (S_1,S_2-1), (S_1-1,S_2+1), (S_1-1,S_2-1). The gradient search procedure evaluates the total cost of each neighbor policy created. The policy with the lowest cost becomes the new policy to search around. The iterative search procedure continues until no neighboring policy that is better than the current policy can be found.

The gradient search procedure's performance in terms of solution time and quality is heavily dependent on the starting solution. We use the approximation developed in Section 4.7 as a starting point. The downside of using a gradient search method is that a narrow solution space is searched; an objective function with many local minima will only find the local minimum associated with the initial solution. In order to explore the different parts of the solution space, multiple initial solutions can be used. The multiple-starting point gradient search procedure uses the approximate solution as a base solution then a set of initial solutions is created consisting of every combination of doubling S or setting it equal to zero in each period. Starting at multiple locations across the solution space potentially increase the probability of finding the optimal solution.

4.6.2. Genetic Algorithm

Genetic Algorithms (GA) generally work well at exploring large areas of the solution space at first then narrowing the search to improve the solutions. Joines et al. (1996) provides a good overview of genetic algorithms. Genetic algorithms make use of evolutional “survival of the fittest” concepts to develop better solutions over time. Current feasible solutions (parents) are modified or bred to produce new solutions (children). The children take the some genetic material (partial solution) from their parents in an attempt to make a better solution. Over a series of iterations (generations), this procedure tends to create better solutions and, in general, near optimal solutions can be found relatively quickly.

When using a GA the modeler must specify the following: the chromosomal representation, the evaluation function (the so-called fitness value of an individual solution), the size of the population, the selection process of the initial population, the genetic operators to be used, the strategy for selecting individuals to undergo genetic operations, and the algorithm termination criteria. The GA used in this paper uses an ordered vector of base stock levels for each of the N periods of the problem as the chromosomal representation of an individual solution. The fitness value is the resulting total cost of the specified base stock levels for each period in the model. The size of the population was set at 10 and the initial population was a set of 10 randomly generated policies.

Our GA uses two genetic operators, crossover and point mutation to manipulate individuals in the current population of solutions. The crossover operator takes two parent solutions then cuts them into two pieces at a randomly selected period n. Then corresponding pieces of the two parents are switched to create two new children solutions each of whom inherits a portion of the solution from each parent. The point at which the solutions are cut and switched is sampled from a discrete uniform distribution (1, N) such that all N period values have an equal chance of being selected. The resulting children are checked to ensure they are feasible solutions. Feasible solutions are solutions whose base-stock values in each period are less than or equal to the value of the prior period. When infeasible solutions are generated, new parents are selected and the process repeats until feasible children are generated.

The point mutation operator randomly selects a period, n, uniformly between 1 and N. Then, the base stock value for that period, S_n, is mutated to a value randomly sampled from a uniform distribution. The limits of the uniform distribution are the values that make the solution feasible. Using these operators, new solutions are generated in each generation.

Candidates for each genetic operator are selected by weighting each parent's fitness. The probability of selecting a given individual in the population to undergo genetic operations is based on a normalized geometric distribution:

$$P\left(selection = i\right) = q^{'}\left(1-q\right)^{i-1}$$

$$q^{'} = \frac{q}{1-\left(1-q\right)^{P}}$$

where q is the probability of selecting the best solution (lowest cost), P is the size of the population, and i is the rank of the solution (i^{th} lowest cost). To ensure that the sum of the probabilities equals one, q' is used. This normalized geometric distribution was shown to work well by Joines et al. (1996).

For our experimentation the GA is implemented for 100 generations. Over this time, the population generally becomes better as the fittest solutions are modified by the genetic operators and the weakest are dropped. We also ensure the fittest individual from each generation survives into the next using an elitist strategy where the worst solution from a generation is replaced with the best solution from the previous generation. This step enables the GA to keep the best solution in the parent set to be modified by genetic operators. After running a specified number of generation, the best solution is selected and the gradient search described earlier is used improve the solution to the nearest local minimum. Increasing the number of generations and parents increases the likelihood that a good solution is found, but there is no guarantee that the optimal solution will be found.

Using more generations and additional initial solutions can increase the likelihood of finding the optimal solution. Another approach is to couple the genetic algorithm with the gradient search procedure described above. The best solution after the last generation of the GA is used as the input to the gradient search procedure, guaranteeing at least a local minimum is found.

4.6.3. Solution Results

Long horizon problems are computational difficult to solve optimally because of the intricate nature of changing realized demand as a function of the demand deflation factor. For two period problems, the demand deflation value can't increase because it is assumed to start at a value of one (we assume initially customers are satisfied); however, with the inclusion of additional periods, a lowered demand deflation value in early periods can increase through high service. Three-period problems are the smallest problems in which the value of the demand deflation factor can increase after first decreasing. Fulfilling demand in the second period will lessen the effects of α_1. The solution procedures described in Sections 4.6.1 and 4.6.2 are tested on the base case for both the three and four period horizons. The base case problem parameters are described in

Table 4-1.

Table 4-3 Base model parameters

Parameter		Value
Unit cost	c	1
Holding cost per unit	h	0.2
Revenue per unit	r	1.5
Deflation intensity	β	1
Deflation persistence	λ	0.5
Underlying Demand Distribution	$G(\xi)$	Neg. Binomial (20, 0.5)

The results from the three-period case are the summarized in Table 4-4. The solution time and cost calculations reported for the genetic algorithm are an average of ten replications. The time recorded is the time needed to find the best solution, not the time it took to run through all the generations. The best solution time was used to demonstrate that using fewer generations may be required to find good solutions.

Table 4-4 Three-period solution comparison

Method	Time (sec)	Total Cost	% from opt cost	[S_3, S_2, S_1]
Enumeration	4,144.90	-22.9829	0	[25, 24,16]
Gradient search	2.58	-22.9829	0	[25, 24,16]
Multiple starting point gradient search	89.60	-22.9829	0	[25, 24,16]
Genetic algorithm	37.56	-22.9761	0.0296%	
Genetic algorithm with gradient search	70.33	-22.9829	0	[25, 24,16]

Four of the five solution methods arrived at the optimal solution every time, while the genetic algorithm without gradient search found the optimal solution eight of the ten runs. Enumeration is slow but is guaranteed to find the global optimal solution. Gradient search procedures simply guarantee that a local optimal is found, but there was only one local minimum for this problem instance. The gradient search is fast because it had a good initial policy, reducing the number of iterations until a minimum was found. The initial solution was [25 25 17] allowing the optimal policy to be found in one step. The genetic algorithm and multiple starting location gradient search methods have the advantage of searching a larger area of the cost function, which may be able to find more local minimums if they exist. The tradeoff is speed; larger search area increased solution time.

In the four-period case, the procedures arrive at the same policy but solution times increase significantly. Enumeration takes too long due to the added number of policies to consider and the increased time for each cost calculation. The gradient search time grows exponentially because of the additional points being evaluated before each step (26 vs. 80). The multiple starting point procedure uses more starting locations as N increases, thus increasing the solution time dramatically as N goes from 3 to 4. The genetic algorithm has an increased solution time due to the time required to calculate the cost with every additional period added. As the number of periods increases, the importance of having a fast solution becomes more important. Table 4-5 gives an overview of the solution method results for both solution time and total cost.

Table 4-5 Four-period solution comparison

Method	Time (sec)	Total Cost	% from best cost	[S_4, S_3, S_2, S_1]
Enumeration	No longer feasible to enumerate			
Gradient search	370.76	-31.1532	0	[25, 24, 24,16]
Multiple starting point gradient search	15,125.06	-31.1532	0	[25, 24, 24,16]
Genetic algorithm	2,632.79	-31.1405	0.0041%	
Genetic algorithm with gradient search	3,471.92	-31.1532	0	[25, 24, 24,16]

4.7. Approximate solution

In order to develop a fast solution for the perturbed demand model, an approximation was developed that only relies on the problem parameters to determine the order policy. The objective of the approximation is to not only provide faster solution time but also to provide a good initial solution for local improvement procedures (gradient search). The lost sales penalty cost version of this problem has a fractile solution (Morton, 1971). The fixed analytical fractile solution takes on the following form:

$$S_n = \begin{cases} F^{-1}\left(\dfrac{\pi - c}{\pi + h}\right) & n = 1 \\ F^{-1}\left(\dfrac{\pi - c}{\pi - c + h}\right) & n > 1 \end{cases}$$

where π is the penalty cost, c is the materials cost, and h is the holding cost per unit. While the perturbed demand model does not have a direct penalty cost, an approximate penalty cost may perform well. This approximation serves two purposes: (1) it provides a fast solution to the problem, and (2) it provides an estimate of a penalty cost to equivalence the models.

It was shown for single period problems that r and π are substitutable. For $N>1$, the penalty cost and revenue terms diverge because of the difference in how the models handle stock-outs. This paper uses a linear regression to estimate the penalty cost, which can be used to determine the order quantity for $N>1$ horizons. The approximation solution uses the penalty cost critical fractile with an estimated penalty cost to determine the order quantity for all periods $n>1$. The final period ($n=1$) uses the single period problem solution developed in section 4.4, which uses revenue in place of the estimated penalty cost.

A linear regression is used to estimate the equivalent perturbed demand model penalty cost. Two hundred randomly generated problems were constructed by sampling from a uniform distribution using the bounds outlined in Table 4-6. These problems were used as the training set for the linear regression. The optimal policy for each problem was determined by solving the N=2 perturbed demand problem using policy enumeration. Then each problem set was resolved in the lost sales penalty cost setting. The penalty cost was adjusted until the order policy was matched that of the perturbed demand solution. As noted in Liberopoulos et al. (2010), it is not always possible to find exact matches in the order policy. In the case where no exact match can be found, we assume that the penalty cost which resulted in the closest order policy is used. The closest policy is defined as the policy that has the smallest absolute difference in order quantities over both periods. This provided a set of data wherein the best penalty cost was found for each problem instance.

Table 4-6 Bounds for test problems

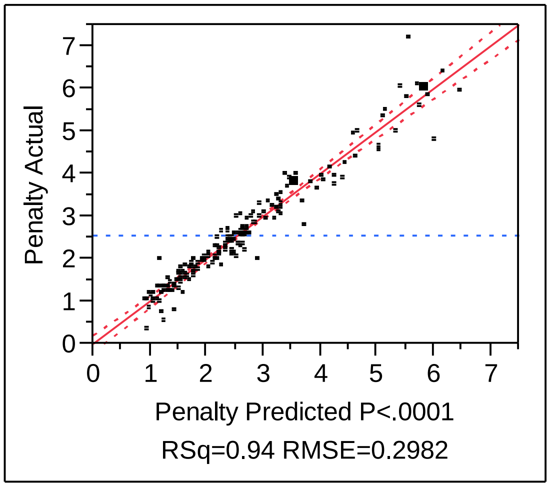
Parameter		Range
Unit cost	c	1
Holding cost per unit	h	[0,1]
Revenue per unit	r	[1, 5]
Product sensitivity	β	[0,2]
Weight parameter	λ	1
Underlying Demand shape	p	[0, 1]
Underlying Demand scale	m	[0, 40]

Using this data set, a multiple variable linear regression was fit to predict the penalty cost. The independent variables are as follows: the negative binomial demand shape (p); scale parameters (m); holding (h); revenue (r); and sensitivity to missed demand parameters (β, λ). The dependent variable is the penalty cost. The equation takes on the following form, where the y_i's are the slope parameters and y_0 is the intercept:

$$\hat{\pi} = y_0 + y_1 h + y_2 r + y_3 \beta + y_4 \ln\left(m\right) + y_5 p + y_6 \ln\left(m\right) p + y_7 pr + y_8 hr + y_9 r\beta$$

SAS JMP software was used to estimate the best fit line through the data. The natural log of the scale parameter *(m)* was used to reduce a non-linear trend in extreme values.

The R^2 value is greater than 0.9, indicating that this model is a good fit to the data. There are a few outliers as seen in Figure 4-4. Some outliers can be explained by the randomly generated problems being unprofitable products. For example, one problem had cost parameters of r=1.035, c= 1, and h=0.815. This results in an extreme implied penalty cost, which makes it difficult to estimate. Figure 4-2 gives the best fit value for each factor, which are statistical significant at the 0.05 confidence level. Models including additional interaction terms were considered but were not statistically significant. The benefit of using this linear regression model is the ability to take the perturbed demand model parameters and estimate a penalty cost without having to solve either model first.



Term	Parameter Value
Intercept	-0.469
h	0.364
r	1.092
β	0.281
$\ln(m)$	0.099
p	-0.685
$(\ln(m)-2.78)*(p-0.49)$	-0.258
$(p-0.49)*(r-2.39)$	-0.502
$(h-0.51)*(r-2.39)$	0.337
$(r-2.39)*(\beta-1.01)$	0.098

Figure 4-2 penalty cost regression

The inherent advantage of this procedure is that an order policy can be generated without solving any total cost functions which makes this solution procedure very fast. The procedure shown here is limited to the negative binomial distribution and the scope of the training set problems. If other distributions need to be used, then additional models will need to be fit to take into account their demand distribution parameters.

4.7.1. Approximation experiments

The robustness of the approximate policy was tested by using a full factorial experimental design. Three levels were used with a high, middle, and low level for each parameter. With five parameters and three levels for each parameter, there are 243 different test cases. The demand scale parameter was adjusted for each problem to ensure a constant mean of twenty. Each problem was solved using the approximation and gradient search methods. The performance of the approximation was measured in terms of percentage difference from the gradient search policy. A separate ANOVA test was completed for each parameter to see if there was a statistical difference in the percentage from the best known solution using the approximation under a specific set of parameters.

Table 4-7 Three and four period 3 level full factorial results

Variable Changed	Low level	Mid level	High level	N = 3 P-value	N = 4 P-value
r	1.3	1.5	1.7	<0.0001	<0.0001
β	0.5	1	1.5	0.003	0.0004
h	0.2	.3	0.4	<0.0001	<0.0001
λ	0.25	0.75	1	0.013	0.0274
p	0.25	0.5	0.75	<0.0001	<0.0001

All the terms had a significant difference between the levels at the 95% confidence level. The approximation performs worst for low-margin products which cover both h high and r low cases. Products that are more sensitive to losing demand (high β, λ) are statistically more costly compared to the gradient search policy. It is harder to estimate problems with these parameters because shifts in the demand are more dramatic. Demand distributions with a low p parameter have the highest coefficient of variation (CV) which might explain the increased cost difference. Distributions with higher CVs are often harder to predict for any type of inventory model, not just our model.

The same full factorial design used in the three-period model was applied to four-period problems. The test instances which caused high deviation from optimal were the same for both the N=3 and 4 periods problems. The λ term is slightly less significant and the β term is slightly more significant for the four period horizon versus the three period horizon; however, all variables are still significantly different across tested levels at the 95% confidence level. The cases with low margin and highly sensitive to missed demand (high β and λ) are still the hardest to estimate. The same test problem that had the maximum error in the three-period also resulted in the maximum error in the four-period case. The worst three-period case tested had a 12.31% cost deviation from optimal while the four-period max deviation was 10.57%. The largest difference was the result of all the parameters having values in which the approximation performs poorly at the same time: a low margin product that is very sensitive to missing demand with a high CV demand distribution. The low profit margin and highly sensitive consumer behavior have a compounding effect on the deviation from the best solution that the gradient search procedure found. Figure 4-3 shows the distribution of deviations for both the three- and four-period problems. Most problems fall within a 2% deviation from the best known solution: for the three-period problems, the deviation from the best known solution was within 2% for 138 of the 243 cases while for the four-period cases, the deviation for 160 of the 243 cases was within 2% of best known solution. Only four three-period and one four-period test cases had deviations of greater than 7%. Generally, additional periods lowered the deviations from optimal for the test set. The approximate solution works well unless multiple extreme parameters occur (low margin and fast changing demand deflation value).

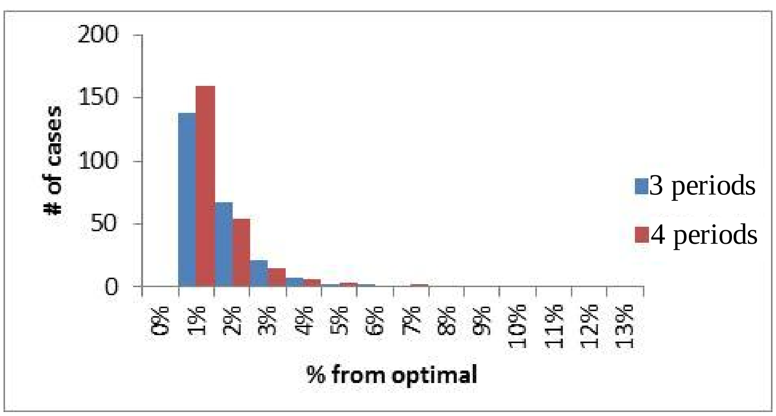


Figure 4-3 Approximation deviation from optimal 3 and 4 periods

4.7.2. Effect of penalty estimation method

The policy generated by using the estimated penalty cost performs well for many problems because it takes all of the problem parameters into consideration. The sensitivity of the penalty cost estimation is tested by using alternative penalty cost estimation methods. One alternative is to use a defined service level to find an implied penalty cost. A commonly used service level is 95% which was tested. Liberopoulus et al. (2010) found that the implied penalty cost was usually zero or infinity for Schwartz's perturbed demand model. In addition to testing these two penalty cost estimators we will also use the revenue (r) as the penalty cost, which implies that there are no long term effects of stock-outs. Additionally, revenue has the largest impact on the regressed penalty cost estimate. Lastly, over and underestimating the regressed penalty cost by 5% is used to test the sensitivity of parameter estimates. All of these estimates are evaluated to see which of them provides the best total cost for a set of problems.

An experimental set of 500 test problems with N=4 were used to evaluate the total cost of implanting the policy generated by each penalty cost estimator. Each problem was randomly generated by sampling each variable from a uniform distribution between the upper and lower bounds provided by Table 4-8. The negative binomial demand distribution scaling factor is adjusted based on the demand parameter p to keep the underlying means constant at twenty units per period. The total cost for each new estimate was compared to the regression estimated penalty cost developed in section 4.7.

Table 4-8 Bounds on random problems		
Parameter		Range
Unit cost	c	1
Holding cost per unit	h	[0,1]
Revenue per unit	r	[1+h,3+h]
Product sensitivity	β	[0,1]
Weight parameter	λ	[0,1]
Underlying demand distribution	p	[0.1, 0.9]

Table 4-9 shows the percentage difference between the total cost for impending estimated penalty cost policy and using the regression generated policy. The top three penalty cost estimation procedures all performed similarly to the regressed penalty cost value developed in Section 4.7.1; in some cases, these alternative methods out-performs our regression estimation procedure. The first three rows have means that are statistically equal to the regression at a 95% confidence level. The estimator based upon the regression estimate minus 5% has a max error of 185.64%, while that based on the regression plus 5% has a max error of only 2.84% for the test set. Thus, when using the estimated penalty cost, it is better to err on the side of over penalizing than under penalizing to reduce the maximum error. Over-ordering is better than under-ordering since stock-outs are reduced keeping the demand deflation value high. With more demand, more units are sold resulting in a flatter cost difference when over-ordering. Penalty cost models are also less sensitive to over-ordering than under-ordering. Use of the revenue value as an estimator of the penalty cost seems to work well because the full cost of missing a sale may not be felt in such short horizons. Over longer horizons, using revenue as the penalty cost results in under-ordering because it does not take into consideration the full effect of missing a sale. The end of the time horizon has an effect on the ordering policy;

specifically, there is no value or cost for ending with any particular demand deflation level. Thus, the system orders less to reduce holding cost. Additionally, lost future demand is truncated by the end of the horizon, meaning stock-outs at the end of the horizon are less costly than ones with more time remaining.

Table 4-9 Error results from penalty cost estimation methods

Penalty estimator	Min Error	Mean Error	Median Error	Max Error
r	-23.45%	-0.24%	0.00%	5.84%
Estimated +5%	-23.29%	-0.32%	0.00%	2.84%
Estimated -5%	-2.67%	0.91%	0.00%	185.64%
95% service level	-1.56%	16.27%	7.11%	439.79%
0	2.01%	54.27%	31.90%	525.00%
Infinity	7.55%	84.53%	54.72%	684.08%

Using a estimator based on (1) the critical fractile for a 95% service target, (2) a penalty cost of $0, and (3) an infinite penalty cost (bottom three rows in

Table 4-9) all have very poor performance compared to the regression estimated penalty cost, i.e. statistically worse than the regressed penalty cost at a 95% confidence level. While penalty costs of zero or infinity might work well for Schwartz's perturbed demand model, they do not work well for ours. Both of these values have at best a 1.79% deviation from the regression generated estimated penalty cost, and potentially a much larger deviation. Using a base stock derived from a 95% service level critical fractile does not result in a good estimate of the implied penalty cost, though it is better on average than either using zero or infinity. Generally, as long as a penalty cost is selected that is reasonably close to the regression estimate, the errors are small, however poor estimates can lead to very poor performance.

4.8. Markov process cost approximation

Calculating the total cost by enumerating all possible demand paths works well for short horizons ($N \leq 4$), but the computational time required for the total cost calculation grows exponentially with N. Using the gradient search method for large N problems is similarly disadvantaged because of the increased number of alternates (3^N) that need to be considered before each step. The genetic algorithm also becomes slow due to the increased number of possible policies. The Markov process (MP) cost approximation can compute the total cost of a fixed policy for long horizon problems, but increasing N makes solving for the optimal policy more difficult.

Using a Markov process to estimate the cost allows for alternative solutions methods to be implemented. Finite horizon Markov Decision Process (MDP) analyses are often solved using the value iteration approach. Value iteration solution time increases linearly with N instead of exponentially. A single period's cost can be calculated if the current inventory and demand deflation value are known. Under the assumption of a known realized demand distribution, it is possible to calculate the single period transition probabilities. However, in order to implement an MDP, the state space must be finite, therefore the demand deflation value must be modeled as a discrete variable whereas in reality, it is continuous. One discretization method is to segment the variable into uniform pieces referred to as a grid (Chow and Tsitsiklis, 1988). Approximating the demand deflation factor as a discrete value results in some error in the total cost calculation due to the rounding required.

The magnitude of the error is a function of two parameters: horizon length (N) and the granularity in which demand deflation factor is discretized. The granularity changes the maximum amount of rounding that is needed in each period. The demand deflation factor, for example, can be segmented into equally spaced intervals (e.g. 0, 0.05, 0.1, 0.15….0.95, 1). The distance between the discretized demand deflation values is defined as Δ. In order to convert continuous demand deflation values into discrete segments, the actual value of the demand deflation factor is rounded to the nearest discrete segment at the end of each period. The maximum possible rounding is $\Delta/2$. When Δ =0.05, if α = 0.93, then the discretization procedure will round to α=0.95. The need to round the demand deflation factor is the root cause of the error in estimating total cost. Smaller Δ values will result in less rounding yielding a better estimate but increase the size of the state space.

A Markov process is used to calculate the total cost over the horizon. The transition (p_{ij}^k) and reward (R_{ij}^k) matrices need to be constructed in order to implement the MP setting. The probability that the system transitions from state *i* to state *j* under decision *k* is defined as p_{ij}^k. The reward for transition from state *i* to state *j* under decision *k* is defined as R_{ij}^k. The immediate expected reward for being in state *i* and ordering *k* is defined as q_i^k where $q_i^k = \sum_{j=1}^{N} p_{ij}^k R_{ij}^k$. The total cost of the system is calculated using $v_i(n+1) = q_i^k + \sum_{\forall j} p_{ij}^k v_j(n)$ where $v_i(0) = 0$ for all *i*. The total cost of the system is $v_l(N)$ where *l* represents the state with zero inventory on hand and α=1.

Fifty randomly generated test problems are used to evaluate the size and trend of the error as a function of *N*. Four different Δ values are tested to determine the change in error due to rounding. The choice of Δ is a balance between realism and solution time. Fifty test problems are generated using the parameter bounds outlined in Table 4-8. A smaller constant underlying demand mean, 5, is used to extend the maximum horizon length to seven for the exact cost calculation. Anything greater than seven periods requires hours to solve using the exact method even with the low mean demand. The regression approximation developed in Section 4.7 was evaluated using the cost calculation method. The error is measured as the absolute percentage difference between the total cost estimation using the MP setting versus the total cost using demand path enumeration.

Figure 4-4 displays the average error over the fifty test problems for a variety of N and Δ values. The average error is less than 1% for all combinations of N and Δ, indicating that not much is lost in terms of accuracy in the short term by discretizing the demand deflation factor. Additionally, errors are increasing at a decreasing rate as a function of N, which is partially noticeable in the $\Delta = 0.1$ case. The average error associated with the Δ=0.1 case leveled off at approximately 0.6%. The average errors for the remaining Δ values are nearly identical, and the right graph indicates the mean error is leveling off as N increases. The mean error is small and increases at a decreasing rate as N increases for all Δ levels tested.

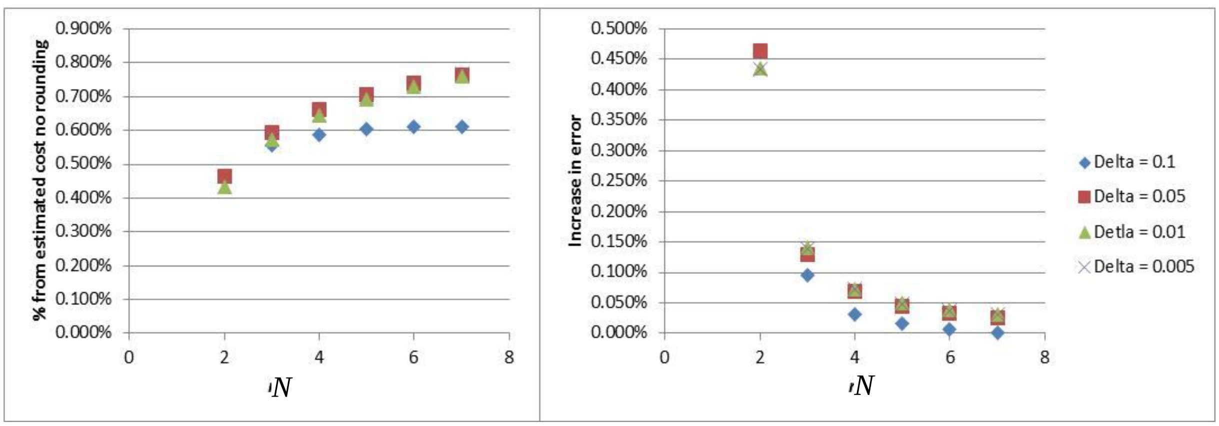


Figure 4-4 MP accuracy test

The MP approximation seems to be a reasonable approximation to the exact method. The results indicate that the error is relatively small and levels off as N increases, making the approximation useful for long horizon problems that could not otherwise be solved.

4.8.1. Infinite horizon penalty cost policy

The optimal policy structure in the infinite horizon is base stock dependent on the demand deflation level; additionally, it has been shown that using a single base stock level works well for most problem instances (Section 2.7). It was shown in Chapter 2 that for a given base stock level there was an implied penalty cost range which could be used to arrive at this policy using traditional models. Any penalty cost higher than the upper end of this penalty cost range would yield a policy that orderstoo much while dropping below the lower end of the range would yield a policy ordering too few. Appling the infinite horizon penalty solution in the finite horizon might be a useful approximation if the finite horizon policy quickly converges to the infinite horizon policy. It is much easier to calculate the infinite horizon penalty than the penalty cost for the finite horizon case. Additionally we have regression models that work well for the infinite horizon setting. Similar to the regression developed in Section 4.7.1, a penalty cost can be estimated based on the problem parameters.

Using a small grid size (Δ=0.005), the best penalty cost in the finite horizon case converges to the infinite horizon solution within twenty periods for the majority of tested problems. Problems where the demand deflation value changes slowly across periods take longer to converge (low β, λ). For the discussion that follows we define goodwill (GW) as the difference between the penalty cost and revenue (sales price, r). This is the additional cost implied by the ordering policy in addition to the margin loss for each missed sale. A plot of goodwill cost as a function of horizon length (Figure 4-5) shows the average goodwill cost for 36 test cases assuming demand follows a negative binominal distribution (20, 0.5), c = 1, and r = 1.5. The other problem parameters vary across problems h = {0.05, 0.2, 0.4}; β = {0.25, 0.5, 1}; and λ = {0.35, 0.5, 0.75, 1}. As the horizon increases the goodwill cost increases, until it converges at around fifty periods. The upper (Inf UB) and lower (Inf LB) bounds give a range of penalty costs which results in the same order policy.

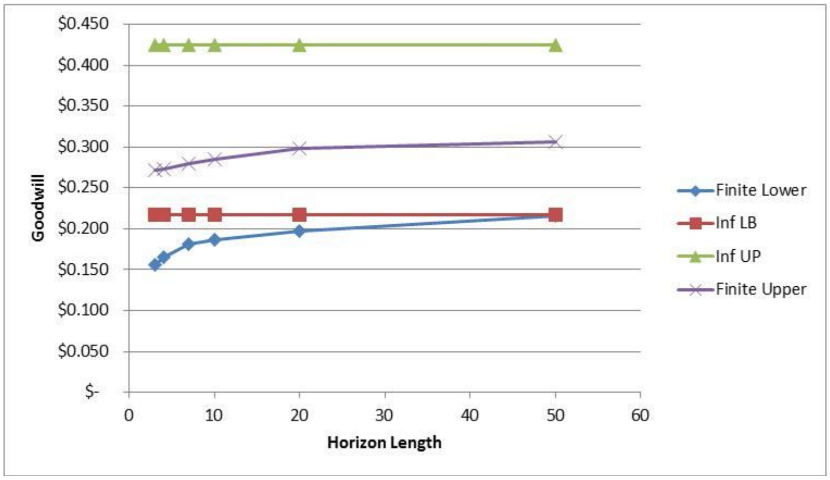


Figure 4-5 Average goodwill bounds for 36 problems

Figure 4-5 suggests that the infinite horizon penalty cost lower bound can be used to approximate the penalty cost for the finite horizon setting. The performance of using the infinite horizon penalty cost in the finite horizon setting was tested by evaluating 100 randomly generated problems (see Table 4-8 for the parameter bounds). The infinite horizon penalty cost was applied in the finite horizon and compared against the best penalty cost policy, as well as the optimal solution. There is overlap in the best finite and infinite penalty cost ranges, i.e. the lower bound of the infinite horizon is contained in the finite horizon penalty cost range on average. Table 4-10 shows the performance of the infinite horizon penalty cost lower bound compared to the best penalty cost policy and the optimal policy. We also tried implanting the infinite horizon penalty cost lower bound plus 10%, to see if there would be an improvement in the policy. The maximum error occurred in problems where λ was low (<0.082). In such problems if demand is lost, not all of it can be gained back due to the discretization of the demand deflation factor.

Table 4-10 Infinite horizon penalty cost applied to finite horizon

Penalty cost		$N = 3$			
		Min	Mean	Median	Max
Optimal Solution	Inf LB GW	0.00%	0.27%	0.16%	3.24%
	Inf LB GW+10%	0.00%	0.25%	0.14%	3.97%
Best penalty cost	Inf LB GW	0.00%	0.15%	0.09%	2.47%
	Inf LB GW+10%	0.00%	0.14%	0.05%	2.47%
Penalty cost		$N = 10$			
		Min	Mean	Median	Max
Optimal Solution	Inf LB GW	0.00%	0.24%	0.13%	2.84%
	Inf LB GW+10%	0.00%	0.18%	0.08%	2.84%
Best penalty cost	Inf LB GW	0.00%	0.13%	0.06%	2.83%
	Inf LB GW+10%	0.00%	0.07%	0.00%	2.83%
Penalty cost		$N = 20$			
		Min	Mean	Median	Max
Optimal Solution	Inf LB GW	0.00%	0.25%	0.15%	3.73%
	Inf LB GW+10%	0.00%	0.18%	0.07%	3.73%
Best penalty cost	Inf LB GW	0.00%	0.15%	0.04%	3.62%
	Inf LB GW+10%	0.00%	0.08%	0.00%	3.62%

The results suggest that there is a small difference between the best penalty cost value for infinite and finite horizon problems. The small differences are due to the end of horizon effects. The optimal policy orders based on the current demand deflation value, whereas the penalty cost policy does not. This difference accounts for the difference in solution performance. However, the benefit of implementing the optimal policy is small and may not be worth the added complexity needing to track the demand deflation value.

4.8.2. MP cost approximation heuristic

We assume the order policy must be a single base-stock level defined before the horizon starts, while the optimal policy orders based on the current demand deflation value and the number of periods remaining. Forcing the value iteration procedure to take on a specific policy structure can be implemented by modeling the problem as a restricted observation Markov process (ROMDP) where the demand deflation value is unknown to the decision maker.

Wei (2005) and Davis et al. (2008) developed an infinite horizon ROMDP solution method; they grouped states to ensure a common decision is made across the group. The solution procedure makes a single order decision for a group of states by weighting each state by its likelihood. In our formulation, all of the states for a given number of periods remaining are grouped together in order to force a common base stock level across demand deflation levels. The base stock policy is allowed to change as function of n but not α.

The ROMDP procedure applied in the infinite horizon problem relies on the steady state probabilities; in the finite horizon version the state probabilities and order decisions are dependent on the number of periods remaining. Both the state of the system and the expected costs are required to make a decision. The increased number of unknown values makes solving for the finite horizon ROMDP more difficult. Define the state probabilities as $z_i(n)$ where i is the state of the system and n is the number of periods remaining. With N periods remaining, the state probability for each state ($z_i(N)$) is assumed to be known, but the cost for each alternative ($v_i(N)$) is unknown. With one period remaining, the cost ($v_i(1)$) are known but the state probabilities $z_i(1)$ are unknown. The interdependence of the cost and state probabilities makes the problem difficult to solve.

To alleviate the problem associated with unknown state probabilities, they are estimated based on the regression approximation policy. The initial state of the system is assumed to be zero inventory with a demand deflation value of 1.0 (x = 0 and α = 1.0). For a given policy, transition probabilities and a starting condition, an initial set of state probabilities can be estimated. The ROMDP value iteration solution procedure can use these estimated probabilities to determine a policy. The resulting policy is then used to estimate a new set of $z_i(n)$. This cycle is repeated until there is no change in the policy compared to the previous ROMDP solution iteration. The solution procedure is outlined in the steps below:

$$z_i(N) = \begin{cases} 1 & x_i = 0 \wedge \alpha_i = 1 \\ 0 & otherwise \end{cases}$$

$$k(S, x) = \min(S - x, 0)$$

Step 1. Use regression approximate policy to estimate z_i for all periods

$$z_i(n-1) = z_i(n)\, p_{i\cdot}^{k(S_n, x_i)}$$

Step 2. Use ROMDP solution method to minimize cost

$$v_i(0) = 0 \quad \forall i$$

$$v_i(n+1) = \min_{S_n} \sum_{\forall i} z_i(n) \left[q_i^{k(S_n, x_i)} + \sum_{\forall j} p_{ij}^{k(S_n, x_i)} v_j(n) \right]$$

Step 3. Use ROMDP policy to updated z_i

Step 4. Repeat steps 2 and 3 until the policy converges

One hundred random test problem were generated using the parameters in

Table 4-6 and tested in the N = 3 and 4 setting. The policy with the best MP estimated cost was always found using the ROMDP procedure. However, the policy resulting in the best cost under the MP cost estimate does not always match the policy with the lowest cost found using the exact cost calculations. Due to the MP cost approximation used in the MP, the policy which has the lowest cost may be different that if using the exact cost calculation. In our study, the small error associated with the MP cost calculation kept the MP low cost policy near the exact cost solution for all the tested problems. The small deviation from the optimal policy yielded small percentage cost differences from the best policy for all tested N=3 and 4 periods problems. For N = 3, 100 problems generated with Δ

=0.05, the max deviation from optimal was 0.14% with the optimal solution found 75% of the time. The same problems tested in the N = 4 setting, the maximum deviation from the best known cost was 0.07% with the optimal solution found 67% of the time.

The first step in the ROMDP solution procedure estimates the state probabilities by using the regressed penalty cost approximation policy. The initial state probability estimates have no impact on the final solution for any of the test problems. Each problem was resolved using the initial state probabilities generated from a policy of ordering up-to one unit in each period in order to test if the final solution was dependent on the initial solution. Changing the initial state probabilities estimates did not change the final solution in the ROMDP procedure. The change in the initial solution did change the number of iterations required for convergence. Using the naïve policy with base stock values of 1 takes more iterations to converge than using the regression estimated penalty cost policy as the initial solution.

For longer horizon problems, there is no alternative search procedure to compare against the ROMDP approach. Gradient search and genetic algorithms are not efficient solution methods because of the additional number of possible policies to consider. Two solution methods considered in this paper that can generate a policy for N = 10 problems are the regression approximation and the ROMDP. Using the same one hundred problems used for N = 3 and 4, the ROMDP always outperforms the regression approximation. The average improvement over the approximation was 0.2% with a best case showing a 1.22% improvement. On average it took over five minutes to compute the ROMDP value iteration policy compared to the fractions of a second that the regression approximation took. Part of the improvement can be explained by the stair-step structure of the optimal policy. In the ROMDP policy as the number of periods remaining, n, increases, the base stock value increases slowly with consecutive periods often having the same base stock level. For N = 10 period problems, there were often three or four different base stock levels used in the ROMDP policy.

The regression approximation uses two bases-stock levels over the horizon (n=1 and n>1) regardless of the horizon length. The multiple levels of order up-to quantities indicate that end of horizon effects are more complex and further reaching than is accounted for with the penalty cost method.

The comparison between the approximation and the ROMDP solution does not provide insight to the quality of the ROMDP policy compared to the optimal policy for N = 10 problems. To this end, a total cost lower bound can be calculated by removing the base-stock structure order restriction and solving the problem using the traditional value iteration approach. Removing the restriction on the order policy allows orders to be made based on the inventory and demand deflation value in each period. Comparing the ROMDP to the unrestricted optimal policy provides an lower bound on the deviation from the optimal cost with the base-stock order policy restriction.

Over the same 100 problems used in the N=3 and 4 horizon length, the average difference in cost between the ROMDP solution and the unrestricted order policy was only 0.1%, with only one test problem having a deviation of greater than 1% (2.15% max deviation). The maximum deviation occurred in a problem with a low margin with future demand sensitive to stock-outs (high β, λ). This type of problem is particularly problematic because the low margin results in a lower order quantity, thereby increasing the likelihood of stock-outs. Additionally the high sensitivity to missed demand changes realized demand dramatically across periods.

Another set of problems was created to test the performance of the ROMDP procedure under difficult problem instances. The revenue was reduced to ensure that problems had a revenue value that is at least 5% above materials cost (r>1.05c). The demand deflation and demand intensity factors were both limited between 0.7 and 1 to ensure that future demand was sensitive to stock-outs. An additional policy is used in this section (referred to as the “no demand loss solution”), is the best policy found assuming that stock-outs do not affect future demand (β=0). Figure 4-6 plots the percentage differences from the optimal policy without order restriction for the ROMDP, no demand loss solution, and regressed penalty cost solutions for sixty test problems across various time horizon lengths. For these particular problems, the parameters are outside the range of parameters used to train the regression equation which leads to very poor solutions. The no demand loss solution orders as if there is no possible way to lose future demand making the penalty cost equal to the revenue and ordering based on the fractile. We see that the no demand loss solution performs well indicating that the cost of goodwill is low. The ROMDP solution seems to be approaching the no demand loss solution as the length of the horizon increases. One possible explanation is that the end of horizon effects, where the largest difference between the two occur, become less of an impact on the total cost as the horizon increases.

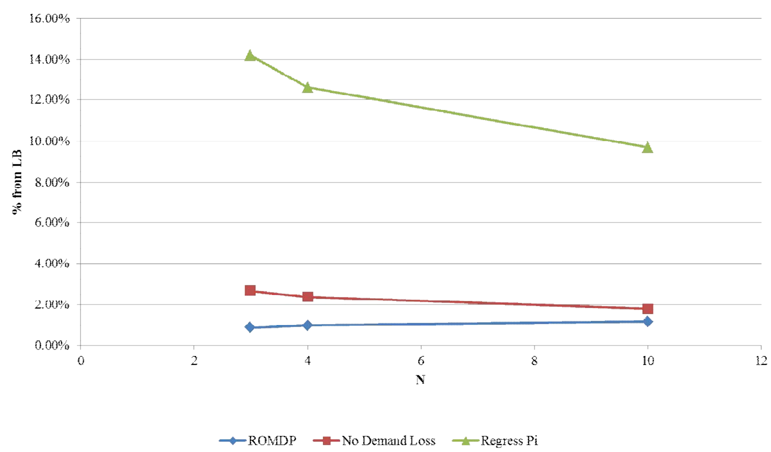


Figure 4-6 Low margin sensitive demand loss problems

4.9. Conclusion

In this paper a perturbed demand model is developed to handle a finite horizon periodic review setting and non-seasonal stochastic demand. The model assumes that a base stock order policy must be established before the start of the season and that once demand begins, the policy is not changed. This assumption was made to force the policy to match traditional penalty cost models. The order policy assumption causes the cost function to potentially be non-convex. This makes finding the optimal order policy slightly more difficult than standard inventory models. Enumeration of all possible order policies quickly becomes impractical as longer horizons are considered. Gradient search and genetic algorithms are developed to solve the problem to replace enumeration as a feasible solution method. Additionally, an approximation to the problem is created that yields a near optimal solution quickly. The approximation also results in a penalty cost estimation linking our model to penalty cost methods. Calculating the total cost of the system becomes computationally intensive as the number of periods grows because of the interdependence of the demand periods. A Markovian approximation to the problem is developed that enables larger problems to be solved without losing much accuracy in the total cost calculations.

Looking back at the relationship between perturbed demand and penalty cost models, we see that the penalty cost should be a function of number of periods remaining. As the number of periods increases, generally the goodwill cost increases as well. With short horizons, the full negative effects of a missed sale cannot be felt. Over a long horizon, the goodwill seems to level out to a steady state goodwill cost. Knowledge of this cost allows more accurate estimates of penalty costs to be determined.

The key research implications of this paper are the model itself and the solution techniques. The model itself extends current perturbed demand literature by creating a new demand loss function and removes steady state assumptions from being considered. Various solution techniques are developed to handle the potentially non-convex nature of the cost function. The approximate solution estimated the implied penalty cost for the perturbed demand model. Our findings are similar to Liberopoulus et al. (2010), with the important exception that our work is applied to a finite horizon setting and stochastic demand. The implied penalty cost takes on a wide range of values unlike their work. The MP solution procedure makes evaluating the cost for longer horizon problems possible.

Chapter 5: Research Contribution and Future Work

5.1. Chapter 2 Contributions

Chapter 2 applied perturbed demand principals to an infinite horizon periodic review inventory model. Specifically, the research was focused on the non-seasonal, no fixed order cost case, with instantaneous order arrival. The optimal policy was proven to have a base-stock structure, where each demand deflation state has its own base stock value. Additionally, the policy was proven to converge to a steady state policy in the infinite horizon setting. These two propositions show that our perturbed demand model behaves similarly to penalty cost inventory models. While the optimal policy does not exactly match, they both have a base stock structure and converge to a steady state order policy in the infinite horizon.

Unlike the penalty cost version of the problem, there is no easy analytical solution to the problem. Thus, the optimal policy is calculated using a Markov decision process (MDP) where the demand deflation value is segmented into discrete values. Using a large number of discrete values for demand deflation makes the model more realistic but increases the size of the state space and consequently the solution time. The increased computational effort makes the use of approximations more attractive.

Multiple approximations were developed that worked well for a variety of test problems. The approximations all significantly decreased solution times without sacrificing accuracy. For the test problems used in Chapter 2, the mean error for any approximation was always under 0.25% from optimal. The increased solution speed allows problems with larger demand distributions and finer demand deflation values to be solved. These findings are in line with the penalty cost model literature; the optimal solution is computationally expensive to generate but approximations work well.

One approximation implemented in Chapter 2 was to force a single base-stock policy structure on the problem, which matches the penalty cost model optimal policy. This policy is much easier to implement than the optimal policy because it does not require demand deflation to be tracked explicitly in practice. Additionally, this policy performs well for many of the test problems. This indicates that using a penalty cost model is a close approximation to the perturbed demand reality as long as the correct penalty cost is used. By aligning the policies for both models, it is possible to determine an implied penalty cost from the perturbed demand reality. The implied penalty cost is generally within 20% of the revenue term. We consistently see that the revenue generated by the sale of a product has the greatest impact on the implied lost sales penalty cost.

5.2. Chapter 3 Contributions

Chapter 3 extended the periodic review model developed in chapter 2 to include order lead time. The inclusion of lead time changes the optimal policy structure and increases the difference between the perturbed demand and penalty cost model optimal policy structure. The instantaneous order model had an easy-to-characterize policy, but when an order lead time is included the order policy is not easily characterized by a traditional policy structure. The order quantity is a function of on-hand inventory and demand deflation value and is not easily characterized. For a given demand deflation value, the order quantity increases then levels off and finally decreases as on-hand inventory increases. This policy structure is unlike the structure of the penalty cost with lead time, with the main deviation occurring at low inventory states. The perturbed demand policy initially increases due to the potential loss of demand caused by stock-outs in the current period. As expected, stock-outs decrease from holding more inventory, the order quantity levels out, and then eventually decline as on-hand inventory increases. The quantity is also dependent on the demand deflation value, each level taking on a unique order quantity curve. The difference in policy structure indicates a significant difference between how penalty cost and perturbed demand models penalize retailers for stock-outs. As in Chapter 2, the optimal policy was calculated in Chapter 3 using MDP solution techniques. However, the solution times are long and the optimal policy is difficult to characterize and therefore easily implement.

Several approximations were developed in an attempt to avoid the computational difficulties associated with determining the optimal order policy. One approximation developed used insight gained from the optimal policy structure; here, the policy order is based on where the system is expected to be when the order arrives. The approximation estimates the demand deflation and on-hand inventory for when the order willarrive and orders stock based on those conditions. This approximation provided a near optimally performing policy for many of our research problems and requires little computational effort. An additional approximation forced the policy to take on a demand deflation dependent base stock level to match the optimal policy structure of the instantaneous order arrival case; however, the model did not always perform well.

Assuming a base stock policy structure for the lost sales penalty cost model is a common approximation, however the performance of the approximation improves as penalty cost is increased. As the cost of a stock-out increases, there is a greater willingness on the retailer's part to hold larger inventories. The optimal policy structure approaches a demand deflation dependent base stock policy as the on-hand inventory increase. Consequently, the base stock approximation works well for high revenue cases where retailers are willing to hold inventory to satisfy demand.. For low revenue problems retailers are not willing to hold inventory because it diminishes profits, thus using a base stock policy does not perform well. Finally, we implemented Morton's (1971) approximation method to the perturbed demand model which turned out to work the best. The approximation matched the optimal policy structure better than a base stock structure. Additionally, as in the single period problem, an implied penalty cost could be determined from this policy structure.

Future research can extend this model into more complex lead time situations. Some penalty cost models have stochastic lead time, and adding this complexity to the model may result in a very different optimal policy structure. Without retailers' knowing when the order will arrive, the policy could be dramatically different.

5.3. Chapter 4 Contributions

Chapter 4 took the model originally developed in the infinite horizon problem and applied it to a finite horizon setting. The finite horizon adds additional complexity to the order policy because of the end-of-horizon effects. As the horizon comes to an end, the cost of a stock-out decreases because there are not enough periods remaining to feel the full effects of lowered demand. While penalty cost models quickly approach their steady state order policy, end of horizon effects in perturbed demand models can last significantly longer. Multiple periods are needed to feel the full effect of lowered future demand in the perturbed demand setting while penalty costs are felt instantly. Consequently, actions taken at the start of the horizon have a much larger effect on the total cost than ones made later in the horizon.

The start of horizon ordering decision carries the most weight and makes searching for the best "penalty cost-like" policy difficult because the search space can be non-convex because ordering large quantities in the first period or ordering just a few in the first period can result in two local optimal policies. This makes finding the best single base stock each period policy difficult. The use of advanced search techniques allowed the researcher to find the optimal policy. We generated a policy using a genetic algorithm, gradient search with multiple starting points, and a heuristic. The curse of dimensionality came into play as the horizon increases making it difficult to use these solution methods to solve long horizon problems. A restricted observation Markov decision process (ROMDP) approximation was used to solve long horizon problems with the assumption that demand deflation can be discretized.

The ROMDP solution enables us to compare the perturbed demand model to a penalty cost model. We found that as the number of periods remaining increases, the base stock level increases as well. Eventually, it levels out as it approaches the infinite horizon policy, but it takes much longer to reach the infinite horizon policy than a traditional penalty cost model. The increase in the base stock level means that the implied penalty cost increases as a function of the number of periods remaining. Exponential smoothing of demand deflation spreads out effects of a stock-out over multiple periods. Thus, the longer the horizon, the more a stock-out can lower the demand, thereby increasing the implied penalty cost.

5.4. Future work

The model developed in this paper provides a good foundation for many future research opportunities. One such extension would be to look at the interactions between multiple products and demand deflation values. Several complications could be addressed such as a shelf space constraint where the total inventory for both products has to be less than a fixed quantity or correlated demand deflation between the products. Stocking out of one product may increase the demand of another, especially if they are substitutes for each other, or it may decrease demand because consumers switch all their purchasing to an alternative store. The interdependence between products will make tracking demand deflation difficult. The size of the state space increases because the inventory and demand deflation for each product must be tracked, doubling the number of state dimensions. The larger state space will make the solution of multiple product problems difficult; an alternative solution method such as simulation may need to be implemented to solve large problems.

Advertising is one tool retailers use to increase demand. In the perturbed demand model, advertising might translate to lessoning the effects of demand deflation. Customers that would ordinarily take their demand elsewhere can be convinced to remain loyal through advertising. The model would have to be relaxed to allow demand deflation to increase above 1 if enough is spent on advertising. This could be modeled as an increase in the decision space, expanded to both order quantity and advertising investment per period. When the advertising dollars are spent most likely will be a function of on-hand inventory and demand deflation. Buying demand deflation back may only be worth the investment when demand deflation drops below a set value and when the retailer expects to have enough inventory on hand to fulfill the increased demand.

Perturbed demand research has many potential opportunities for future study. The potential to extend the perturbed demand model to include complications that penalty cost models have already considered opens many avenues for future research. On the consumer behavior side of the model, a more accurate model to predict consumer actions would help perturbed demand models increase their prediction of reality. Our hope is that the model developed in this paper can be used to reinvigorate the perturbed demand inventory model literature and thereby positively affect the retail industry.

REFERNCES

Aksen, D. (2007). Loss of customer goodwill in the uncapacitated lot-sizing problem. *Computers & Operations Research, 34*(9), 2805-2823.

Coca-Cola Research Council/Andersen Consulting (1996), *Where to Look for Incremental Sales Gains: The Retail Problem of Out-of-Stock Merchandise,* The Coca-Cola Research Council, Atlanta, GA,.

Anderson, E. T., Fitzsimons, G. J., & Simester, D. (2006). Measuring and mitigating the cost of stockouts.*Management Science*, *52*(11), 1751-1763.

Arrow, K. J., Harris, T., & Marschak, J. (1951). Optimal inventory policy. *Econometrica*, *19*(3), 250-272.

Arrow, K. J., Karlin, S., & Scarf, H. (1958). *Studies in mathematical theory of inventory production.* Stanford California: Stanford University Press.

Balakrishnan, A., Pangburn, M. S., & Stavrulaki, E. (2004). "stack them high, let 'em fly": Lot-sizing policies when inventories stimulate demand. *Management Science*, *50*(5), 630-644.

Bell, C. E. (1970). Improved algorithms for inventory and replacedment-stocking problems. *SIAM Journal on Applied Mathematics*, *18*(3), 558-566.

Bellman, R., Glicksberg, I., & Gross, O. (1955). On the optimal inventory equation. *Management Science*, *2*(1), 83-104.

Bursk, E. C. (1966). View your customers as investments, *Harvard Business Review*, *44*(3), 91-94.

Caine, G. J., & Plaut, R. H. (2006). Optimal inventory policy when stockouts alter demand. *Naval Research Logistics, 23*(1), 1-13.

Campo, K., Gijsbrechts, E., & Nisol, P. (2000). Towards understanding consumer response to stock-outs. *Journal of Retailing*, *76*(2), 219-242.

Campo, K., Gijsbrechts, E., & Nisol, P. (2003). The impact of retailer stockouts on whether, how much, and what to buy. *International Journal of Research in Marketing*, *20*(3), 273-286.

Campo, K., Gijsbrechts, E., & Nisol, P. (2004). Dynamics in consumer response to product unavailability: do stock-out reactions signal response to permanent assortment reductions? *Journal of Business Research*, *57*(8), 834-843.

Chang, Y. S., & Niland, P. (1967). A model for measuring stock depletion costs, *Operations Research, 15*(3), 427-447.

Chow, C. S., & Tsitsiklis, J. N. (1988). An optimal multigrid algorithm for continuous state discrete time stochastic control. In *Proceeding 27th conference on decision and control* (pp. 1908-1912). IEEE.

Chow, C. S., & Tsitsiklis, J. N. (1991). An optimal one-way multigrid algorithm for discrete-time stochastic control. *IEEE Transactions on Automatic Control*, *36*(8), 898-914.

Corsten, D., & Gruen, T. (2003). Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail & Distribution Management*, *31*(12), 605-717.

Dana, J. D., & Petruzzi, E. (2001). Note: The newsvendor model with endogenous demand. *Management Science*, *47*(11), 1488- 1497.

Davis, L. B., Hodgson, T. J., King, R. E., & Wei, W. (2008). Technical Note: A computationally efficient algorithm for undiscounted markov decision processes with restricted observations. *Naval Research Logistics*, *56*(1), 86-92.

Ding, F. Y., Hodgson, T. J., & King, R. E. (1988). A methodology for computation reduction for specially structured large scale Markov decision problems. *European Journal of Operational Research*, *34*(1), 105-112.

Dorfman, R. B., & Sterner, P. D. (1954). Optimal advertising and optimal quality. *American Economic Review*, *44*(5), 826-836.

Ehrhardt, R. (1979). The power approximation for computing (s, S) inventory polices. *Management Science*, *25*(8), 777-786.

Ehrhardt, R., & Mosier, C. (1985). A revision of the power approximation for computing (s, S) policies. *Management Science*, *30*(5), 618-622.

Emmelhainz, M., & Stock, J. (1991). Consumer response to retail stock-outs. *Journal of Retailing*, *67*(2), 138-147.

Ernst, R., & Powell, S. G. (1995). Optimal inventory policies under service-sensitive demand. *European Journal of Operations Research*, *87*(2), 316-327.

Gardner, E. (1980). Inventory theory and the gods of Olympus. *Interfaces*, *10*(4), 42-45.

Gerchak, Y., & Wang, Y. (1994). Periodic-review inventory models with inventory-level-dependent demand. *Naval Research Logistics*, *41*(1), 99-116.

Hadley, H. G., & Whitin, T. M. (1963). *Analysis of inventory systems.* New Jersey: Prentice Hall.

Hanssmann, F. (1959). Optimal inventory location and control in production and distribution networks. *Operations Research*, *7*(4), 483-493.

Hill, T. W. Jr. (1976). A simple perturbed demand inventory model with ordering cost. *Management Science*, *23*(1), 38-42.

Howard, R. (1960). *Dynamic Programming and Markov Processes*. New York: Wiley.

Hue, W. T., Janakiraman, G., & Rusmechientong, J. A. (2009). Asymptotic optimality of order-up-to policies in lost sales inventory systems, *Management Science*, *55*(3), 404-420

Iglehart, D. (1963). Optimality of (s, S) policies in the infinite horizon dynamic inventory problem, *Management Science*, *9*(2), 259-267.

Janakiraman, G., & Roundy, R. O. (2004). Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Operations Research*, *52*(5), 765-803.

Johnson, E. L. (1968). On (s, S) policies. *Management Science*, *15*(1), 80-101.

Joins, J. A., Culbreth, C. T., & King, R. E. (1996). Manufacturing cell design: an integer programming model employing genetic algorithms. *IIE Transactions*, *28*(1), 69-85.

Karlin, S. (1958). One stage inventory models with uncertainty. In *Studies in the mathematical theory of inventory and production* (pp. 131-151). Stanford: Stanford University Press.

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *The Journal of Marketing*, *59*(2), 71-82.

Liberopoulus, G., Tsikis, I., & Delikouras, S. (2010). Backorder penalty cost coefficient “b”: What could it be? *International Journal of Production Economics*, *123*(2), 166-178.

Ma, W. M., & Lu, W. (2010). Forecasting Daily Cash Turnover of Bank with EWMA and SVR. *Advances in Neural Network Research and Applications*, 83-90.

Mittal, R., & Lasser, W. M. (1998). Why do customers switch? The dynamics of satisfaction versus loyalty. *Journal of Services Marketing*, *12*(3), 188-194.

Morton, T. E. (1971). The near-myopic nature of the lagged-proportional-cost inventory problem with lost sales. *Operations Research*, *19*(7), 1708 1716.

Nahmias, S. (1979). Simple approximations for a variety of dynamic leadtime lost-sales inventory models. *Operations Research*, *27*(5), 904-924.

Nandakumar, P., & Morton, T. E. (1993). Near myopic heuristic for the fixed-life perishability problem. *Management Science*, *39*(12), 1490-1498.

Oral, M., Salvador, S., Reisman, A., & Dean, B. V. (1972). On the evaluation of shortage cost for inventory control of finished goods. *Management Science*, *18*(6), B344-B351.

Raman, A. (1999). Managing inventory for fashion products. International Series in Operations Research and Management Science, 17, 789-806.

(2003). *Quantitative models for supply chain management.* In S. Tayur, R. Ganeshan & M. Magazine (Eds.), Norwell, MA: Kluwer Academic Publishers Group.

Roberts, D. (1962). Approximations to optimal policies in a dynamic inventory model. *Studies in Applied Probability and Management Science*. In K. Arrow, S. Karlin, & H. Scarf (Eds.), Stanford California: Stanford University Press.

Robinson, L. W. (1990). Appropriate Inventory policies when service affects future demand. *Working Paper #88-08,* Johnson Graduate School of Management, Cornell University.

Scarf, H. (1960). The optimality of (s, s) policies in dynamic inventory problem. In E. Sheshinski & Y. Weiss (Eds.),*Optimal pricing, inflation, and the cost of price adjustment* (pp. 49-56). MIT Press.

Schary, P. B., & Becker, B. W. (1978). The impact of stock-outs on market share: temporal effects. *Journal of Business Logistics*, *1*(1), 31-44.

Schneider, H., & Ringuest, J. (1990). Power approximation for computing (s, S) policies using service level. *Management Science, 36*(7), 822-834.

Schwartz, B. L. (1966). A new approach to stockout penalties. *Management Science*, *12*(12), B538-B544.

Schwartz, B. L. (1970), Optimal inventory policies in perturbed demand models. *Management Science*, *16*(8), B509-B518.

Sethi, S., & Cheng, F. (1993). Optimality of (s, S) policies in inventory models with Markovian demand. *Operations Research*, *45*(6), 931-939.

Tijms, H. (1972). Analysis of (s, S) inventory models. *Mathematical Centre Tracts.* Amsterdam: Mathematishce Centrum.

Urban, T. (1995). Inventory models with the demand rate dependent on stock and shortage levels. *International Journal of Production Economics*, *40*(1), 21-28.

Van Woensel, T., Van Donselaar, K., Broekmeulen, R., & Fransoo, J. (2007). Consumer responses to shelf out-of-stocks perishable products. *International Journal of Physical Distribution & Logistics Management*, *37*(9), 704-718

Veinott, A., & Wagner, H. (1965). Computing optimal (s, S) inventory polices. *Management Science*, *11*(5), 525-552.

Veinott, A. (1966). On the optimality of (s, S) inventory policies: new conditions and a new proof. *SIAM Journal on Applied Mathematics*, *14*(5), 1067-1083.

Verbeke, W., Farris, P., & Thurik, R. (1998). Consumer response to the preferred brand out-of-stock situation. *European Journal of Marketing*, *32*(11), 1008-1028.

Wagner, H., O'Hagan, M., & Lundh, B. (1965). An empirical study of exactly and approximately optimal inventory polices. *Management Science*, *11*(7), 690-723.

Walter, C. K., & Grabner, J. R. (1975). Stockouts cost models: Empirical tests in retail situations. *The Journal of Marketing*, *39*(3), 56-60.

Warren, B. P. (2011). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, New York: Wiley.

Wei, W. (2005). Quantifying shared information in a supply chain using decentralized Markov decision processes with restricted observations, Ph.D. dissertation. North Carolina State University, Raleigh, NC

Zhang, N. L., & Zhang, W. (2001). Speeding up the convergence of value iteration in partially observable Markov decision processes. *Journal of Artificial Intelligence Research*, *14,* 29-51.

Zheng, Y. (1991). A simple proof for optimality of (s, S) policies in infinite-horizon inventory system. *Journal of Applied Probability*, *28*(4), 802-810.

Zinn, W., & Liu, C. P. (2001). Consumer response to retail stockouts. *Journal of Business Logistics*, *22*(1), 49-71.

Zipkin, P. (2000). *Foundations of Inventory Management*. New York: McGraw-Hill.

Zipkin, P. (2008a). On the structure of lost-sales inventory models. *Operations Research*, *56*(4), 937-944.

Zipkin, P. (2008b). Old and new methods for lost-sales inventory systems. *Operations Research*, *56*(5), 1256 – 1263.

APPENDICES

APPENDIX A: Chapter 2

The policy is Base-stock policy when

L is convex and $\left[\begin{array}{l}\displaystyle\int_0^{\frac{y}{\alpha_n}} f_{n-1}\left(y-\alpha_n\xi_n, \lambda+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n \\ \displaystyle+\int_{\frac{y}{\alpha_n}}^{\infty} f_{n-1}\left(0, \lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n\end{array}\right]$ is convex (Scarf 1960 pp 53)

Show L is convex:

$$L(y \mid \alpha)=\int_0^{\frac{y}{\alpha}} h(y-\alpha\xi) g(\xi) d\xi-\int_0^{\frac{y}{\alpha}} r\alpha\xi g(\xi) d\xi-ry\left(1-G\left(\frac{y}{\alpha}\right)\right)$$

$$\frac{dL(y \mid \alpha)}{dy}=(h+r) G\left(\frac{y}{\alpha}\right)-r$$

$$\frac{d^2 L(y \mid \alpha)}{d^2 y}=\frac{1}{\alpha}(h+r) g\left(\frac{y}{\alpha}\right)$$

The second derivative is always positive because the probability density function is always positive.

Show $\left[\begin{array}{l}\displaystyle\int_0^{\frac{y}{\alpha_n}} f_{n-1}\left(y-\alpha_n\xi_n, \lambda+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n \\ \displaystyle+\int_{\frac{y}{\alpha_n}}^{\infty} f_{n-1}\left(0, \lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n\end{array}\right]$ is convex

Because L is convex then f_1 is convex

Derivative of just the top integral

$$\int_0^{\frac{y}{\alpha_n}} f_{n-1}\left(y-\alpha_n\xi_n, \lambda+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n$$

$$\frac{1}{\alpha_n} f_{n-1}\left(0, \lambda+(1-\lambda)\alpha_n\right) g\left(\frac{y}{\alpha_n}\right) - c\omega G\left(\frac{y}{\alpha_n}\right)$$

Derivative of the bottom integral

$$\int_{\frac{y}{\alpha_n}}^{\infty} f_{n-1}\left(0, \lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right) g\left(\xi_n\right) d\xi_n$$

$$-\frac{1}{\alpha_n} f_{n-1}\left(0, \lambda+(1-\lambda)\alpha_n\right) g\left(\frac{y}{\alpha_n}\right) + \int_{\frac{y}{\alpha_n}}^{\infty} \left[\begin{array}{l} -(h+r)\lambda\frac{\beta}{\alpha_n\xi_n} \displaystyle\int_0^{\frac{y_{n-1}}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)}} \xi_{n-1} g\left(\xi_{n-1}\right) d\xi_{n-1} \\ +\omega\displaystyle\int_0^{\infty} \frac{df_{n-2}\left(x_{n-2}, \alpha_{n-1}\right)}{dy} g\left(\xi_{n-1}\right) d\xi_{n-1} \end{array}\right] g\left(\xi_n\right) d\xi_n$$

Combined them

$$-c\omega G\left(\frac{y}{\alpha_n}\right) - (h+r)\lambda\frac{\beta}{\alpha_n}\int_{\frac{y}{\alpha_n}}^{\infty}\left[\frac{1}{\xi_n}\int_0^{\frac{y_{n-1}}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)}} \xi_{n-1} g\left(\xi_{n-1}\right) d\xi_{n-1}\right] g\left(\xi_n\right) d\xi_n$$

$$+\omega\int_{\frac{y}{\alpha_n}}^{\infty}\left[\int_0^{\infty}\frac{df_{n-2}\left(x_{n-2}, \alpha_{n-1}\right)}{dy} g\left(\xi_{n-1}\right) d\xi_{n-1}\right] g\left(\xi_n\right) d\xi_n$$

2nd Derivative

$$-c\omega g\left(\frac{y}{\alpha_n}\right)\frac{1}{\alpha_n}+(h+r)\lambda\frac{\beta}{\alpha_n}\frac{1}{y}g\left(\frac{y}{\alpha_n}\right)\int_0^{\frac{y_{n-1}}{(\lambda+(1-\lambda)\alpha_n)}}\xi_{n-1}g(\xi_{n-1})\,d\xi_{n-1}$$

$$+(h+r)\lambda^2y_{n-1}^2\frac{\beta^2}{\alpha_n^2}\int_{\frac{y}{\alpha_n}}^{\infty}\frac{1}{\xi_n^2}\frac{1}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)^3}g\left(\frac{y_{n-1}}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)}\right)g(\xi_n)\,d\xi_n$$

$$+\omega\int_{\frac{y}{\alpha_n}}^{\infty}\left[\int_0^{\infty}\frac{d^2f_{n-2}(x_{n-2},\alpha_{n-1})}{d^2y}g(\xi_{n-1})\,d\xi_{n-1}\right]g(\xi_n)\,d\xi_n$$

When the L function is added the new second derivative is

$$(h+r-c\omega)\frac{1}{\alpha}g\left(\frac{y}{\alpha}\right)+(h+r)\lambda\frac{\beta}{\alpha_n}\frac{1}{y}g\left(\frac{y}{\alpha_n}\right)\int_0^{\frac{y_{n-1}}{(\lambda+(1-\lambda)\alpha_n)}}\xi_{n-1}g(\xi_{n-1})\,d\xi_{n-1}$$

$$+(h+r)\lambda^2y_{n-1}^2\frac{\beta^2}{\alpha_n^2}\int_{\frac{y}{\alpha_n}}^{\infty}\frac{1}{\xi_n^2}\frac{1}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)^3}g\left(\frac{y_{n-1}}{\left(\lambda\left(1-\beta+\frac{y\beta}{\alpha_n\xi_n}\right)+(1-\lambda)\alpha_n\right)}\right)g(\xi_n)\,d\xi_n$$

$$+\omega\int_{\frac{y}{\alpha_n}}^{\infty}\left[\int_0^{\infty}\frac{d^2f_{n-2}(x_{n-2},\alpha_{n-1})}{d^2y}g(\xi_{n-1})\,d\xi_{n-1}\right]g(\xi_n)\,d\xi_n$$

The function is convex when $r+h>\omega c$ and f_{n-2} is convex this pattern will repeat until f_1 which is convex.

Penalty Cost

Penalty cost is continuous and monotonic increasing.

$$P(S \mid \alpha) = -\int_0^{\frac{S}{\alpha}} r\alpha\xi g(\xi)\,d\xi - rS\left(1 - G\left(\frac{S}{\alpha}\right)\right)$$

$$\frac{dP(S \mid \alpha)}{dS} = -r + rG\left(\frac{S}{\alpha}\right)$$

$$\frac{d^2P(S \mid \alpha)}{d^2S} = rg\left(\frac{S}{\alpha}\right)\frac{1}{\alpha}$$

$$T_\alpha(\alpha, S, \xi) = \begin{cases} \lambda + (1-\lambda)\alpha & S \geq \xi \\ \lambda\left(1 - \frac{(\xi - S)\beta}{\xi}\right) + (1-\lambda)\alpha & S < \xi \end{cases}$$

$$\frac{dT_\alpha(\alpha, S, \xi)}{dS} = \begin{cases} 0 & S \geq \xi \\ \lambda\left(\frac{\beta}{\xi}\right) & S < \xi \end{cases}$$

Revenue and demand deflation increases as a function of *S*. Penalty cost is the cost of missing sales therefore ordering fewer will increase penalty and decrease the demand deflation.

APPENDIX B: Chapter 4

Two-period problem

$$\frac{df_2\left(x_2,\alpha_2\right)}{dS_2}=c-r+\left(h+r-c\right)G\left(\frac{S_2}{\alpha_2}\right)-\omega\left(h+r\right)\frac{\lambda\beta}{\alpha_2}\int_{\frac{S_2}{\alpha_2}}^{\infty}\left[\frac{1}{\xi_2}\int_{0}^{\frac{S_1}{\left(\lambda\left(1-\frac{\left(\alpha_2\xi_2-S_2\right)\beta}{\alpha_2\xi_2}\right)+\left(1-\lambda\right)\alpha_2\right)}}\xi_1 g\left(\xi_1\right)d\xi_1\right]g\left(\xi_2\right)d\xi_2$$

$\beta \le 1$

$$\frac{df_2\left(x_2,\alpha_2\right)}{dS_1}=\omega\left\{\begin{array}{l}c-r+\left(h+r\right)G\left(\dfrac{S_1}{\lambda+\left(1-\lambda\right)\alpha_2}\right)G\left(\dfrac{S_2}{\alpha_2}\right)\\[2ex] +\displaystyle\int_{\frac{S_2}{\alpha_2}}^{\infty}\left[\left(h+r\right)G\left(\frac{S_1}{\left(\lambda\left(1-\dfrac{\left(\alpha_2\xi_2-S_2\right)\beta}{\alpha_2\xi_2}\right)+\left(1-\lambda\right)\alpha_2\right)}\right)\right]g\left(\xi_2\right)d\xi_2\end{array}\right\}$$

$$\frac{df_2\left(x_2,\alpha_2\right)}{dS_2}=c-r+\left(h+r-c\right)G\left(\frac{S_2}{\alpha_2}\right)-\omega\left(h+r\right)\frac{\lambda\beta}{\alpha_2}\int_{\frac{S_2}{\alpha_2}}^{\frac{S_2\beta}{\alpha_2\left(\beta-1\right)}}\left[\frac{1}{\xi_2}\int_{0}^{\frac{S_1}{\left(\lambda\left(1-\frac{\left(\alpha_2\xi_2-S_2\right)\beta}{\alpha_2\xi_2}\right)+\left(1-\lambda\right)\alpha_2\right)}}\xi_1 g\left(\xi_1\right)d\xi_1\right]g\left(\xi_2\right)d\xi_2$$

$\beta > 1$

$$\frac{df_2\left(x_2,\alpha_2\right)}{dS_1}=\omega\left\{\begin{array}{l}c-r+\left(r+h\right)\left(G\left(\dfrac{S_1}{\lambda+\left(1-\lambda\right)\alpha_2}\right)+1-G\left(\dfrac{S_2\beta}{\alpha_2\left(\beta-1\right)}\right)\right)\\[2ex] +\displaystyle\int_{\frac{S_2}{\alpha_2}}^{\frac{S_2\beta}{\alpha_2\left(\beta-1\right)}}\left[\left(h+r\right)G\left(\frac{S_1}{\left(\lambda\left(1-\dfrac{\left(\alpha_2\xi_2-S_2\right)\beta}{\alpha_2\xi_2}\right)+\left(1-\lambda\right)\alpha_2\right)}\right)\right]g\left(\xi_2\right)d\xi_2\end{array}\right\}$$