5-4-2016

# Optimizing Inventory for Profitability and Order Fulfillment Improvement

Liu Yang
*University of Missouri-St. Louis*, lygm8@mail.umsl.edu

# OPTIMIZING INVENTORY FOR PROFITABILITY AND ORDER FULFILLMENT IMPROVEMENT

Integrating Inventory Classification and Control Decisions under Non-Stationary Demand For Profit Maximization
&
Integrating Inventory Classification and Control Decisions to Maximize Order Fulfillment Measures

Liu (Dorothy) Yang
M.B.A, Durham University Business School
B.Eng., Computer Science & Applications, Xi'an Jiaotong University

A Dissertation Submitted to The Graduate School at the University of Missouri-St. Louis in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration with an emphasis in Logistics and Supply Chain Management

May 2016

<u>Advisory Committee</u>

Haitao Li, Ph.D.
Chairperson

James F. Campbell, Ph.D.

Donald C. Sweeney II, Ph.D.

Keith Womer, Ph.D.

Mitch Millstein, Ph.D.

# Dedication

To my mother, Li Li, who used to teach higher mathematics and physics in the community college and retired as an Associate Professor. I know you would be very proud to see I earned a doctoral degree. And I'm equally proud to follow your path to pursue a career in academia. I love you forever.

To William and Fanyu, my family and my life.

# Acknowledgements

My sincere thanks to my advisor Dr. Haitao Li for his patience, guidance and full support throughout my dissertation research. I'm grateful for the opportunities of working with you in the past few years, which were both fun and equipped me with essential skills in operations optimization.

My heartiest thanks to Drs. James Campbell and Donald Sweeney for guiding and inspiring me in various ways. I'm deeply indebted to you for always being there to provide me with your critical advices; especially in the past year, you have guided me through an important journey, for dissertation and beyond.

I owe my thanks to Dr. Keith Womer for offering me many valuable comments and suggestions since I started my dissertation. I wish to convey my thanks to Dr. Mitch Millstein for initiating the class project, listening to my ideas and providing thoughtful feedback.

I would like to thank Dr. Ray Mundy for providing a nice office space for me at the Center for Transportation Studies for my entire PhD study. I would also like to thank Karen Walsh for having always been supportive. Things in her hands never go wrong.

# Abstract

Despite the extensive research on inventory management, few studies have investigated the optimization of inventory classification and control policies for maximizing the net present value of profit and order fulfillment performance. This dissertation aims to fill the gaps, and consists of two main essays. Essay One (Chapter 1) presents a new multi-period optimization model to explicitly address nonstationary demand, arbitrary review periods, and SKU-specific lead times, with the objective of maximizing the net present value of profit. A real-world application and computational experiments show that the optimal dynamic inventory classification and control decisions obtained from the model significantly reduce both safety stock and base stock levels compared to a multi-criteria inventory classification scheme and the traditional ABC approach. Essay Two (Chapter 2) examines two order-based fulfillment performance measures: the order fill rate, defined as the percentage of orders that are completely filled from available inventory; and the average customer-order fill rate, defined as the mean percentage of total units in a customer order that can be filled from on-hand inventory. Novel optimization models are developed to maximize the order fulfillment performance. Computational results indicate that a commonly used item-based measure in general does not adequately indicate order-based performance, and the tradeoffs between profit and order-based measures vary with inventory investment. This research contributes to the existing literature by providing new approaches to optimize inventory classification and control policies with various performance criteria. It also provides practitioners with a viable way to manage inventory with nonstationary demand, general review periods and lead times, and further allows companies to quantity the tradeoffs of different performance measures.

# Table of Contents

# Table of Figures

# Table of Tables

# Chapter 0 : Overview of the Thesis

Profitability and customer satisfaction are the paramount goals of any for-profit business aiming for long-term success. Inventory, which is generally the second largest expenditure for a company, has a direct impact on both the company's bottom line and the customer service level. This research provides manufacturers, distributors, and retailers with a viable way to effectively and efficiently manage their inventory so that the company's profitability can be improved and the customer demand can be better satisfied.

The main body of this dissertation is comprised of two essays and each of them is presented as a separate chapter. Chapter 1 includes Introduction, Literature Review, Model Development, Real World Application, Computational Experiments, and Conclusions. Chapter 2 has the similar structure except for the section of Real World Application.

Chapter 1 focuses on the financial aspect of inventory management by presenting a new multi-period inventory optimization model that explicitly addresses nonstationary demand, arbitrary review periods, and SKU-specific lead times, with the objective of maximizing the net present value of profit. The model is evaluated against an "advanced" multi-criteria inventory classification scheme through a real-world case, and is also compared with the more commonly used ABC approach through extensive computational experiments. The impacts of key inventory parameters (including demand pattern, demand variability, inventory holding cost, lead time, and cost of capital) on model performance are examined.

Chapter 2 focuses on the customer service aspect of inventory management by examining two order-based fulfillment performance measures: the order fill rate, defined as the percentage of orders that are completely filled from available inventory; and the average customer-order fill rate, defined as the mean percentage of total units in a customer order that can be filled from on-hand inventory. Two optimization models are developed to maximize the order fulfillment performance measures by endogenously selecting the appropriate inventory classes to which SKUs are assigned, subject to the inventory capital constraint and the minimum profit requirement. An extensive sensitivity analysis is performed over different levels of inventory investments, profits, and order fill rates to gain insights on the tradeoff between order-based measures and profit, the relationship of the two order fulfillment measures, and their relationships with item fill rate.

Chapter 3 summarizes the model development framework, key findings, major contributions, and future research opportunities.

# Chapter 1 Integrating Inventory Classification and Control Decisions under Non-Stationary Demand for Profit Maximization

## 1 Introduction

Manufacturers, distributors and retailers often have to deal with a large variety of items, known as stock-keeping units (SKUs). A large industrial supply wholesaler on average has over 100,000 different items in inventory (First Research, 2014). In the retail sector, a mid-size supermarket may offer about 20,000 SKUs and a hypermarket carries 80,000 items or more (Goic et al., 2015). To efficiently and effectively manage inventory with such a large number of SKUs, a common practice is to group the items into a limited number of classes and then set common target service levels and inventory policies per class (Teunter et al., 2010; van Kampen et al., 2012). This enables companies to specify, monitor and control inventory performance for SKU classes rather than for each SKU individually.

Because stock control methods are primarily based on SKU classes, inventory classification has a direct impact on inventory cost and service levels, which are essential to the company's short-term profitability and long-term customer relationships. A large body of existing literature (e.g., Soylu and Akyol, 2014; Hatefi et al., 2014; Molenaers et al., 2012, Hadi-Vencheh; 2010; Ramanathan, 2006; Ernst and Cohen; 1990) has studied the classification process and proposed various approaches and techniques to classify inventories based on pre-defined criteria, such as demand volume, unit cost, lead time, etc. The vast majority of this classification research focuses exclusively on the classification process of developing SKU ranking methods, leaving unresolved the fundamental question of how inventory performance measures can be improved through SKU classification (Mohammaditabar et al., 2012; Lajili et al., 2012). In parallel, another line of inventory research (e.g., Neale and Willems, 2009; Stanford and Martin, 2007; Cormier and Gunn, 1996) has focused primarily on inventory cost minimization for SKU groups that are predetermined upon the classical or extended ABC analysis with arbitrarily assigned service levels. These research works overlook the possibility that the ABC classification itself may be a barrier for superior inventory performance (Teunter et al., 2010). This disconnect between inventory classification and performance measures raises the question of whether the current inventory management practices are most appropriate. Companies generally have a better chance to achieve optimal performance when adopting a holistic view where various processes and strategies are simultaneously evaluated and aligned (Closs et al., 2009).

A handful of research works (e.g., Teunter et al., 2010; Millstein et al., 2014) that have attempted to integrate inventory classification with performance measures all assume stationary demand. In the real world, demand is rarely stationary due to the seasonality, product life cycles, and business cycles. For example, Neale and Willems (2009) note that two thirds of the annual demand for Xbox video game consoles occur in the last quarter. A smartphone has an average product life cycle of 8 to 10 months (Brightstar Intelligence, 2014) with the demand fluctuating as it goes through the introduction, growth, maturity, and decline stages. Furthermore, many companies observe month-end, quarter-end, or year-end surges in sales because of the business review cycles. Nonstationary demand, characterized by means and variances both changing over time, is the norm in reality (Silver, 2008). Managing inventory under nonstationary, stochastic demand can be arduous. The inherent complexity of nonstationary problems complicates the analysis and model development, and may present significant computational challenges (Neale and Willems, 2009). Companies often adopt stationary inventory policies even in a nonstationary demand environment because of their relative simplicity. Tunc et al. (2011) however, demonstrate that in most cases, the cost of using a stationary ($s$, $S$) policy is significantly higher than the optimal non-stationary policy.

The inventory performance measure of interest in this research is the expected profit. The consideration of profit maximization instead of the more commonly used cost minimization is due to the following arguments. First, to businesses, the ultimate financial measure is profit, and lower cost does not necessarily result in higher profit. In fact, the inventory parameters generated from cost minimization and those from profit maximization may drastically differ unless all items have the same profit margin, which is unlikely for most companies. To illustrate, consider the case of minimizing inventory holding cost subject to the service level requirement. Intuitively, the model would encourage holding a minimum level of safety stock for high cost SKUs; but if these SKUs are also high profit items, the potential loss on sales due to stockouts can be significant. For companies that carry highly substitutable products, if a customer order cannot be served immediately from stock, it is often lost to competitors. Second, profit is the best measure of the tradeoff strategy of inventory investment and lost sales. Profit in this research is represented by subtraction of the inventory holding cost from the expected gross profit which takes item or unit fill rate into account. To an extent, optimizing profit is equivalent to balancing inventory holding cost and unit fill rate.

To support businesses to develop an effective inventory system that optimizes financial performance and to address the aforementioned gaps in the inventory literature, this chapter proposes a multi-period, integrated inventory model that simultaneously classifies SKUs and determines inventory control policies in response to the change of expected demand and forecast uncertainty across the planning horizon, with the objective to maximize the net present value (NPV) of profit. The model takes into account the time value of money, and allows an arbitrary review period and SKU-based inventory holding cost and lead time. The proposed model is evaluated against a multi-criteria inventory classification (MCIC) scheme through a real-world case, considering that the MCIC approach has been extensively studied in the literature and is regarded as an advanced classification scheme by practitioners. The optimization model is also compared with the more commonly used ABC approach, and the individual and interactive effects of various inventory parameters on the model performance are investigated through comprehensive computational experiments.

The remainder of this chapter is organized as follows: the next section provides a review of inventory literature, specifically inventory classification and modeling of nonstationary demand, and identifies the literature gaps that motivate this study. Section 3 presents the design and underlying assumptions of the multi-period inventory optimization model. In Section 4, an empirical investigation is conducted based on a real-world case, followed by Section 5 which compares the optimization model with the ABC approach through comprehensive computational experiments and addresses the question of under what scenarios the proposed model significantly outperforms the traditional approach. Section 6 summarizes the findings and provides managerial insights.

# 2 Literature Review

## 2.1 Inventory Classification

The primary reason for applying inventory classification is that the number of SKUs is too large for companies to implement a specific inventory control policy for each item (Ernst and Cohen; 1990). Classification has been intensively studied in the inventory management literature. According to the number of criteria used for classification, the existing research comprises two classification methods: single-criterion inventory classification and multi-criteria inventory classification. One of the earliest and most widely used single-criterion inventory classification methods is the ABC analysis that is developed upon the Pareto principle and uses either the demand volume or the demand value to group SKUs (Syntetos et al., 2009). Its underlying rationale is that a large portion of the company's sales comes from a small percentage of its products (known as the 80-20 rule), and the high sales products should be managed most intensively. Typically, class A items are few in number but account for a significant percentage of the annual sales, class C items may consist of over half of the SKUs but only contribute a small percentage of the sales, and class B items fall in between. Determining the cutoff value for each class is more an arbitrary process, with little visibility on the ultimate effects these values may have on the overall inventory cost (Stanford and Martin, 2007). Whether A items deserve a higher stock availability or C items should be given a higher service level receives differing views in the literature (Viswanathan and Bhatnagar, 2005). On the one hand, some textbooks (e.g., Knod and Schonberger, 2001; Chase et al., 2001) suggest to keep the highest inventory level for C items and a lower level for A items because A items incur high inventory carrying cost and C items are not worth the effort of frequent review and replenishment. On the other hand, others (e.g., Stock and Lambert, 2001; Ballou, 2003) believe that A items are most critical to the company's survival and hence should receive the highest service level to avoid stockouts. Teunter et al. (2010) argue that the fundamental reason for such discrepancy is that the traditional ABC classification has not been developed from an inventory cost perspective.

Recently, many scholars have advocated the use of multi-criteria classification methods, where different SKU-based characteristics such as price, demand uncertainty, lead time, criticality, etc. are combined to rank and group SKUs. Various judgmental and statistical methodologies have been proposed, including analytic hierarchy process (AHP) by Flores et al. (1992), case-based distance modeling by Chen et al. (2008), cluster analysis by Ernst and Cohen (1990), decision trees by Porras and Dekker (2008), and heuristic methods such as artificial neural networks by Partovi and Anandarajan (2002). Among all these methods, the most notable one is the weighted linear optimization model developed by Ramanathan (2006), which considers four criteria, namely, annual sales revenue, unit cost, product criticality, and lead time. The model that he proposed automatically generates a set of criterion weights for each item such that its performance score can be maximized, subject to the constraints that the weighted sum, computed using the same set of weights, must be less than or equal to one for all the other items. A major advantage of Ramanathan's model is that it offers an objective approach for MCIC by eliminating the impact of subjectivity involved in the AHP method when rating the criteria and the inventory items. The model has been extended and improved by a number of researchers, including Zhou and Fan (2007), Ng (2007), Hadi-Vencheh (2010), and

Soylu and Akyol (2014). All these research works focus on the development of pure SKU ranking methods. Lajili et al. (2012) conduct a follow-up investigation of their effects on inventory cost and conclude that a seemingly good classification method does not guarantee superior inventory performance. Moreover, while the criteria considered for classification are believed to have an impact on inventory performance, the extent to which each criterion may affect performance and whether interactive effects may exist have not been discussed in the classification literature.

To connect inventory classification with performance measures, Teunter et al. (2010) proposed a new approach based on the objective of minimizing total inventory cost, subject to the constraint of meeting a required average fill rate. Their classification criterion is computed by taking into account demand volume, inventory holding cost, and order quantity of each SKU. The authors construct a non-linear optimization model to simultaneously classify SKUs and determine optimal cycle service levels for each group. Their empirical investigations demonstrate the inventory performance superiority of such an approach as compared to the traditional ABC method. SKU criticality, described in the shortage cost, is considered in criterion development, but not included in their empirical testing, due to the difficulty of quantifying shortage cost in practice.

Mohammaditabar et al. (2012) include two objectives in their optimization function: minimizing the inventory cost composed of inventory holding cost and inventory replenishment cost, and at the same time, minimizing dissimilarity of items classified in the same class through a pre-defined dissimilarity index. Each SKU group is assumed to have a unique order interval, hence the model looks to simultaneously classify SKUs and decide the order interval for each class. The model is non-linear and solved by simulated annealing. Service levels and lost sales are excluded from the model.

Tsai and Yeh (2008) use another heuristic method, particle swarm optimization, to classify inventory based on a specific objective or multiple objectives, including inventory cost minimization, inventory turnover ratio maximization, and SKU demand correlation maximization. They also use the order interval as the distinction between inventory groups, and their total cost minimization function reflects a tradeoff between inventory holding cost and ordering and setup cost. Neither service level nor lost sales is taken into account, as the model assumes instantaneous replenishment and known demand rates that are spread evenly throughout the year, which rarely hold in reality.

Millstein et al. (2014) develop a mixed integer linear programming (MILP) model to simultaneously make inventory classification and service level decisions, and allocate inventory capital across SKU classes in a way to maximize profitability. However, it is not a truly profit maximization model because inventory holding costs are not considered. In addition, the model is limited to stationary demand, assumes lead time is fixed to one time period for all SKUs, and is not designed to cope with an arbitrary review period or take into account the time value of money. Similar to many of the stationary models, the structure of the model itself in Millstein et al. (2014) prohibits it from being executed in a rolling horizon fashion in the face of nonstationary demand.

This research presents a new model to explicitly address nonstationary demand, arbitrary review periods, and SKU-specific lead times, with the objective of maximizing the net present value (NPV) of profit.

## 2.2 Nonstationary Demand

Nonstationary demand is characterized by means and variances that may both change over time, in contrast to stationary demand that has a constant long-term mean and a constant variance independent of time. Seasonal demand for products such as ski equipment, as described in Figure 1.1, is a typical nonstationary demand pattern, which has low mean demand during spring and summer but high mean demand in autumn and winter. Another common demand pattern in the business-to-business environment is intermittent demand as shown in Figure 1.2 due to batch ordering and other factors, which is also nonstationary. In the real world, the most practical demand patterns are nonstationary (Silver, 2008).

*Figure 1.1: Seasonal demand*

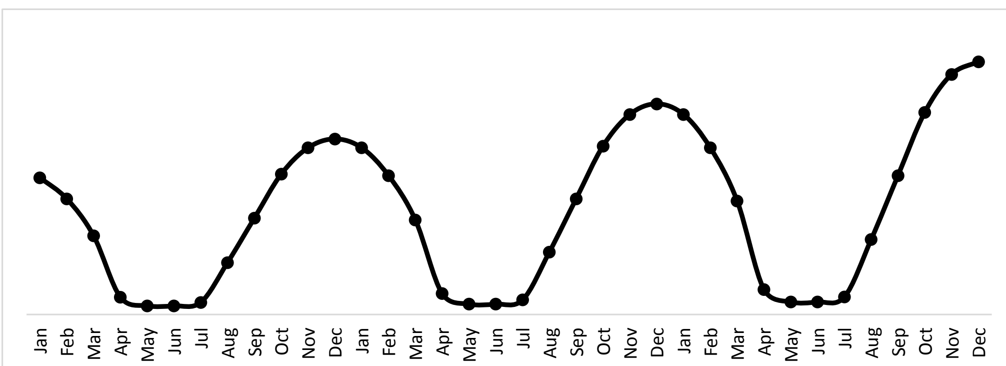
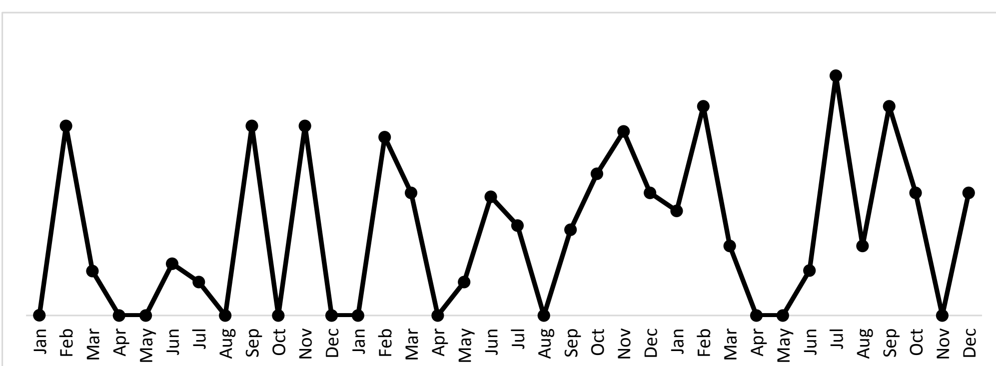



*Figure 1.2: Intermittent demand*



Graves (1999) states that a “major theme in the continuing development of inventory theory is to incorporate more realistic assumptions about product demand into inventory models”. The inventory literature that has taken into account nonstationary demand can roughly be classified into two lines of work: one analyzing the effects that a particular demand process and a forecast technique have on the inventory policies by addressing both demand forecast and inventory modelling; and the other focusing exclusively on inventory policy setting and inventory performance optimization.

In a nonstationary demand environment, the optimal inventory control policy is often time varying (Shang, 2012). With fixed ordering costs, Scarf (1959) has proven that the time-based ($s_t$,$S_t$) policy is optimal. Under the ($s_t$,$S_t$) policy, if the inventory position at the beginning of the period $t$ is lower than or equal to the reorder point $s_t$, a replenishment order is triggered to increase the inventory to a target level $S_t$. The inventory position $s_t$ and the target inventory level $S_t$ remain constant when demand is stationary. When ordering cost can be neglected, Karlin (1960) has proven that a time-varying base-stock policy is optimal for a single stage system in terms of cost minimization.

Graves (1999) investigates the adaptive base-stock policy for a single-item inventory system in which the demand process can be modeled as an integrated moving average. He demonstrates that the relationship between safety stock requirements and replenishment lead times in the case of nonstationary demand behaves dramatically differently from that of stationary demand. Kurawarwala and Matsuo (1996) focus on demand forecasting and inventory management of short life-cycle products by first constructing a seasonal trend growth model to forecast monthly demand over the entire life cycle of the product, and then use these demands as input to a finite-horizon inventory model. Their research is based on products that have a very short life cycle (e.g., one year) and a relatively long lead time (e.g., six months or longer). They determine production and inventory policies for the entire life cycle before the product is introduced, and these decisions made in advance are not expected to change in the later stage. Obviously, the accuracy of the demand forecast is particularly critical to the effectiveness of the inventory model in such a problem. The authors paid special attention to the forecasting techniques as data required by traditional time-series models do not exist for new products. In a similar vein, Treharne and Sox (2002) consider the problem of nonstationary demand with partial information observed, i.e., the actual demand is observed but the probability distribution of the demand in each period is determined by the state of a Markov chain. They suggest that a state-dependent base-stock policy is optimal for a single-stage system with a Markov-modulated demand process.

The aforementioned works and a number of other research articles (e.g., Chen and Song, 2001; Lovejoy, 1990) all consider a single-item inventory system and leverage the critical fractile (the ratio of the cost of being understocked to the total costs of being either overstocked or understocked) of the newsvendor problem to develop optimal policies. That is, to minimize the expected cost, the probability of not stocking out (i.e., cycle service level) should be equal to $b/(h+b)$, in which $b$ is the penalty cost or shortage cost per unit of stockout, and $h$ is the unit holding cost. There are two issues with this simple structure: a) for a multi-period model, it needs to allow the order quantity to be negative (Graves, 1999). That is, if the desired inventory level of this period $(t)$ calculated upon the critical fractile is less than the on-hand inventory at the end of the last period $(t-1)$, the order quantity would become negative, which is obviously not practical; and b) it is not applicable to a multi-item inventory system facing a limited inventory capital, because the critical fractile is for computing the optimal inventory level of individual SKUs and does not consider any capital constraint. Both issues may need to utilize dynamic programming to find optimal solutions.

A second line of research tackling nonstationary demand explicitly focuses on inventory policy modelling and directly utilizes the demand process information generated by forecasting models. Ettl et al. (2000) and Neale and Willems (2009) are indicative of this approach. Ettl et al. (2000) minimize the total expected inventory capital of a multi-stage inventory system by modeling each stocking location as an infinite-server queue operating under a base-stock control policy, subject to the end-customer service level requirements. They use discretized time units to address nonstationary demand and assume that all demands follow a normal distribution, but the mean and the standard deviation may vary with time. To incorporate the latest available demand information, they adopt a rolling-horizon approach in which the optimization is performed for each time period. Similarly, the method which Neale and Willems (2009) used to model nonstationary demand is to divide the planning horizon into a set of phases. The time span of each phase can vary and be as small as necessary, so that within each phase the demand can be characterized with an expected value and a standard deviation (i.e., forecast error) measured by statistics such as root mean squared error (RMSE). The authors assume that every stage in the supply chain operates in a common periodic-review, make-to-stock environment, and promises to 100% fulfill the downstream demands within an agreed service time. The proposed nonlinear optimization model aims to identify the optimal placement of safety stock in a multi-echelon system and the level of safety stock such that the total supply chain inventory holding costs are minimized, subject to end-customer service constraints. The authors argue that the average safety stock over the planning horizon at a location can be a concave function of the sum of the replenishment lead time and the order cycle time under certain conditions; hence the proposed models could be solved approximately by using spanning trees or the algorithms for general acyclic networks within a reasonable time.

The approach used in this research to incorporate nonstationary demand is most similar to the model developed by Neale and Willems (2009): the planning horizon is split into different phases in which both the mean and the variance of the demand may differ; the model optimizes the entire planning horizon, but it also allows inventory parameters to be updated in a rolling horizon fashion as new information becomes available. A number of key differences exist as follows. First and foremost, Neale and Willems (2009) exclusively address inventory policy setting and treat SKU classification and associated service levels as given, whereas my research integrates classification and inventory policy setting in such a way that inventory performance can be truly optimized. A major pitfall of separating classification from inventory modelling is that classification itself may prevent superior inventory performance. Second, the model structure differs. Previous inventory optimization models, including Neale and Willems (2009) and Ettl et al. (2000), are mostly nonlinear and thus face significant computational challenges in finding optimal solutions. The models formulated in this research are MILP models produced by discretizing cycle service level (CSL) into a finite number of levels. Finally, the inventory performance measures addressed are different. The existing inventory research mostly focuses on minimizing inventory costs comprised of ordering cost, inventory holding cost and in some cases backordering cost (Silver, 2008), whereas my research looks to maximize expected profit, echoing practical interest of businesses.

## 2.3 Summary of Literature Review

There is limited research that addresses nonstationary demand, very few articles investigate inventory classification and policy setting concurrently, and almost none of the authors aim to optimize profit, especially the NPV of profit. The integration of all the above aspects in one model constitutes a major contribution of this research, in response to the fact that companies typically face all these issues at the same time. Table 1.1 compares the features of this research with that of the works that are most relevant.

*Table 1.1: Comparison of features of different inventory research*

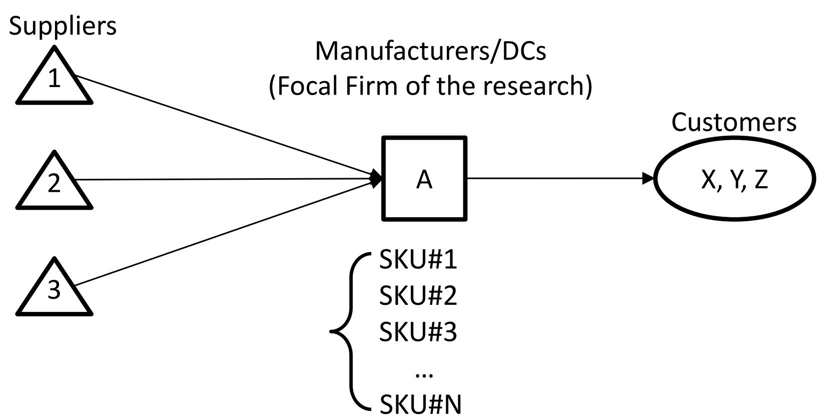
| | Teunter et al. (2010) | Neale and Willems (2009) | Millstein et al. (2014) | This Research |
|---|---|---|---|---|
| Integrated Approach | Yes | - | Yes | Yes |
| Nonstationary Demand | - | Yes | - | Yes |
| Linear Model | - | - | Yes | Yes |
| Profit Maximization | - | - | Yes | Yes |
| NPV of Profit | - | - | - | Yes |
| Multi-Period | - | - | - | Yes |
| Arbitrary Review Period | - | Yes | - | Yes |
| Comparison with ABC | Yes | - | Yes | Yes |
| Comparison with MCIC | - | - | - | Yes |

The performance of the proposed model is compared against the MCIC approach in a real-world case. Further, extensive computational experiments are conducted to evaluate the individual and interactive effects that various inventory parameters may have on the model performance relative to the traditional ABC method. It offers insights into under what circumstances companies may benefit significantly by employing the multi-period optimization model and in which situations the ABC approach may be sufficient.

## 3 Model Development

This research focuses on a three-stage supply chain setting as depicted in Figure 1.3, in which the focal firms are manufacturers or distribution centers (DCs) that source typically from multiple suppliers and sell to a number of downstream customers. While the objective of this research is to integrate SKU classification and service policies such that the profit of the focal firms is optimized, the companies at any stage of the supply chain may implement the proposed model independently to determine their optimal inventory policies based on the downstream demands, the upstream lead times, and their present review periods.

*Figure 1.3: Three-stage supply chain*



### 3.1 Assumptions

The addressed Multi-Period Inventory Classification with Non-Stationary Demand (MPIC-NSD) problem has the following assumptions.

Time Period: A finite planning horizon consisting of a set of discrete time periods is assumed, corresponding with the way inventory decisions are made in practice. The length of the discretized time periods can be customized and as granular as necessary. In most cases, a time period should be the basic unit of measure for time-based inventory parameters, such as replenishment lead time, inventory review period, and order delivery time. In other words, these time-based inventory parameters should be able to be described in non-negative integer multiples of the time periods. This approach has been commonly used in the inventory literature, including Hausman et al. (1998), Graves and Willems (2008), and Tarim and Kingsman (2006), to list a few. During a time period, the following events typically take place in sequence: inventory review, procurement order placement, replenishment receipt, and customer order fulfillment.

Demand Process: Demand is modeled as a nonstationary process with means and standard deviations both changing with time. The demand in each time period is assumed to be a realization of a random variable with a known probability distribution function, in line with the approach taken by Neale and Willems (2009) and Ettl et al. (2000). Demand characteristics generated by the forecast model are the input for the inventory optimization model. This research uses a normal distribution to model the probability distribution with forecast means and variances for each period. Forecast means are derived from time-series forecast models and variances are measured as mean squared forecast error (MSE).

The impact of inventory policies on demand is not straightforward and difficult to quantify. It is thereby assumed in this research that the future demand is only a function of the past demand pattern, and is not affected by the inventory decisions of the current planning horizon and the immediate last planning horizon. The effect of inventory decisions on short-term and long-term demand shall be studied as future research.

Replenishment Policy: Inventory is managed under a periodic-review, order-up-to inventory policy. Each item may have its own review interval, which can be one time period or multiple time periods. All else being equal, the longer the review intervals, the higher the on-hand inventory required. Inventory review is always conducted at the beginning of a time period, followed by procurement orders being generated according to the target base-stock levels.

Customer Order Fulfillment: If customer orders are fulfilled in the same period as the orders are received, it is considered as immediate delivery. If some SKUs in an order are in shortage, partially fulfilling the order is acceptable, and unfulfilled demands are treated as lost sales. The model assumes no backordering. Replenishment shipments received in a period are used to fulfill customer orders in the same period and in some cases, also the following time period(s), depending on the length of the review period.

Replenishment Lead Time: Deterministic replenishment lead times are adopted in the model, given that it is common to have planned lead times in a business-to-business environment (Hoen et al., 2011). Lead time is measured in units of the time period. Each SKU may have its own replenishment lead time.

Inventory Capital: The objective function is constrained by the maximum value of the inventory that can be carried during any given time period and is referred as the inventory capital or inventory investment throughout this dissertation. The level of inventory investment is the decision of the company's management. The inclusion of this constraint reflects the practical application of the model because the amount that a company can borrow is typically restricted and in most cases is in proportion to its own capital. The required inventory investment for profit maximization may or may not be practical for the company. Furthermore, the inventory capital restriction can also be used to approximate the storage limit in the manufacturers or distribution centers.

Costs Incurred: Costs incurred at the focal firms include inbound cost, inventory holding cost, outbound cost, and capital cost.

- Following the common industry practice, it is assumed here that suppliers are responsible for shipping and retain ownership of the goods until the goods arrive at the buyers' warehouses. The item costs in the model are thereby the total inbound costs including shipping and material costs.
- Inventory holding cost includes warehousing cost (e.g., rent, equipment, labor and utilities costs) and other inventory related cost (e.g., insurance, shrinkage, and obsolescence). A common practice to measure the holding cost is to take a percentage of the value of the goods in stock. Products with a shorter shelf-life tend to have a higher inventory holding cost percentage than those with a

longer shelf-life. In the model, inventory holding cost is calculated as $h_i \times c_i \times (IP\_End_{it} + IP\_A_{it})/2$, in which $h_i$ is the holding cost percentage, $c_i$ is the item cost, $IP\_Endi_t$ denotes the inventory position at the end of period $t$, and $IP\_A_{it}$ denotes the inventory position at the beginning of period $t$ after the receipt of the replenishment shipment. For a planning horizon of one year, the item cost and the holding cost percentage of an SKU in most cases remain constant; hence $c_i$ and $h_i$ in the proposed model are regarded independent of the time. The model can be easily extended to account for time-based costs in the case where they do vary over time.

- Cost of capital is separated from inventory holding cost in this research in order to examine the time value of money. The model assumes that an interest charge occurs at the end of a time period (which could be a week, a month, or any other defined length of time), consistent with the approach taken in the literature addressing discounted cash flow (e.g., Ammar, 2010).
- Outbound cost is mainly the shipping cost to the customers. Most companies cannot quantify the outbound logistics cost by SKUs, hence this is excluded from the current model. If the data are available, outbound cost can be easily incorporated through a reduced profit margin of individual SKUs.
- Besides the above four types of costs, the earlier inventory literature often considers ordering costs. Nowadays the orders are mostly placed over the Internet from the company's ERP system, which is part of the company's regular IT expenses. The incremental costs incurred with an increase of the

number of the orders in general are negligible. Therefore this research does not take into account ordering costs in the model.

- Another type of cost is the management cost, which is primarily related to physical inventory control such as cycle count policies based on the inventory stratification, but not necessarily directly related to the number of inventory classes. The primary goal of the inventory classification in this research is to set the most appropriate CSL and the safety stock level for each SKU, rather than the physical management of inventory. Companies can always combine the inventory classes suggested by the optimization model to fewer groups for the purpose of physical inventory handling. The management cost therefore is not considered in the model.

Pipeline Stock: As suppliers are responsible for shipping and retain ownership of the goods until the goods arrive at the buyers' warehouses, the pipeline stock is not considered in calculating inventory holding costs.

SKU Classification: Many companies, as well as inventory control software, use SKU classification to set service levels (Teunter et al., 2010). Each class is assigned a unique service level, and all the SKUs in that class target the same service level. In this research, the CSL instead of item fill rate is used to define SKU classes, in line with the approach taken by Teunter et al. (2010) and many commercial software products, because the safety stock and base-stock calculations are all based on CSL. To resolve the complexity of non-linearity, the CSL is discretized so that the model can be constructed as a mixed-integer linear programming (MILP) model, for which optimal solutions can be found and proved.

This research considers 109 discrete CSL levels. Integer-percentage service levels are used when CSL is less than 99%, but beyond 99%, companies typically use a more granular measure in tenths from 99.1% to 99.9%. Since the service level can never be truly 100%, a 99.99% level is used to approximate a perfect service level. In reality, changing the CSL by 5% from 60% to 65% may require a smaller increase in inventory investment compared to a 0.1% increase from 99.1% to 99.2%. As the service level grows higher, the unit of measure needs to be finer because a small improvement in CSL may lead to a substantial increase in inventory. The use of 109 levels for CSL ensures that the difference in safety stock derived from two adjacent classes is relatively small and the model can be solved within a reasonable time.

### 3.2 Problem Description

Let $N$ be the set of SKUs in the inventory, $T$ be the set of time periods in a planning horizon, and $J$ be the set of inventory classes to which an SKU can be assigned. The demand for SKU $i$ at time $t$ is assumed to follow a normal probability distribution with a forecast mean demand ($d_{it}$) and a standard deviation ($\sigma_{it}$). Both the mean and the standard deviation may fluctuate over time. Each SKU has its own review frequency ($f_i$) and lead time ($l_i$), from which $v_{it}$, a parameter indicating whether or not a replenishment order for SKU $i$ can arrive at time $t$, is derived. To elaborate, consider an SKU that is reviewed every other week ($f_i = 2$) and has a lead time of three weeks ($l_i = 3$). At week 3, its inventory is examined and a replenishment order is placed; three weeks later at week 6 the ordered supply arrives, setting $v_{i6}$ to 1. Then, this SKU is reviewed again at week 5 and another replenishment order is placed; at week 8, a new supply arrives, turning $v_{i8}$ to 1. Since there is no replenishment coming in during week 5 and 7, $v_{i5}$ and $v_{i7}$ are 0. Order quantity $O_{it}$ is a decision variable, denoting the replenishment quantity arriving at time $t$ when $v_{it}$ is 1; when $v_{it}$ is 0, the value of $O_{it}$ is also equal to 0. The model uses $O_{it}$ only as the product of $v_{it}$ and $O_{it}$, which further ensures that no supply arrives at time $t$ when $v_{it}$ is 0. It should be noted that in the optimization model, $v_{it}$ is a parameter, recording the arrival of orders only, though it is derived by using both replenishment lead time and review frequency.

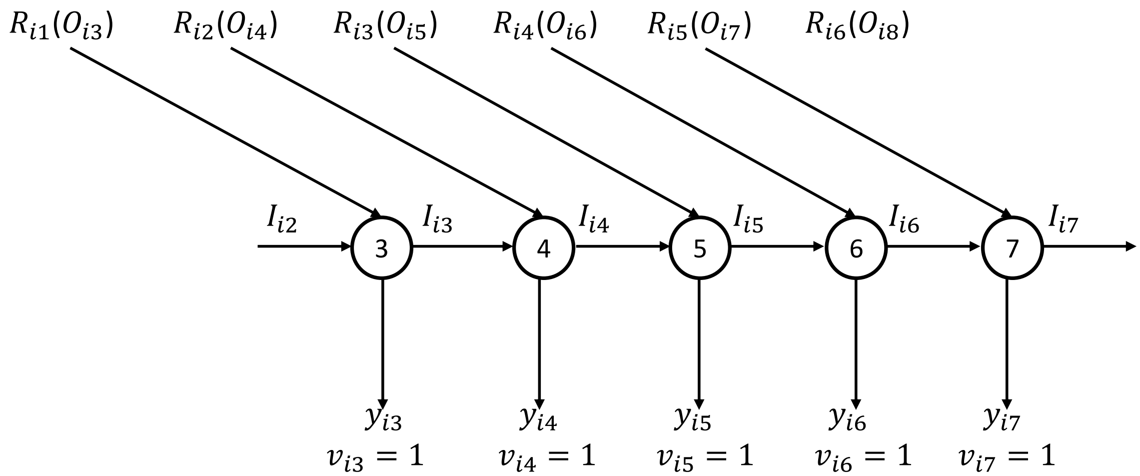
Each inventory class $j$ has a corresponding service level ($\alpha_j$), a derived z-value ($z_j$) to determine safety stock, and a derived value from the standard loss function ($e_j$) to decide the expected lost sales. SKU $i$ at time $t$ can only be assigned to at most one inventory class and the total inventory value in any given time period shall not be higher than the inventory capital $\omega$. The goal is to balance inventory holding cost and lost sales under a limited inventory capital by simultaneously classifying SKUs and setting the CSL for each inventory class such that the NPV of profit is maximized.

### 3.3 Model Formulation

The formulation of the multi-period model is built upon the NETFORM framework (Glover et al., 1992), which conveniently models nonstationary demand, arbitrary review periods, and different lead times (see Figure 1.4). Each node represents a time period which can be a month, week, or any other defined length of time. Inclined arcs directed into the nodes represent arriving replenishment orders, while vertical arcs directed out of the nodes represent the mean demand that is expected to be filled. Horizontal arcs directed into the nodes represent on-hand inventory that is carried from the last period to this period, whereas horizontal arcs directed out of the nodes represent on-hand inventory that is carried from the current period to the next period.

Figure 1.4 provides an example in which an SKU is reviewed on a monthly basis and has a 2-month lead time. At the beginning of Month 5, the inventory is comprised of two components: the inventory carried over from last month ($I_{i4}$; i.e., the safety stock of Month 4), and the replenishment order ($O_{i5}$) placed in Month 3 ($R_{i3}$ is for descriptive purpose only, indicating when an order is placed). This inventory is expected to satisfy demand $y_5$ and have $I_5$ of inventory (safety stock) left for Month 6. The quantity of the replenishment order placed in Month 3, the expected fulfilled demand $y_5$, and the safety stock $I_5$ are all determined by the class to which the SKU is assigned and the corresponding service level. Because of the conversion required from service level to z-value in the inventory calculation, the vast majority of inventory models, if not all, are nonlinear. The discretization of service level allows the model in this research to be a linear model.

*Figure 1.4: Inventory Flowchart of an SKU (Review Frequency = 1; Lead Time = 2)*




The model formulation is based on the following notation.

<u>*Sets*</u>

$N$: set of SKUs in the inventory

$T$: set of time periods in a planning horizon

$J$: set of inventory classes to which an SKU can be assigned

<u>*SKU-related Parameters*</u>

$d_{it}$: expected demand of SKU *i* in time period *t* (forecast demand), $\forall i \in N, t \in T$.

$\sigma_{it}$: standard deviation of the forecast demand in time period *t*, $\forall i \in N, t \in T$.

$v_{it}$: =1 if there is an arriving replenishment order at time $t$ (placed at time $t - l_i$); 0 otherwise, $\forall i \in N, t \in T$.

$\pi_i$: unit gross profit, $\forall i \in N$.

$c_i$: cost of goods per unit, $\forall i \in N$.

$h_i$: inventory holding cost (% of the cost of SKU $i$) per time unit, excluding interest rate, $\forall i \in N$.

*Other Parameters*

$\alpha_j$: CSL associated with class $j$, $\forall j \in J$.

$z_j$: z-value associated with CSL $\alpha_j$, $\forall j \in J$.

$e_j$: value in standard loss function (corresponding to CSL $\alpha_j$), $\forall j \in J$.

$r$: capital cost per time unit

$\omega$: inventory capital. On-hand inventory at the end of a time period $t$ must be equal to or less than $\omega$.

*Decision Variables*

$x_{ijt} = 1$ if SKU $i$ is assigned to class $j$ in time period $t$, $\forall i \in N, j \in J, t \in T$.

$O_{it} \geq 0$: replenishment order that arrives at the beginning of time period $t$, $\forall i \in N, t \in T$.

$I_{it} \geq 0$: on-hand inventory (safety stock) at the end of time period $t$ for SKU $i$, $\forall i \in N, t \in T$.

$y_{it}$: demand that is expected to be satisfied in time period $t$, $\forall i \in N, t \in T$.

$AveI_{it}$: average inventory level of SKU $i$ at time $t$, $\forall i \in N, t \in T$.

*Objective Function*

$$\max \sum_{t \in T} \sum_{i \in N} (\underbrace{\pi_i \times y_{it}}_{①} - \underbrace{h_i \times c_i \times AveI_{it}}_{②} - \underbrace{I_{i(t-1)} \times c_i \times r}_{③}) / (1+r)^t \qquad (1)$$

*Constraints*

$$\sum_{j \in J} x_{ijt} \leq 1 \qquad \forall i \in N, t \in T \qquad (2)$$

$$y_{it} \leq d_{it} - \sigma_{it} \times \sum_{j \in J} x_{ijt} \times e_j \qquad \forall i \in N, t \in T \qquad (3)$$

$$y_{it} \le d_{it} - \sigma_{it} \times e_1 \times \left(1 - \sum_{j\in J} x_{ijt}\right) - \varepsilon \quad \forall i \in N, t \in T \tag{4}$$

$$I_{i(t-1)} + O_{it} \times v_{it} = I_{it} + y_{it} \quad \forall i \in N, t \in T \tag{5}$$

$$I_{i(t-1)} + O_{it} \times v_{it} \ge \sum_{j\in J} x_{ijt} \, (d_{it} + \sigma_{it} z_j) \quad \forall i \in N, t \in T \tag{6}$$

$$I_{it} \ge 0 \quad \forall i \in N, t \in T \tag{7}$$

$$O_{it} \ge 0 \quad \forall i \in N, t \in T \tag{8}$$

$$\sum_{i\in N} c_i \times (I_{i(t-1)} + O_{it} \times v_{it}) \le \omega \quad \forall t \in T \tag{9}$$

$$AveI_{it} = (I_{i(t-1)} + O_{it} \times v_{it} + I_{it})/2 \quad \forall i \in N, t \in T \tag{10}$$

$$x_{ijt} \in \{0, 1\} \quad \forall i \in N, j \in J, t \in T \tag{11}$$

$$y_{it} \ge 0 \quad \forall i \in N, t \in T \tag{12}$$

$$AveI_{it} \ge 0 \quad \forall i \in N, t \in T \tag{13}$$

The objective function (1) maximizes the NPV of expected profit. The first part of the equation describes the expected gross profit calculated upon the expected sales volume ($y_{it}$); the second part is the inventory holding cost, computed by taking the average of the beginning and end inventory (equation(10)); and the third part is the interest charge on on-hand inventory at the end of a time unit.

Constraint (2) restricts that an SKU cannot be assigned with more than one CSL at any given time period. If an SKU is not assigned with any CSL ($x_{ijt} = 0$), it means that the SKU either has zero CSL (when there is no inventory) or has CSL less than 1%.

Constraints (3) and (4) determine the demand that is expected to be satisfied. Constraint (3) ensures that when an SKU is assigned a CSL, the mean satisfied demand

($y_{it}$) is the subtraction of the expected lost sales (determined by the CSL) from the forecast mean demand. If an SKU is not assigned any CSL, $y_{it}$ is restricted by constraint (4), which requires the mean satisfied demand must be less than the one when CSL is 1%. $\varepsilon$ denotes a small positive infinitesimal quantity.

Constraint (5) is a flow-balancing constraint that ensures the amount of each SKU in each time period coming into the node equals the amount flowing out of the node. Constraint (6) ensures on-hand inventory covers both regular stock and safety stock. For example, if the forecast mean demand is 400 and the standard deviation is 40, then to make sure the expected satisfied demand is close to 400, the CSL needs to be 99.99%, which requires on-hand inventory to be $d_{it} + \sigma_{it} z_j = 400 + 40 \times 3.7190 \approx 549$, where 3.7190 is the z-value corresponding to a 99.99% CSL.

Constraints (7) and (8) require on-hand inventory and order quantity to be non-negative. Constraint (9) ensures that the total value of the inventory at any given time period does not exceed the available inventory capital. When this constraint is relaxed, the model identifies the optimal inventory capital needed in order to maximize profit. An inventory investment less than the optimal amount will result in a reduced profit. Constraint (10) determines the average on-hand inventory of an SKU at a given time period.

The model is an MILP model, which is NP-hard. There is no known polynomial algorithm to solve it to optimality. The proof of NP-hardness is established by transforming the formulation into an uncapacitated facility location problem (UFLP, cf. Drezner and Hamacher 2004).

## 3.4 Summary of Model Development

The proposed MILP formulation of the MPIC-NSD model utilizes the NETFORM structure to integrate inventory classification and service policies, eliminating the tedious process associated with using the ABC approach to classify inventory and set service levels, and offering management a decision-support tool to manage inventory holistically from classification to profit optimization. More importantly, the model incorporates a number of practical issues faced by management: nonstationary demand, limited inventory capital, arbitrary review periods, SKU-varying holding costs, and time value of money. The model is a generalized model designed for profit-driven corporations dealing with finished goods inventory.

# 4 Real World Application: A Case Study

The above MPIC-NSD model is evaluated on a real world database from a St. Louis Missouri based cheese processor and distributor selling imported and domestic cheeses and specialty foods to retail grocers and wholesale brokers in the Midwest. The company has been in operation for over 70 years and experienced a number of major expansions. It presently purchases from over 160 suppliers and serves about 150 customers, with an annual sales revenue of nearly $100 million in 2015. To stay ahead of the competition, it constantly introduces new products and revises its product mix. Out of its 608 active SKUs, nearly half were introduced within the past four years.

## 4.1 Case Description

The company presently uses a weighted-multi-criteria inventory classification scheme to manage its inventory, which takes into account the sales volume, coefficient of variation in demand, number of orders, shelf life, and gross profit. The goal of the classification is to improve revenue and reduce waste (caused by obsolescence) so that the total profit can be increased. Obviously this classification scheme is much more comprehensive than the traditional ABC approach. However, as a pure classification technique, how this MCIC method may improve inventory performance measures such as cost and profit is not straightforward and difficult to quantify during the planning. Also, it presents the same challenge as the ABC approach in that the classification process doesn't suggest what service level should be assigned to a class. In most cases, companies start with an arbitrary service level for each class based on experience, and then make a number of rounds of adjustments based on the available inventory capital. The St. Louis-based cheese distributor followed this approach. They worked with a supply chain consultant to first weigh each criterion according to its importance that they perceived, and then calculate a weighted-multi-criteria score for each SKU. The SKUs were split into two groups based on their scores, and each group was further divided into four categories. The categories in the group with higher scores are named as Class A+, A, B, and C, and the categories in the other group are named as Class A+(2), A(2), B(2), and C(2). Through trial and error, they assigned the following CSLs to each class as described in Table 1.2.

*Table 1.2: Multi-Criteria Inventory Classification and Service Level of a St. Louis-based Cheese Distributor*

| **Class** | A+ | A | B | C | A+(2) | A(2) | B(2) | C(2) |
|---|---|---|---|---|---|---|---|---|
| **CSL** | 98% | 95% | 90% | 80% | 65% | 60% | 55% | 50% |

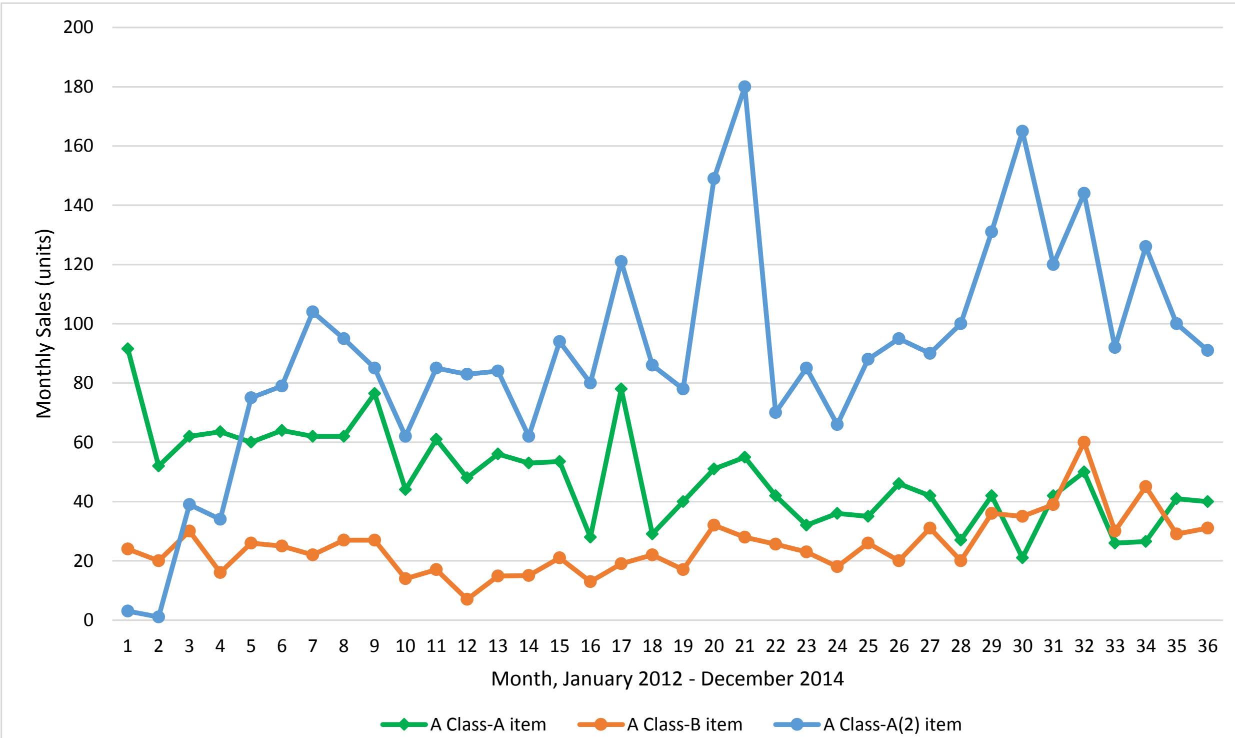
## 4.2 Application of MPIC-NSD Model

To evaluate the performance of the MPIC-NSD model against their current MCIC approach, the following SKU-level information is required: the forecast monthly mean demand and standard deviation, the actual monthly demand, unit cost, profit margin, lead time, review/order frequency, and inventory holding cost (which incorporates shelf-life). The database has 608 SKUs, among which 284 are made-to-stock items with full four-year (January 2012 to December 2015) monthly sales history, while the rest are either made-to-order or introduced to the market only in the last four years. The evaluation therefore is conducted on the 284 SKUs with four years of sales data.

The data from January 2012 to December 2014 are used to forecast the monthly mean demands and standard deviations for the year of 2015. Monthly sales history of three SKUs (randomly chosen) from Class A, B and A(2) respectively is provided in Figure 1.5 as an example. The company employs the triple exponential smoothing method to forecast monthly sales. The root mean squared forecast errors computed by comparing the actual and forecast monthly sales in 2013 and 2014 are used to represent the monthly demand uncertainty for 2015. This demand information is the input for both the optimization model and the MCIC method to calculate target base stock levels in each month. The actual monthly demands in 2015 are found mostly within 3 standard deviations of the forecast means.

The initial inventory is set at the target base stock levels as the company is unable to trace back the amount of on-hand inventory at the end of 2014. The inventory capital for the 284 SKUs is set at $9.2 million based on the calculated target base stock levels using the MCIC method. The company is a private firm and its cost of capital is not revealed. The annual interest rate in this case study uses 9.07%, the weighted average cost of capital of food wholesalers industry in 2015 according to Damodaran (2016).

*Figure 1.5: Historical monthly sales of three randomly-chosen SKUs*




The MPIC-NSD optimization model was solved with the branch-and-cut method using GAMS/CPLEX 12.4 on a laptop that has an 8GB RAM and runs on Intel Core i7-4500 CPU with a maximum frequency of 2.4GHz. The relative optimality tolerance parameter was set as .00001. It took the solver 228.20 CPU seconds to find a near-optimal solution (relative gap to the best possible solution is .000001).

Table 1.3 below presents the financial results of using the MPIC-NSD model and the MCIC approach respectively to determine service levels, safety stock and base stock levels for the 284 SKUs. Although the company's current approach has already taken into account profit, demand variation and shelf life, which the traditional ABC method does not consider, the proposed model is able to improve the NPV of profit by 2.8% through increased sales and reduced inventory levels. With the MPIC-NSD model, the monthly inventory capital required ranges from $4.58 million to $5.90 million, on average 13% less than the MCIC method each month. The optimal solution doesn't reach the company's current inventory capital limit.

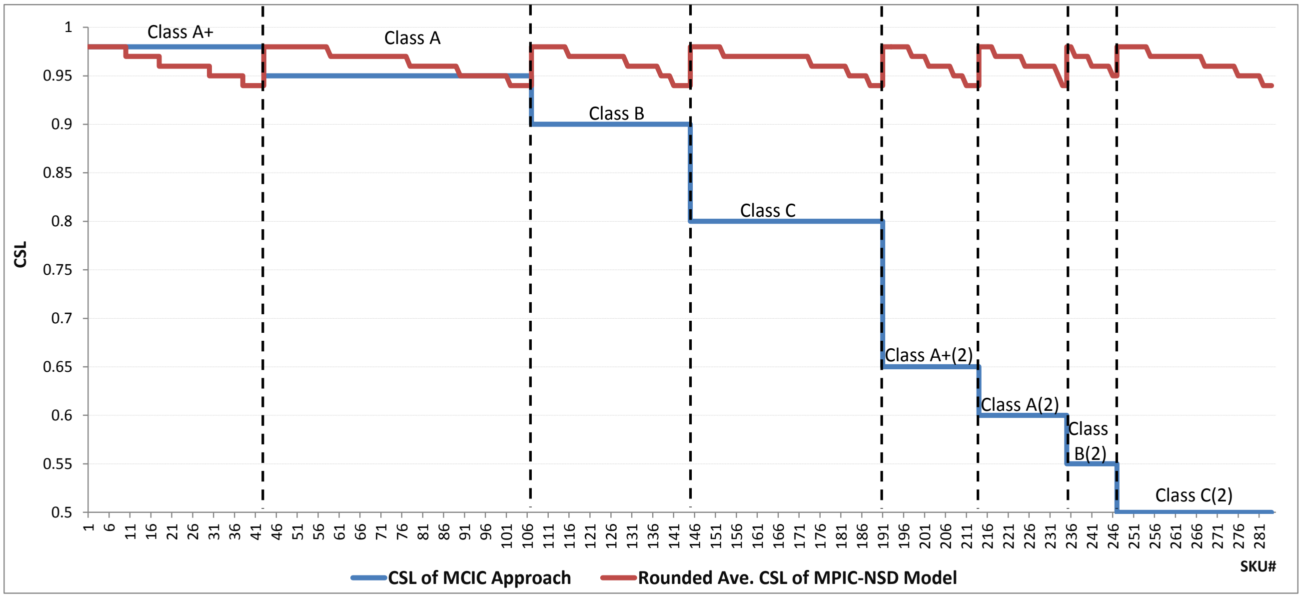
*Table 1.3: Financial Performance Comparison: MPIC-NSD Model vs. MCIC Method*

| | NPV | | | | Inventory capital Required | |
|---|---|---|---|---|---|---|
| | Total Revenue | Inventory Cost | Interest Charge | Net Profit | Maximum | Minimum |
| **MPIC-NSD** | $59,935,103 | $1,061,613 | $196,139 | $25,648,428 | $5,895,430 | $4,527,425 |
| **MCIC** | $59,014,340 | $1,293,051 | $268,141 | $24,961,481 | $6,443,148 | $4,677,659 |
| **Improvement ($)** | $920,763 | -$231,437 | -$72,002 | $686,947 | -$547,719 | -$150,234 |
| **Improvement (%)** | 1.56% | -17.90% | -26.85% | **2.75%** | -8.50% | -3.21% |

The expected-demand-weighted average monthly CSL of an SKU suggested by the optimization model ranges from 94% to 98%, significantly differing from the service levels that the company is currently using, as described in Figure 1.6. For example, the current Class A+ has 43 items with a unified service level of 98%, while the model suggests that the company is better off to differentiate those items with 5 different service levels, from 94% to 98%. Similarly, the Class C(2) is believed to deserve a service level of only 50%, but the optimization model shows that there should be 5 different classes and none of the SKUs have a weighted average CSL less than 94% in order to maximize the NPV of net profit. Out of the 284 SKUs, only 22 SKUs have a current CSL the same as the weighted average CSL generated by the optimization model and many of the optimal CSLs differ by a very large amount (over 40%).

The two most important classes perceived by the company, Class A+ and A, account for about 63% of the total sales volume. Although the CSLs suggested by the model for Class A+ and A appear to be less substantial than the other classes (compared to their current CSLs), these two classes contribute about 76% of the improvement in NPV of profit.

*Figure 1.6: SKU CSL/Classification comparison: MCIC approach vs. MPIC-NSD model*




A major goal of the MPIC-NSD model is to obtain dynamic optimal inventory solutions so as to cope with the change of the expected demand and forecast uncertainty across the planning horizon. Over the 12-month planning cycle, 22 different CSLs are proposed by the model, ranging from 88% to 99.99%. 181 SKUs are expected to have a consistent CSL throughout the year, whereas the remaining 103 SKUs are assigned with 2 to 7 different CSLs over time, as described in Table 1.4.

*Table 1.4: Expected number of different CSLs by SKUs over the planning horizon*

| Number of SKUs | 181 | 22 | 53 | 14 | 11 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| Expected number of different CSLs over planning horizon | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Figure 1.7 provides an example of the different CSLs assigned to a specific SKU over the year. This SKU is currently in Class C with a CSL of 80%. Its demand fluctuates from month to month and some months are more difficult to forecast than others. The CSLs assigned to it by the model range from 88% to 99.99%. The large drop in CSL in February is likely caused by its high demand uncertainty (coefficient of variation is close to 100%) as shown in Table 1.5. Demands in May and June are expected to be relatively stable, and as their mean values are much less than the safety stock in April, high CSLs are anticipated. This is consistent with what have been observed in businesses: when a company transits from a high-demand season to a low-demand season, inventory is often high and can well satisfy the demand in the first one or two months of the low season. For this specific SKU, adopting the dynamic inventory solutions generated by the optimization model provides 50% higher profit (in terms of NPV) than using its current approach.

*Figure 1.7: Assigned CSL over the planning horizon to a specific SKU*

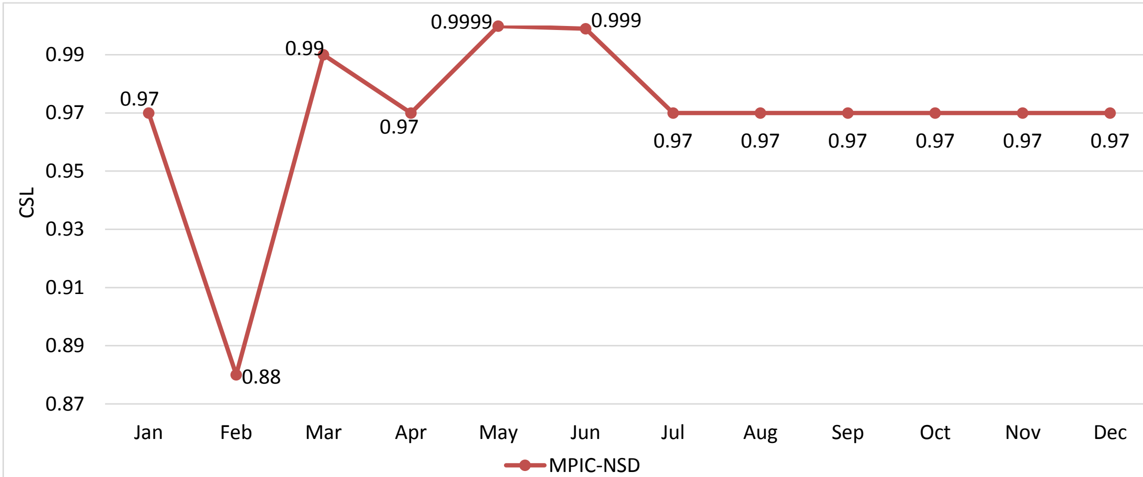



*Table 1.5: Expected demand pattern of a specific SKU*

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Expected Demand | 67 | 46 | 21 | 15 | 6 | 5 | 5 | 17 | 15 | 16 | 17 | 33 |
| Standard Deviation | 39 | 45 | 15 | 11 | 3 | 3 | 4 | 10 | 8 | 9 | 9 | 27 |
| Cofficient of Variation | 57.45% | 99.58% | 72.28% | 71.97% | 60.92% | 63.98% | 82.32% | 58.46% | 53.73% | 52.23% | 55.55% | 84.54% |

Table 1.6 compares the item fill rate of the two approaches. With the MPIC-NSD model, the average monthly fill rate across all the SKUs is 95.5%, more than 1% higher than the current MCIC approach. As the minimum service level suggested by the optimization model is 94%, significantly higher than the Class C(2) service level (50%), it is not surprising that the minimum average monthly fill rate by SKUs and the number of SKUs with 100% fill rate are all higher with the MPIC-NSD model. It's worth noting though that these are achieved with a smaller inventory capital requirement.

*Table 1.6: Item Fill Rate Comparison: MPIC-NSD Model vs. MCIC Method*

| | **Ave. Fill Rate** | **Min. Fill Rate** | **# SKUs with 100% Fill Rate** |
|---|---|---|---|
| **MPIC-NSD** | 95.5% | 24.2% | 112 |
| **MCIC** | 94.3% | 19.4% | 109 |

## 4.3 Impact of the Number of Classes

Out of the 22 different CSLs suggested by the model, 5 levels are most commonly used, which are 94%, 95%, 96%, 97% and 98%. Every SKU has been given at least one of these CSLs in some or all the months during the planning horizon. From Figure 1.8, the CSLs of 96% and 97% are most popular, used in 1,794 (or 53%) assignment decisions, followed by the CSLs of 95% and 98% which are used in 1,111 (or 33%) assignment decisions; and the CSL of 94% is used in 350 (or 10%) assignment decisions. The remaining 17 CSLs only account for about 4% of the assignment decisions, and are allocated to 103 SKUs (out of 284 SKUs) for only one or two months of the planning horizon. This raises the question about whether a small number of CSLs might be sufficient as far as profit is concerned.

*Figure 1.8: Relative frequency distribution of CSLs used in the model*

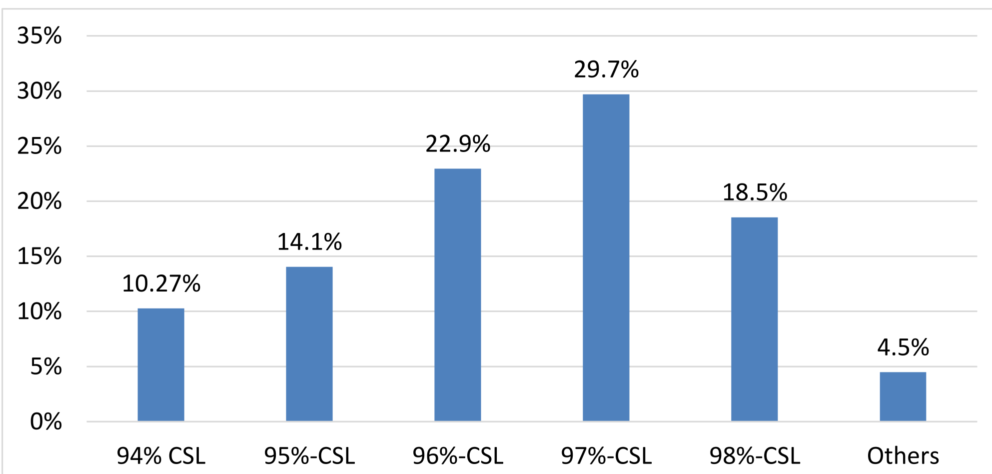



Table 1.7 summarizes the results of limiting the total number of CSLs that can be chosen while relaxing the limited inventory capital constraint. When the number of classes is restricted to 5, the impact on the NPV of profit is almost negligible with only

0.13% or $32,306 less than the optimal value. When the number of classes is limited to 3, the resulting NPV of profit is -1.37% or $351,144 less than the optimal result. Although there are only a small number of CSLs to allocate, over one third of the SKUs are assigned with multiple CSLs over the planning horizon. For this particular dataset, its NPV of profit appears to be insensitive to the number of CSLs, while what is most important is how to choose the appropriate CSLs and then how to allocate those service levels to SKUs.

*Table 1.7: Impact of the number of CSLs*

| Total # of CSLs | **Chosen CSLs** | **Comparison to optimal solution (%)** | | | |
|---|---|---|---|---|---|
| | | **NPV of Profit** | **Total Revenue** | **Inventory Cost** | **Interest Charge** |
| **3 CSLs** | 83%, 95%, 97% | -1.37% | -1.72% | -5.43% | -6.75% |
| **4 CSLs** | 96%, 97%, 98%, 99% | -0.70% | 0.03% | 14.44% | 16.94% |
| **5 CSLs** | 95%, 96%, 97%, 98%, 99.3% | -0.13% | 0.04% | 3.49% | 4.10% |

### 4.4 Impact of Inventory Capital

The expected monthly inventory capital required to maximize the NPV of profit ranges from $4.27 million to $5.21 million with an average of $4.78 million. To gain an insight into how the choice of CSLs and assignment decisions may be affected by the inventory investment, the model is tested with two levels of capital constraint, $4 million and $3 million, respectively. Table 1.8 summarizes the results. When the monthly inventory capital is limited to $4 million, the model chooses 56 different CSLs (vs. 22 CSLs when there is no capital limit), ranging from 48% to 99.9% (vs. 88% to 99%), and the monthly average CSL across all SKUs is lowered to 81% (vs. 96%). None of the SKUs have a consistent CSL with vast majority (about 95%) being assigned with 9 to 11 different CSLs over time. Such dynamic management of inventory allows the expected NPV of profit to only reduce by 1.73% or $0.55 million, compared to the situation where sufficient inventory capital is available.

Further reducing the inventory investment to $3 million caused 79 different CSLs to be used, including 0% which suggests that no inventory needs to be held. The SKU's average CSL is only about 45% and 277 SKUs (out of 284) are assigned with 8 to 10 different CSLs over time. The expected NPV of profit is 12.3% or $3.92 million less than the optimal result.

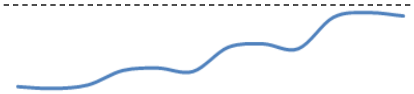
*Table 1.8: Impact of Inventory Investment*

| Inventory Investment | **Chosen CSLs** | **SKU's Average CSL over Planning Horizon** | **Δ to Expected Optimal NPV of Profit** |
|---|---|---|---|
| **$4 million** | 56 in total, ranging from 48% to 99.9% | 68% to 89% | -1.73% |
| **$3 million** | 79 in total, ranging from 0% to 83% | 6% to 68% | -12.30% |

## 5 Computational Experiments

Further computational experiments are performed to examine the impact of demand pattern, demand variability, replenishment lead time, inventory holding cost, and interest rate on the profit improvement of the MPIC-NSD model over the ABC approach. The goal is to provide insight into under what circumstances the ABC approach may be sufficient and under what situation a company should employ the MPIC-NSD model for integrated inventory management. Table 1.9 illustrates the $4 \times 5 \times 3 \times 5 \times 3$ experimental design used to generate data and perform sensitivity analysis.

*Table 1.9: Experimental design*

| Parameters | # of Levels | Description |
|---|---|---|
| Demand pattern<br>(Detailed in Table 1.12) | 4 | 1 <br>2 <br>3 <br>4  |
| Demand uncertainty (measured by coefficient of variation) | 5 | 5%, 10%, 15%, 20%, 25% |
| Replenishment lead time | 3 | 1 month, 2 months, 3 months |
| Annual inventory holding cost | 5 | 10%, 15%, 20%, 25%, 30% |
| Annual interest rate | 3 | 5%, 10%, 15% |

The dataset has 900 SKUs, determined by the unique combinations of demand pattern, demand variability, replenishment lead time, inventory holding cost, and interest rate. Item costs are randomly generated following a uniform distribution $U[10, 135]$, and gross margins follow $U[0.5, 1.0]$, both of which remain constant throughout the experiments. The realized monthly demand of an SKU is assumed to follow a normal distribution $N(d_{it}, \sigma_{it})$, where $d_{it}$ is derived from the mean annual demand and demand pattern, and $\sigma_{it}$ is generated according to coefficient of variation.

In order to compare with the ABC approach which is developed upon the Pareto principle (i.e., 80-20 rule), about 20% (187) of the SKUs are assigned to Class A with each having an average expected monthly demand ranging from 10,000 to 32,500 units.

Class A as a whole accounts for about 80% of the total sales volume. Class B contains about another 20% (177) of the SKUs with an average expected monthly demand ranging from 2,700 to 5,700, and accounts for about 15% of the total sales units. The remaining 60% of SKUs are Class C items with an average expected monthly demand ranging from 250 to 680 units, contributing 5% of the sales volume. The decisions of what CSL is given to each class is made through the Solver tool in Excel to optimize the expected profit, although in practice the CSL is often arbitrarily assigned to each class. The sales-volume-weighted average CSL of the ABC approach is equivalent to the one generated by the MPIC-NSD model to ensure a fair comparison. Inventory capital in the MPIC-NSD model is set at the highest level required by the ABC approach, which turns out to be a non-binding constraint because the inventory investment required by the optimal solution is much less than that of the ABC approach. Table 1.10 summarizes the "optimized" ABC classification used in the experiments.

*Table 1.10: ABC classification*

| Class | Average monthly mean demand | % of total SKUs | % of total revenue | Assigned CSL |
|---|---|---|---|---|
| A | $U$ (10000, 32500) | 20.8% | 79.7% | 97% |
| B | $U$ (2700, 5700) | 19.7% | 15.1% | 94% |
| C | $U$ (250, 680) | 59.6% | 5.1% | 86% |
| Weighted average | | | | 96% |

To accurately evaluate the profit improvement of the optimization model over the ABC approach, a simulation model is constructed using Excel. Simulation is a powerful tool for analyzing inventory systems and comparing the alternate policies of an inventory system (Mahamani et al., 2008), and many researchers have utilized Excel spreadsheets to simulate supply chain networks (Chwif et al., 2002), the bullwhip effect (Boute and

Lambrecht, 2005), a warehouse system (Sezen and Kitapci, 2007), and inventory management (Jung et al., 2007). In the experiments, a set of actual monthly demand data are randomly generated for 1000 replications through Monte Carlo simulation using the forecast means ($d_{it}$) and the standard deviations ($\sigma_{it}$). The decision variables of the MPIC-NSD model consist of an SKU's CSL, replenishment order quantity, and on-hand inventory at month $t$, upon which the target inventory level ($T_{it}$) is determined. With the ABC approach, the order-up-to inventory level at month $t$ ($T_{it}$) is a function of replenishment lead time ($l_i$), inventory class (upon which $z_i$ is defined), forecast monthly mean demand ($d_{it}$) and standard deviation ($\sigma_{it}$).

$$T_{it} = \sum_t^{t+l_i} d_{it} + z_i \times \sqrt{\sum_t^{t+l_i} \sigma_{it}^2} \tag{14}$$

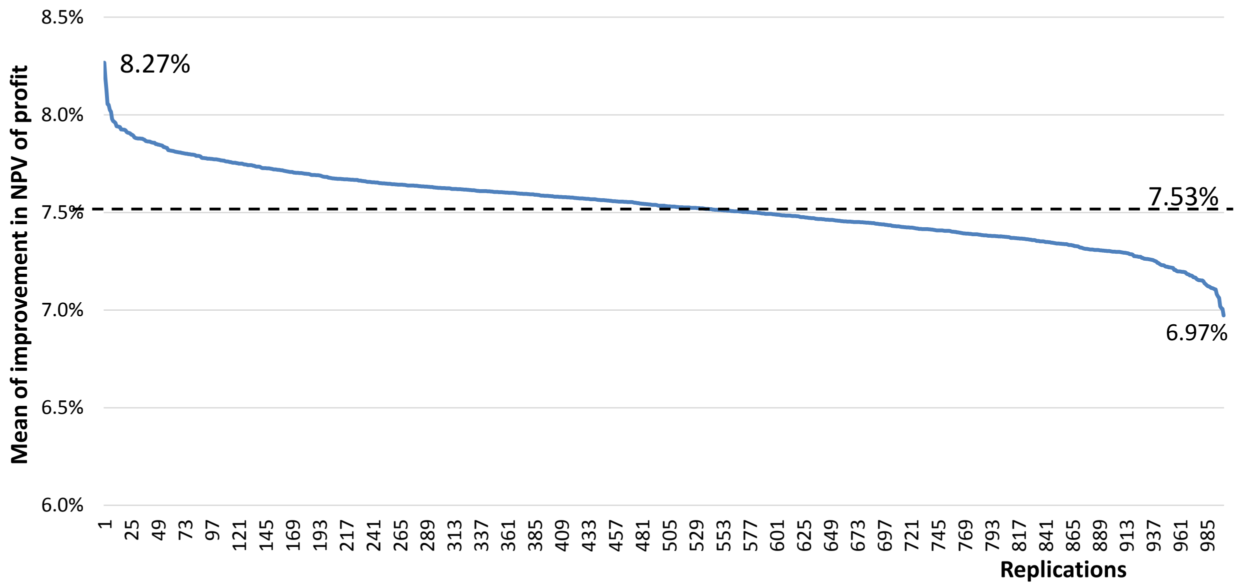
The ABC classification is assumed to be combined with dynamic inventory policies (changing the order-up-to level according to the forecast mean and standard deviation) in the experiments, which provides a better inventory performance than the ABC classification alone would otherwise be able to achieve. Consequently, the resulting gap in profit of the two approaches is a conservative assessment.

The replenishment order is placed according to the target inventory level and on-hand inventory. Since no backorder is allowed, unfulfilled demand is treated as lost sales. Figure 1.9 presents the percentage increase in the NPV of profit of the MPIC-NSD model over the ABC approach after reordering the results. Of the 1000 simulation runs, the minimum improvement is 6.97% and the maximum is 8.27%, while the average is 7.53%. Table 1.11 summarizes the simulation results.

*Table 1.11: Simulation results: MPIC-NSD model vs. ABC approach*

| MPIC-NSD | NPV of Profit | | | | |
|---|---|---|---|---|---|
| over ABC | Max | Min | Mean | Std. Dev. | C.V. |
| Improvement ($) | 216,319,550 | 187,123,886 | 200,005,077 | 4,209,684 | 2.10% |
| Improvement (%) | 8.27% | 6.97% | 7.53% | 0.19% | 2.47% |

*Figure 1.9: Improvement in NPV of profit of MPIC-NSD model over ABC approach*




## 5.1 Effects of Individual Parameters on Model Performance

Effects of demand pattern, interest rate, replenishment lead time, inventory holding cost, and demand variability on model performance are examined through simulation. Except for demand variability, all the other parameters appear to have a linear relationship with the magnitude of profit improvement, in comparing the MPIC-NSD model with the ABC approach.

<u>*Demand pattern*</u>

Four demand patterns that are commonly observed in the real world are examined in the experiment:

(1) Demand is high in the first quarter, and then slows down gradually in the next three quarters;

(2) Demand is low initially, then starts to pick up gradually in the second quarter, and finally reaches the peak in the last quarter;

(3) Demand is relatively high in the second and fourth quarters;

(4) Demand is relatively stable across the year.

*Table 1.12: Four demand patterns*

| | Expected monthly demand as a % of expected annual sales volume | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| Demand Pattern 1 | 16.7% | 16.7% | 16.7% | 10.0% | 10.0% | 10.0% | 5.0% | 5.0% | 5.0% | 1.7% | 1.7% | 1.7% |
| Demand Pattern 2 | 1.7% | 1.7% | 1.7% | 5.0% | 5.0% | 5.0% | 10.0% | 10.0% | 10.0% | 16.7% | 16.7% | 16.7% |
| Demand Pattern 3 | 5.0% | 5.0% | 5.0% | 14.0% | 14.0% | 14.0% | 1.7% | 1.7% | 1.7% | 12.7% | 12.7% | 12.7% |
| Demand Pattern 4 | 7.7% | 6.7% | 7.6% | 8.1% | 8.3% | 8.0% | 9.4% | 8.9% | 8.1% | 9.5% | 8.3% | 9.2% |

Figure 1.10 displays the distribution of the results (i.e., the improvement of the NPV of profit) by demand patterns, where each dot represents the average improvement of an SKU from 1,000 replications. The MPIC-NSD model sees the highest average performance improvement of 16.97% over the ABC method for demand pattern (1), but it is over a wide range from -5% to over 70%. Demand pattern (3) shows an average improvement of 9.5%, with a relatively smaller range from -4% to 55%. Demand patterns (1) and (3) both have a relatively high demand in the first half of the year. Demand (2) and (4) observe smaller average improvements of less than 4%, but their variations are also much less. The results of one-way ANOVA are presented in Table 1.13. With a $p$-value less than 0.001, the differences in the average profit improvement of different demand patterns are confirmed statistically significant.

From Figure 1.10 and Table 1.13, when the expected demand is high in the early of the year (demand pattern (1) and (3)), a closely managed inventory may lead to a large improvement in NPV of profit. This might be attributed to the interest rate: higher mean demands and higher standard deviations in the first half of the year lead to higher safety stock which in turn results in higher interest charge in comparison with the higher safety stock in the later of the year. The data from experiments suggest:

*Finding 1: The performance of the MPIC-NSD model relative to the ABC approach varies by demand pattern. A higher performance improvement is expected when demand tends to be large early in the year (planning horizon).*

*Figure 1.10: Improvement in NPV of expected profit by demand patterns*

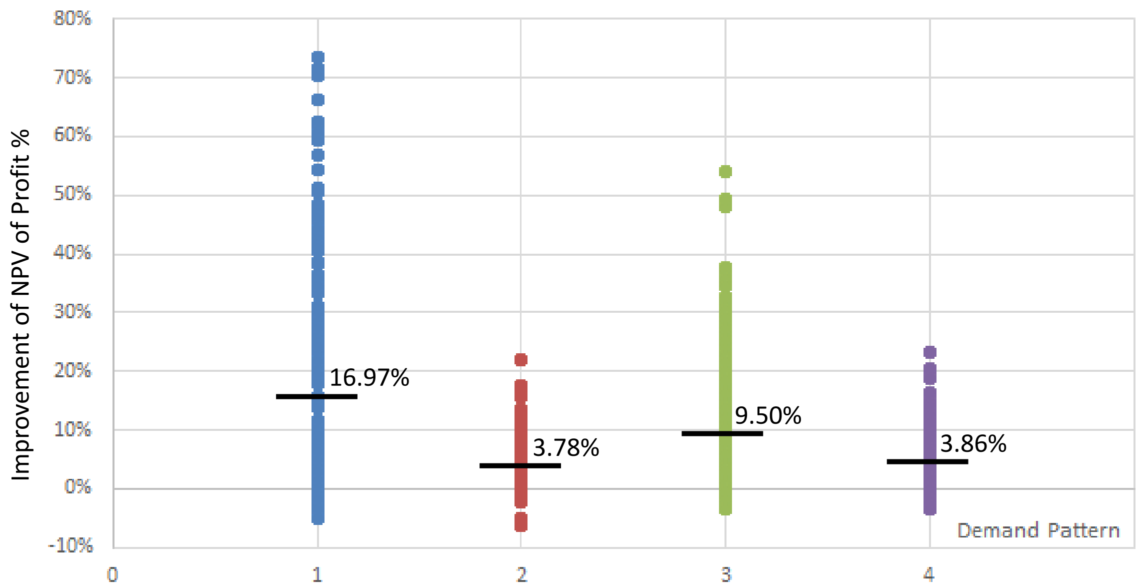


*Table 1.13: ANOVA test of performance equality of demand patterns*

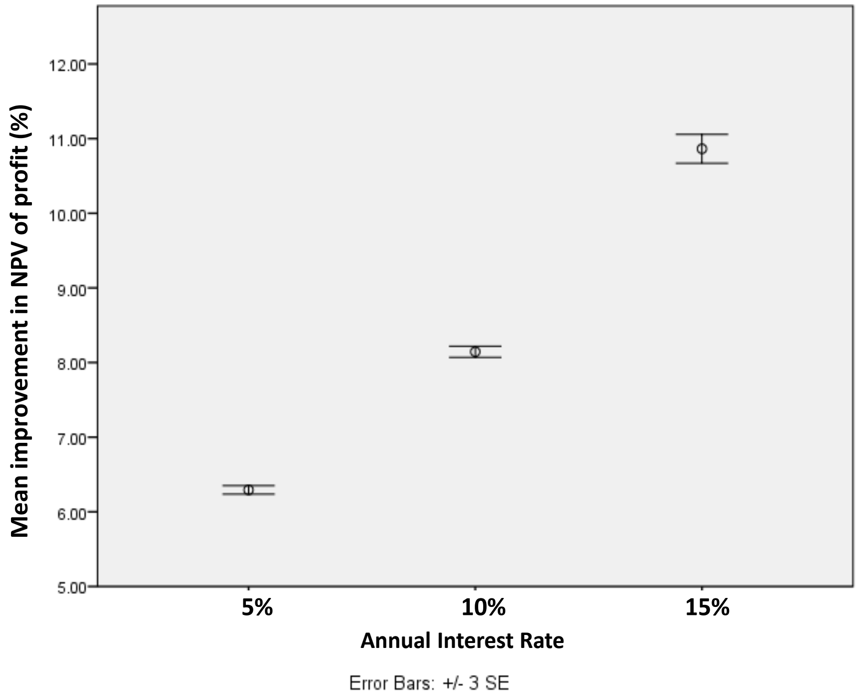
| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 25606847.633 | 3 | 8535615.878 | 17471.365 | 0.000 |
| Within Groups | 439692053.100 | 899996 | 488.549 | | |
| Total | 465298900.733 | 899999 | | | |

*<u>Interest rate</u>*

There are three levels of annual interest rates: 5%, 10%, and 15%. The error bar chart in Figure 1.11 shows that average profit improvement and variation in performance increase with interest rates. The ABC approach overall has much higher order-up-to inventory levels, resulting in higher inventory management cost and higher capital cost. Thus, as the interest rate grows, the ABC method performs significantly worse than the MPIC-NSD model which directly considers capital costs. The experiments on the interest rate indicate:

*Finding 2: The benefit of the MPIC-NSD model relative to the ABC approach increases with interest rate.*

*Figure 1.11: Improvement in NPV of expected profit by interest rates*




*<u>Inventory holding cost percentage</u>*

Inventory holding cost is measured by a percentage of the value of the goods, including warehousing cost, insurance, shrinkage, obsolescence cost (in relation to shelf life), etc., but excludes interest rate in this chapter. Five levels of annual holding cost percentages are examined. Figure 1.12 and Table 1.14 demonstrate that both mean profit improvement and variation in performance increase with inventory holding cost. The higher the inventory holding cost, the more significant is the benefit of the MPIC-NSD model. This is not surprising since the ABC method does not take inventory holding cost into consideration in either its classification procedure or the class CSL decision, while the MPIC-NSD model has incorporated inventory costs in determining classification and setting inventory policies. The results from the experiments confirm:

*Finding 3: The benefit of the MPIC-NSD model relative to the ABC approach increases with inventory holding cost.*

*Figure 1.12: Improvement in NPV of expected profit by inventory holding costs*

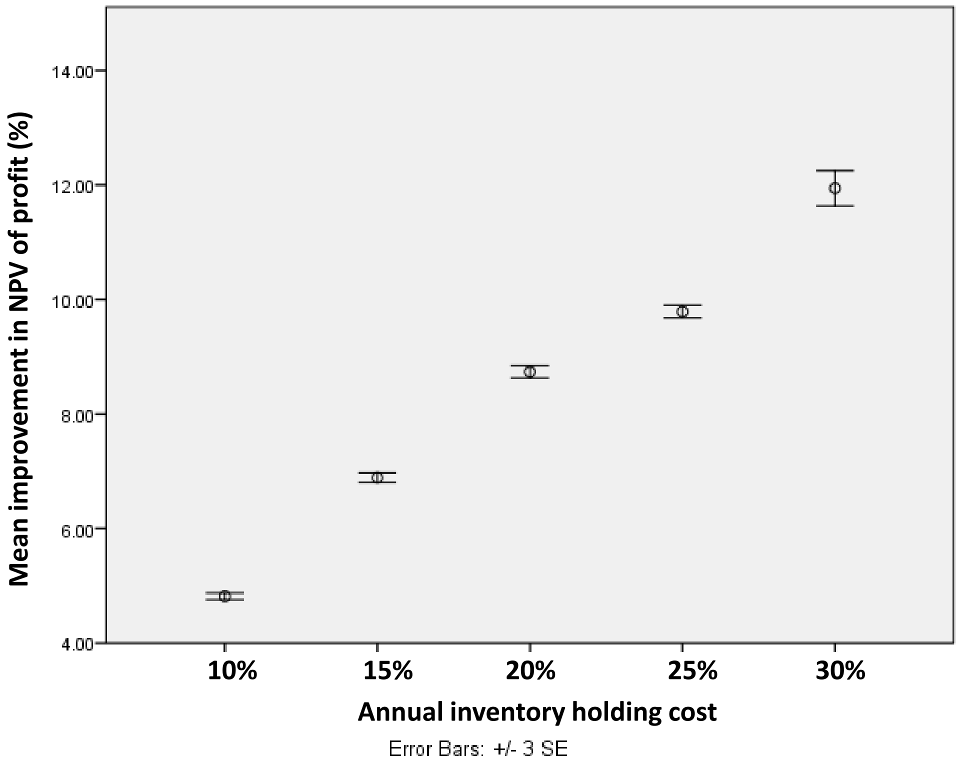



*Table 1.14: Mean comparison by inventory holding costs*

| Annual Inv. Holding Cost | Mean | N | Std. Deviation | Std. Error of Mean |
|---|---|---|---|---|
| 10% | 4.8151 | 180000 | 8.14084 | .01919 |
| 15% | 6.8879 | 180000 | 12.06354 | .02843 |
| 20% | 8.7364 | 180000 | 15.01383 | .03539 |
| 25% | 9.7874 | 180000 | 15.95876 | .03762 |
| 30% | 11.9426 | 180000 | 43.16698 | .10175 |
| Total | 8.4339 | 900000 | 22.73762 | .02397 |

*Replenishment lead time*

From Figure 1.13, replenishment lead time appears to have very different effects on the model performance. When lead time is one time unit, the MPIC-NSD model has marginal improvement on profit compared with the ABC approach. As the lead time becomes longer however, the increase in profit is much more significant. One reason is that the traditional order-up-to inventory model sets a much higher inventory level as the lead time increases, regardless of how frequently the inventory is replenished, while the multi-period model is able to adjust the order quantity and target inventory level based on the order frequency. From the experiments, it is observed that:

*Finding 4: The benefit of the MPIC-NSD model relative to the ABC approach increases with replenishment lead time.*

*Figure 1.13: Improvement in NPV of expected profit by replenishment lead time*

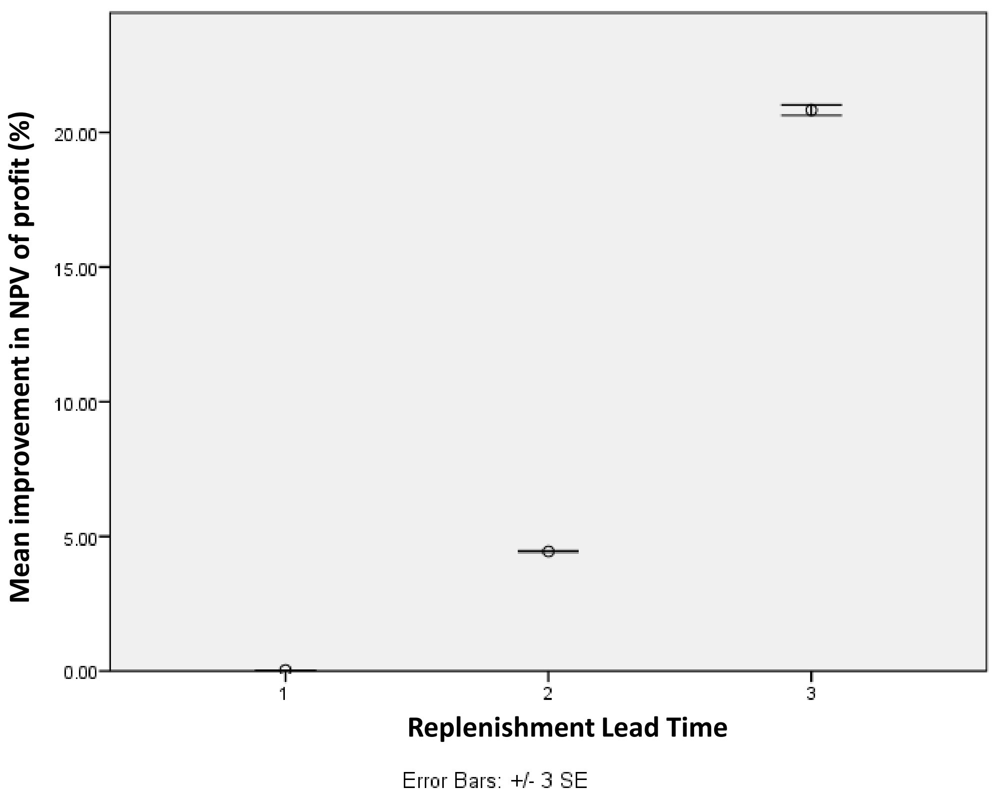


*Table 1.15: Mean comparison by replenishment lead time*

| Lead Time | Mean | N | Std. Deviation | Std. Error of Mean |
|---|---|---|---|---|
| 1 | .0295 | 300000 | .78711 | .00144 |
| 2 | 4.4412 | 300000 | 3.69789 | .00675 |
| 3 | 20.8309 | 300000 | 36.00618 | .06574 |
| Total | 8.4339 | 900000 | 22.73762 | .02397 |

<u>*Demand uncertainty*</u>

Demand uncertainty is reflected through five levels of coefficients of variation (CV): 5%, 10%, 15%, 20% and 25%. A simple error bar chart (Figure 1.14) shows that the level of demand variation is not linearly related to the profit improvement: 5% and 15% levels of uncertainty have little difference in terms of average profit improvement, and the difference is not statistically significant either (as shown in Table 1.16); but for the other levels of demand uncertainty, the extent to which the MPIC-NSD model outperforms the ABC approach appears to increase with the demand variation. The analysis of interaction effects among inventory parameters and the regression analysis in the latter section show that demand uncertainty does not have major impact on the model performance, although it is statistically important. The results from the experiments reveal that:

*Finding 5: The performance of the MPIC-NSD model relative to the ABC approach is only marginally affected by the demand uncertainty and the relationship appears nonlinear.*

*Figure 1.14: Improvement in NPV of expected profit by demand uncertainty*

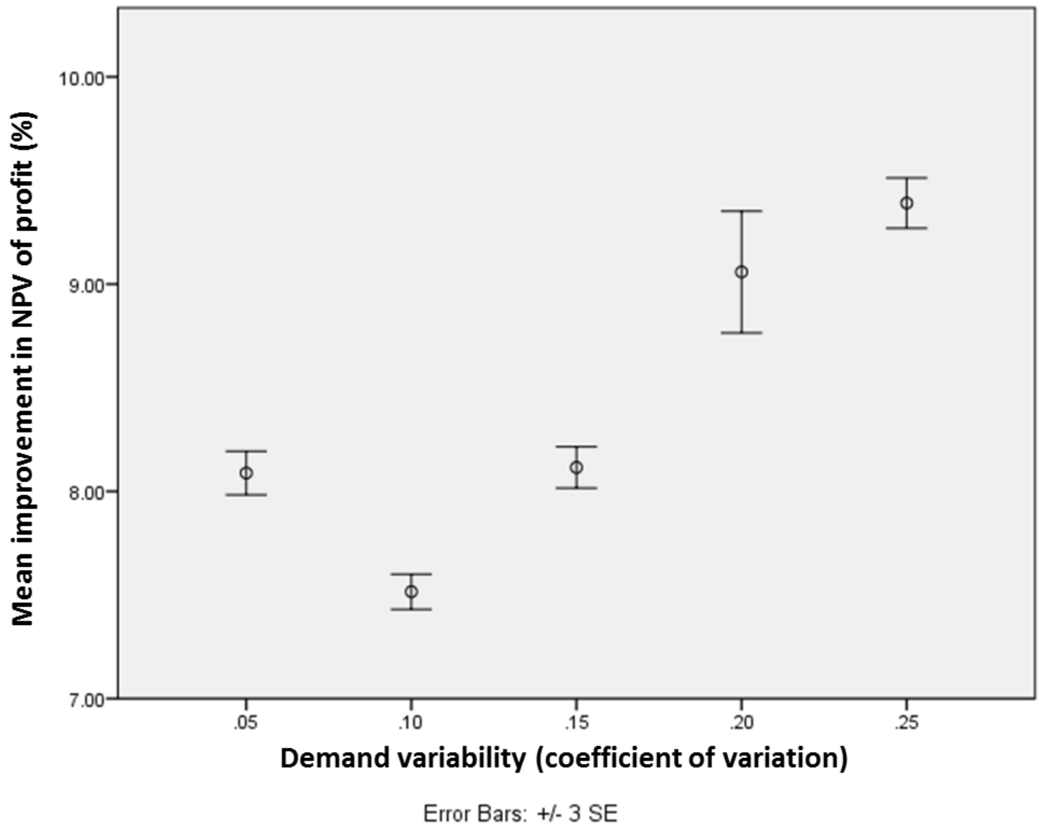


*Table 1.16: Tukey Honestly Significant Difference Test of demand uncertainty*

| (I) Demand_Uncertainty | | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .05 | .10 | .57278* | .07576 | .000 | .3661 | .7794 |
| | .15 | -.02690 | .07576 | .997 | -.2336 | .1797 |
| | .20 | -.97020* | .07576 | .000 | -1.1768 | -.7635 |
| | .25 | -1.30280* | .07576 | .000 | -1.5094 | -1.0961 |
| .10 | .05 | -.57278* | .07576 | .000 | -.7794 | -.3661 |
| | .15 | -.59968* | .07576 | .000 | -.8063 | -.3930 |
| | .20 | -1.54297* | .07576 | .000 | -1.7496 | -1.3363 |
| | .25 | -1.87557* | .07576 | .000 | -2.0822 | -1.6689 |
| .15 | .05 | .02690 | .07576 | .997 | -.1797 | .2336 |
| | .10 | .59968* | .07576 | .000 | .3930 | .8063 |
| | .20 | -.94330* | .07576 | .000 | -1.1499 | -.7366 |
| | .25 | -1.27589* | .07576 | .000 | -1.4825 | -1.0692 |
| .20 | .05 | .97020* | .07576 | .000 | .7635 | 1.1768 |
| | .10 | 1.54297* | .07576 | .000 | 1.3363 | 1.7496 |
| | .15 | .94330* | .07576 | .000 | .7366 | 1.1499 |
| | .25 | -.33260* | .07576 | .000 | -.5392 | -.1259 |
| .25 | .05 | 1.30280* | .07576 | .000 | 1.0961 | 1.5094 |
| | .10 | 1.87557* | .07576 | .000 | 1.6689 | 2.0822 |
| | .15 | 1.27589* | .07576 | .000 | 1.0692 | 1.4825 |
| | .20 | .33260* | .07576 | .000 | .1259 | .5392 |

*. The mean difference is significant at the 0.05 level.

## 5.2 Interactive Effects on Model Performance

<u>*Interaction of demand pattern with lead time*</u>

The plot in Figure 1.15 shows the marginal mean profit improvements at the factor combinations of the demand pattern and lead time. The fact that the seasonal lines are not parallel to each other indicates a strong interaction effect between the two predictor variables. When the replenishment lead time is one time unit, the model performance appears to be independent of demand pattern; while when the lead time is long, the MPIC-NSD model shows significant advantage for certain types of demand patterns.

Table 1.17 confirms the statistical and practical significances of two factors and their interaction. *p*-values for demand pattern, lead time, and their interaction are all less than .05, indicating they are statistically significant in explaining variation in profit improvement. The partial eta squared statistic reports the practical significance of each term: demand pattern accounts for approximately 7.3% of the variation, lead time accounts for 18.1% and their interaction accounts for 11.3%. The results indicate that:

*Finding 6: The effect of demand pattern on the performance of the MPIC-NSD model relative to the ABC approach increases with lead time.*

*Figure 1.15: Interaction of demand pattern with lead time*

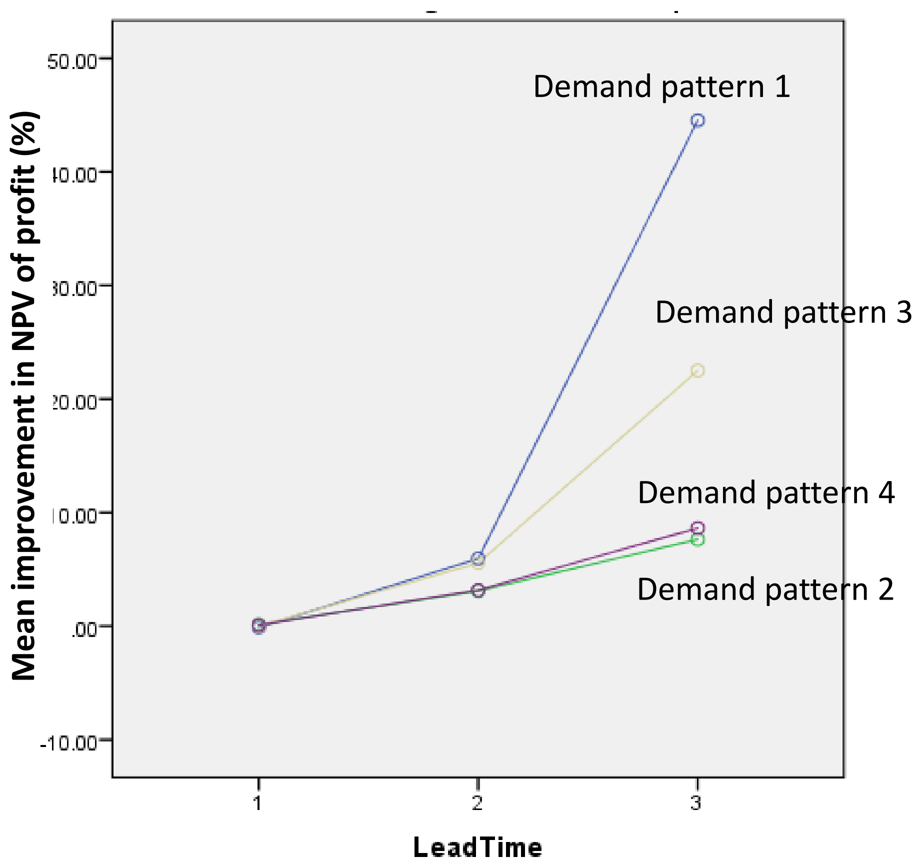


*Table 1.17: Test of between-subjects effects: individual and interactive effects of demand pattern and lead time*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 139195515.641[a] | 11 | 12654137.786 | 34923.195 | 0.000 | .299 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 176676.802 | 0.000 | .164 |
| Demand_Pattern | 25606847.633 | 3 | 8535615.878 | 23556.799 | 0.000 | .073 |
| Lead_Time | 72078595.903 | 2 | 36039297.951 | 99462.125 | 0.000 | .181 |
| Demand_Pattern * Lead_Time | 41510072.105 | 6 | 6918345.351 | 19093.417 | 0.000 | .113 |
| Error | 326103385.093 | 899988 | 362.342 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .299 (Adjusted R Squared = .299)

<u>*Interaction of demand pattern with interest rate*</u>

Figure 1.16 shows that the interaction effect exists between the demand pattern and interest rate. The MPIC-NSD model outperforms the ABC approach in all the twelve scenarios, and the gap grows with interest rate. The difference in profit improvement between demand pattern (1) and (3) is greater at an interest rate of 15% than that of 10%. When the demand follows pattern (2) or (4), the benefit of the optimization model only increases slightly with the interest rate. The tests of between-subjects effects confirm the statistical importance of both factors and their interaction, and demand pattern appears to be more influential than the interest rate as suggested by partial eta squared values in Table 1.18. The results suggest that:

*Finding 7: The effect of demand pattern on the performance of the MPIC-NSD model relative to the ABC approach increases with interest rate.*

*Figure 1.16: Interaction of demand pattern with interest rate*

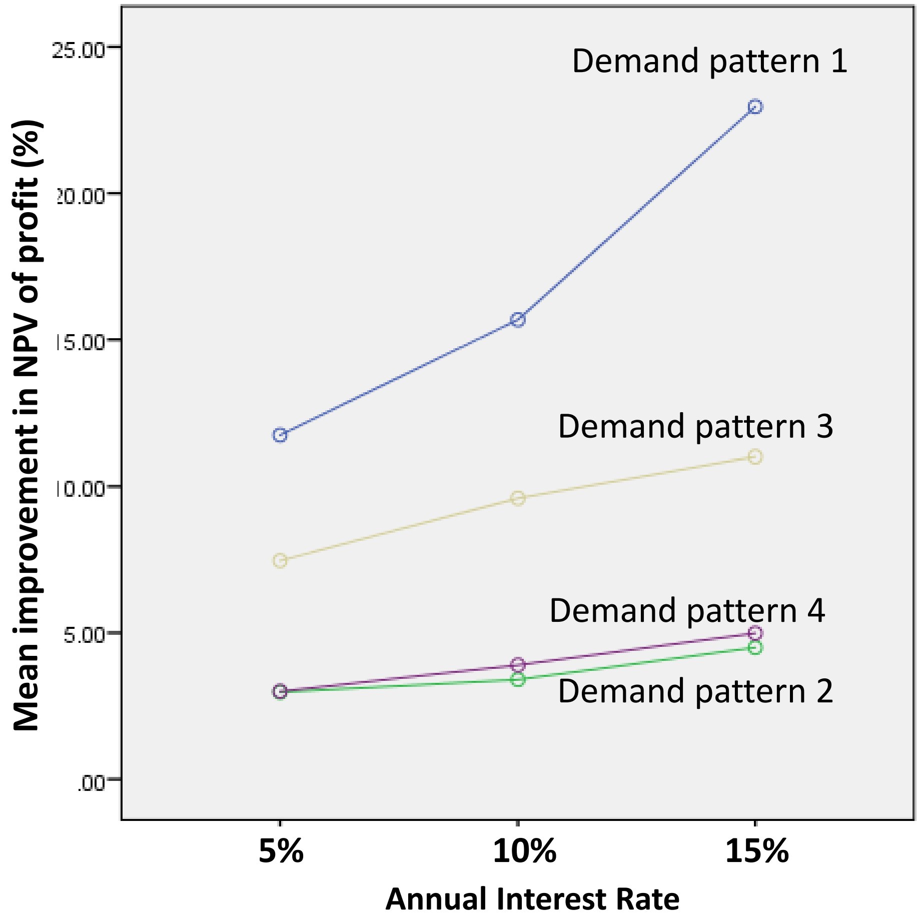



*Table 1.18: Test of between-subjects effects: individual and interactive effects of demand pattern and interest rate*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 31185588.224[a] | 11 | 2835053.475 | 5877.530 | 0.000 | .067 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 132718.582 | 0.000 | .129 |
| Demand_Pattern | 25606847.633 | 3 | 8535615.878 | 17695.730 | 0.000 | .056 |
| Interest_Rate | 3169464.424 | 2 | 1584732.212 | 3285.409 | 0.000 | .007 |
| Demand_Pattern * Interest_Rate | 2409276.167 | 6 | 401546.028 | 832.471 | 0.000 | .006 |
| Error | 434113312.509 | 899988 | 482.355 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .067 (Adjusted R Squared = .067)

*Interaction of demand pattern with inventory holding cost*

The effect of inventory holding cost on the model performance remains largely the same for the stable demand and the demand that is high towards the end of the year. When the peak of the demand happens in the first quarter (demand pattern (1)), higher inventory holding cost puts the ABC approach in a particular disadvantage. Demand pattern, the inventory holding cost, and their interaction are all statistically significant (*p*-value less than .001), with demand pattern showing a relatively stronger effect. Figure 1.17 and Table 1.19 demonstrate that:

*Finding 8: The effect of demand pattern on the performance of the MPIC-NSD model relative to the ABC approach increases with inventory holding cost.*

*Figure 1.17: Interaction of demand pattern with inventory holding cost*

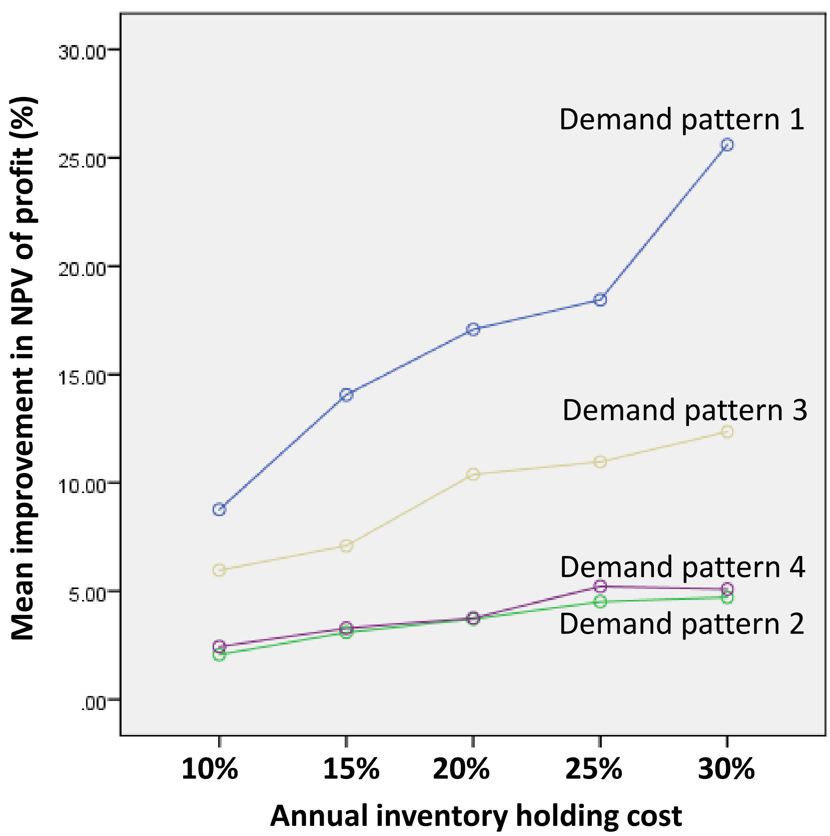


*Table 1.19: Test of between-subjects effects: individual and interactive effects of demand pattern and inventory holding cost*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 34239515.667[a] | 19 | 1802079.772 | 3762.442 | 0.000 | .074 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 133657.665 | 0.000 | .129 |
| Demand_Pattern | 25606847.633 | 3 | 8535615.878 | 17820.940 | 0.000 | .056 |
| Inv. Holding Cost % | 5349668.183 | 4 | 1337417.046 | 2792.303 | 0.000 | .012 |
| Demand_Pattern * Inv. Holding Cost % | 3282999.851 | 12 | 273583.321 | 571.196 | 0.000 | 0.76% |
| Error | 431059385.067 | 899980 | 478.966 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .074 (Adjusted R Squared = .074)

<u>*Interaction of demand pattern with demand uncertainty*</u>

The interaction between demand pattern and demand uncertainty is statistically significant as demonstrated in Figure 1.18 and Table 1.20 but from the perspective of practical impact, the partial eta squared statistic shows that neither demand uncertainty nor its interaction with demand pattern can sufficiently explain the variation in model performance. It’s concluded that:

*Finding 9: The interaction of demand pattern and demand uncertainty only has marginal impact on the performance of the MPIC-NSD model relative to the ABC approach.*

*Figure 1.18: Interaction of demand pattern with demand uncertainty*

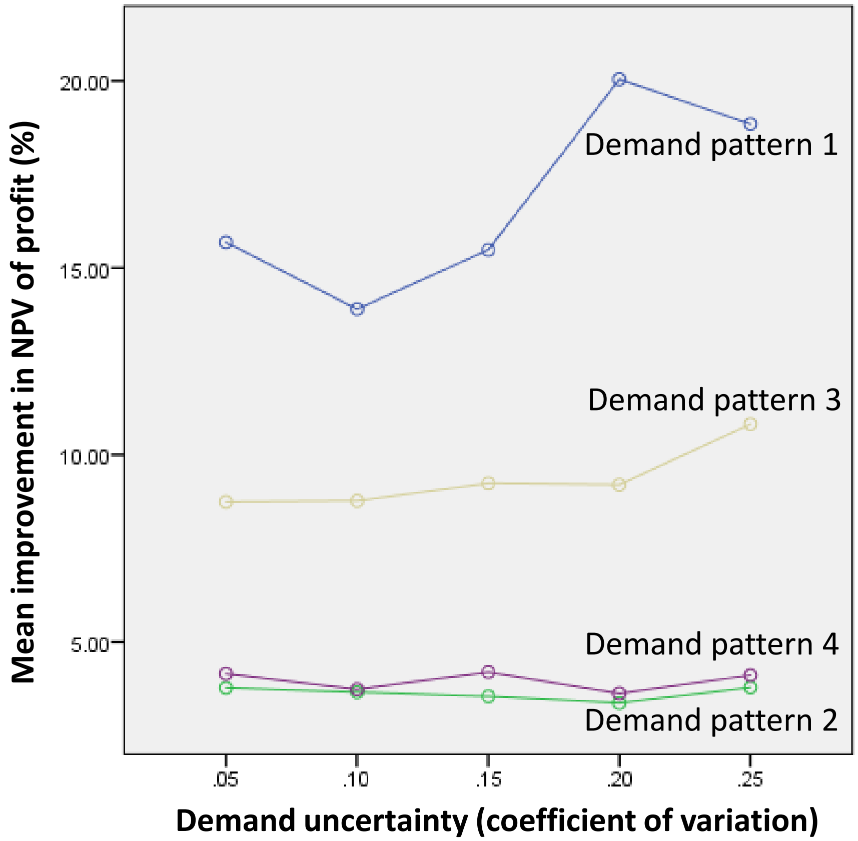



*Table 1.20: Test of between-subjects effects: individual and interactive effects of demand pattern and demand uncertainty*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 26930706.675[a] | 19 | 1417405.614 | 2909.966 | 0.000 | .058 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 131429.223 | 0.000 | .127 |
| Demand_Pattern | 25606847.633 | 3 | 8535615.878 | 17523.816 | 0.000 | .055 |
| Demand_Variation | 426738.334 | 4 | 106684.584 | 219.026 | .000 | .001 |
| Demand_Pattern * Demand_Variation | 897120.707 | 12 | 74760.059 | 153.484 | 0.000 | .002 |
| Error | 438368194.059 | 899980 | 487.087 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .058 (Adjusted R Squared = .058)

*<u>Interaction of lead time with inventory holding cost</u>*

Figure 1.19 shows that when replenishment lead time is one time unit, the model performance is insensitive to the inventory holding cost. As the lead time increases, the effect of inventory holding cost becomes more significant. When the lead time is long and inventory holding cost is high, the MPIC-NSD method presents significant advantage over the ABC approach. The results indicate that:

*Finding 10: The effect of inventory holding cost on the performance of the MPIC-NSD model relative to the ABC approach increases with lead time.*

*Figure 1.19: Interaction of lead time with inventory holding cost*

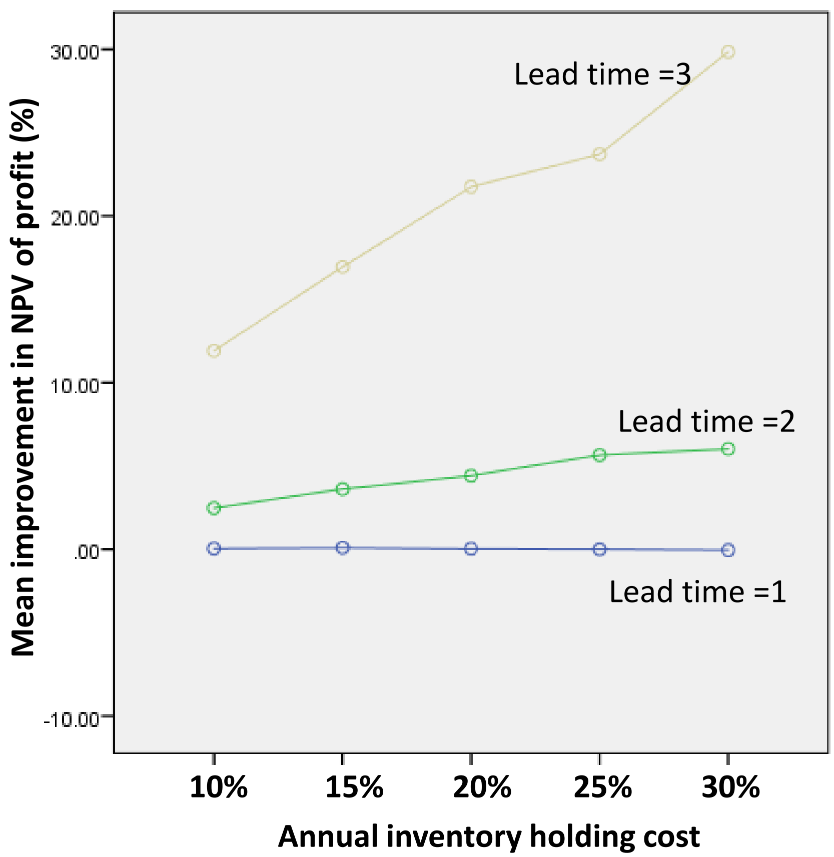


*Table 1.21: Test of between-subjects effects: individual and interactive effects of lead time and inventory holding cost*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 83682840.484[a] | 14 | 5977345.749 | 14096.685 | 0.000 | .180 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 150975.593 | 0.000 | .144 |
| Lead_Time | 72078595.903 | 2 | 36039297.951 | 84993.351 | 0.000 | .159 |
| Inv. Holding Cost % | 5349668.183 | 4 | 1337417.046 | 3154.100 | 0.000 | .014 |
| Lead_Time * Inv. Holding Cost % | 6254576.399 | 8 | 781822.050 | 1843.812 | 0.000 | .016 |
| Error | 381616060.249 | 899985 | 424.025 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .180 (Adjusted R Squared = .180)

*Interaction of lead time with interest rate*

From Figure 1.20, interest rate affects the model performance only when lead time is longer than one time unit. The longer the lead time, the more significant impact the interest rate has on the profit improvement. The tests of between-subjects effects confirm the statistical importance of both factors and their interaction, and lead time presents a much higher impact. The results suggest that:

*Finding 11: The effect of interest rate on the performance of the MPIC-NSD model relative to the ABC approach increases with lead time.*

*Figure 1.20: Interaction of lead time with interest rate*

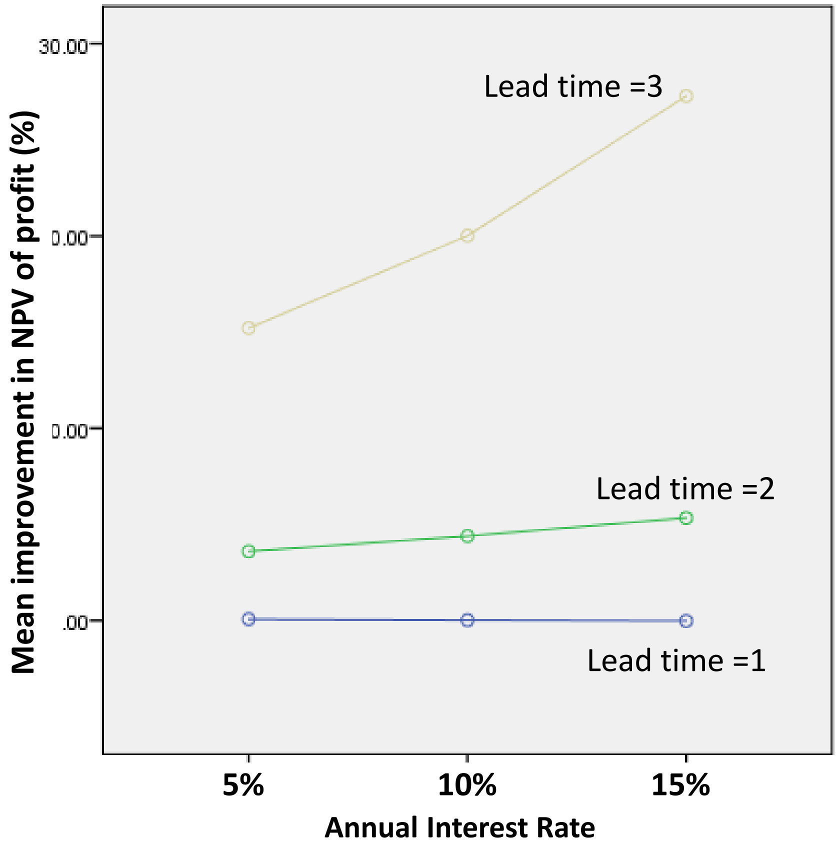



*Table 1.22: Test of between-subjects effects: individual and interactive effects of lead time and interest rate*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 79606145.883[a] | 8 | 9950768.235 | 23219.523 | 0.000 | .171 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 149380.808 | 0.000 | .142 |
| LeadTime | 72078595.903 | 2 | 36039297.951 | 84095.549 | 0.000 | .157 |
| InterestRate | 3169464.424 | 2 | 1584732.212 | 3697.878 | 0.000 | .008 |
| LeadTime * InterestRate | 4358085.556 | 4 | 1089521.389 | 2542.333 | 0.000 | .011 |
| Error | 385692754.851 | 899991 | 428.552 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .171 (Adjusted R Squared = .171)

*Interaction of lead time with demand uncertainty*

Figure 1.21 shows that the model performance is insensitive to the demand uncertainty when the lead time is only one time unit, similar to the effect of inventory holding cost and interest rate. Overall, while the individual effect of demand uncertainty and its interaction effect with lead time appear to be statistically important, they do not exert much practical influence on the model performance, as indicated in Table 1.23. It is concluded that:

*Finding 12: The interaction of demand uncertainty and lead time only has marginal effect on the performance of the MPIC-NSD model relative to the ABC approach.*

*Figure 1.21: Interaction of lead time with demand uncertainty*

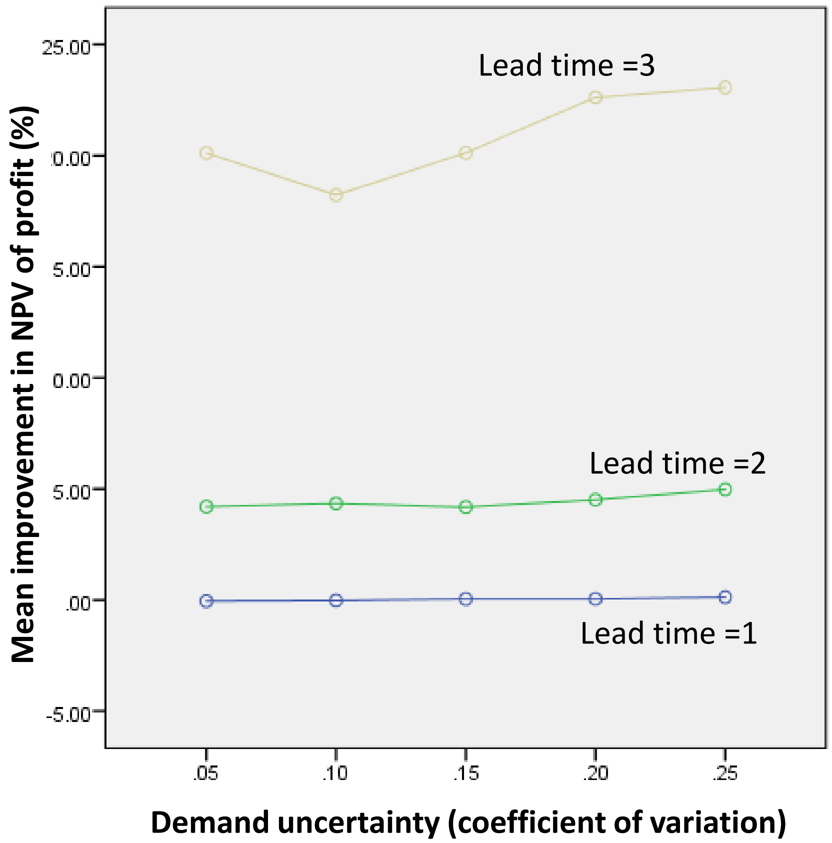



*Table 1.23: Test of between-subjects effects: individual and interactive effects of lead time and demand uncertainty*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 73065208.019[a] | 14 | 5218943.430 | 11974.929 | 0.000 | .157 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 146888.736 | 0.000 | .140 |
| LeadTime | 72078595.903 | 2 | 36039297.951 | 82692.609 | 0.000 | .155 |
| Variation | 426738.334 | 4 | 106684.584 | 244.789 | .000 | .001 |
| LeadTime * Variation | 559873.782 | 8 | 69984.223 | 160.580 | .000 | .001 |
| Error | 392233692.714 | 899985 | 435.822 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .157 (Adjusted R Squared = .157)

*<u>Interaction of interest rate with inventory holding cost</u>*

Figure 1.22 shows that the benefit of the MPIC-NSD model increases with inventory holding cost and interest rate, but little variation in the model performance is accounted for by the interaction of two variables (as indicated in Table 1.24) though it is statistically significant. It is concluded that:

*Finding 13: The interaction of inventory holding cost and interest rate only has marginal effect on the performance of the MPIC-NSD model relative to the ABC approach.*

*Figure 1.22: Interaction of Interest rate with inventory holding cost*

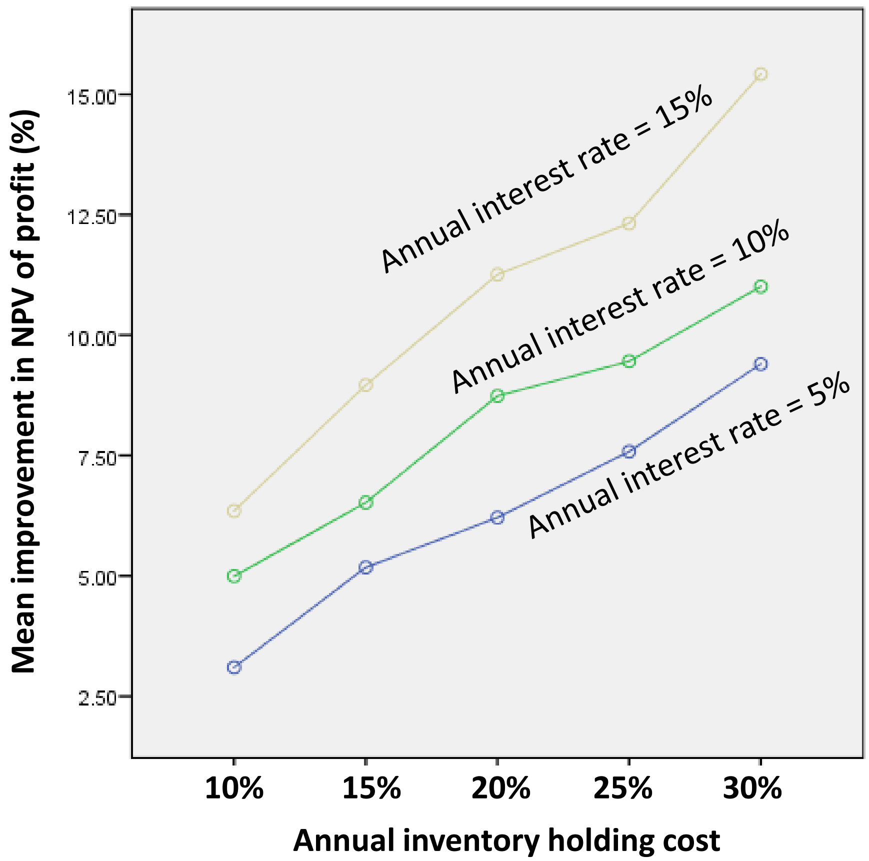


*Table 1.24: Test of between-subjects effects: individual and interactive effects of interest rate and inventory holding cost*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 8726103.733[a] | 14 | 623293.124 | 1228.620 | 0.000 | .019 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 126189.540 | 0.000 | .123 |
| Interest_Rate | 3169464.424 | 2 | 1584732.212 | 3123.785 | 0.000 | .007 |
| Inv. Holding Cost % | 5349668.183 | 4 | 1337417.046 | 2636.283 | 0.000 | .012 |
| Interest_Rate * Inv. Holding Cost % | 206971.127 | 8 | 25871.391 | 50.997 | .000 | .000 |
| Error | 456572797.000 | 899985 | 507.312 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .019 (Adjusted R Squared = .019)

<u>*Interaction of interest rate with demand uncertainty*</u>

From Figure 1.23, the MPIC-NSD model significantly outperforms the ABC approach when interest rate is 15% and coefficient of variation of demand is 20%. The MPIC-NSD model in general provides higher benefit as the interest rate increases, but does not evidence a linear relationship with demand uncertainty for any level of interest rate. The interaction effect between interest rate and demand uncertainty is marginal on the model performance as shown in Table 1.25. Statistically it can be concluded that:

*Finding 14: The effect of demand uncertainty on the performance of the MPIC-NSD model relative to the ABC approach increases with interest rate.*

*Figure 1.23: Interaction of Interest rate with demand uncertainty*

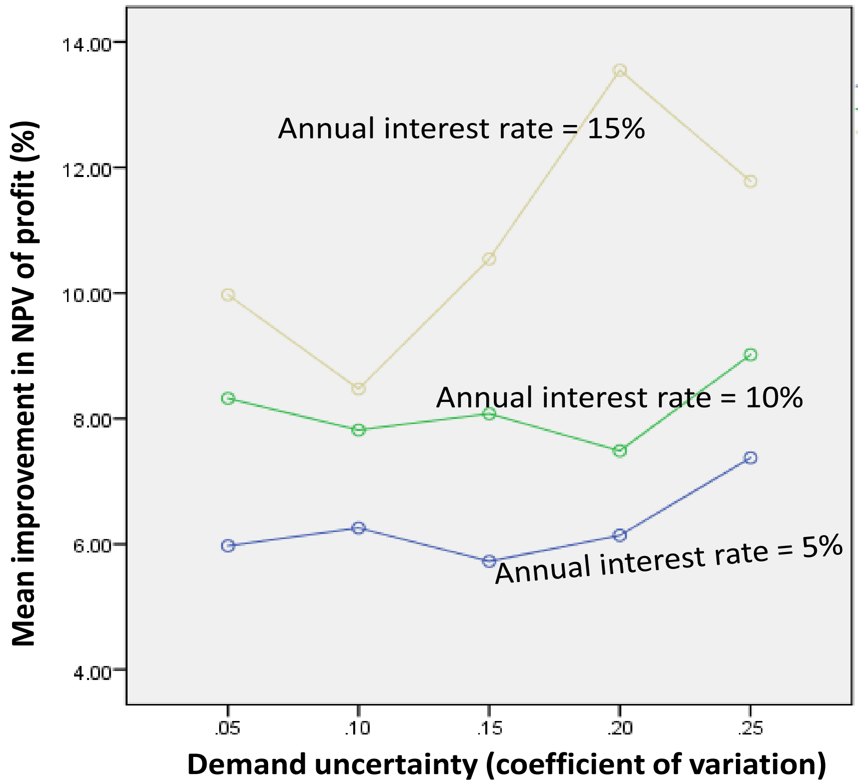



*Table 1.25: Test of between-subjects effects: individual and interactive effects of interest rate and demand uncertainty*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 4227304.561[a] | 14 | 301950.326 | 589.390 | 0.000 | .009 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 124958.275 | 0.000 | .122 |
| InterestRate | 3169464.424 | 2 | 1584732.212 | 3093.305 | 0.000 | .007 |
| Variation | 426738.334 | 4 | 106684.584 | 208.242 | .000 | .001 |
| InterestRate * Variation | 631101.803 | 8 | 78887.725 | 153.984 | .000 | .001 |
| Error | 461071596.173 | 899985 | 512.310 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .009 (Adjusted R Squared = .009)

*<u>Interaction of inventory holding cost with demand uncertainty</u>*

Holding inventory holding cost constant, demand uncertainty in general is not strictly positively related to the model performance except for the annual holding cost of 15%

and 20%. While demand uncertainty and its interaction with inventory holding cost are both statistically significant to the model performance, they only account for a tiny portion of the variation. Figure 1.24 shows that:

*Finding 15: The effect of demand uncertainty on the performance of the MPIC-NSD model relative to the ABC approach changes with inventory holding cost.*

*Figure 1.24: Interaction of demand uncertainty with inventory holding cost*

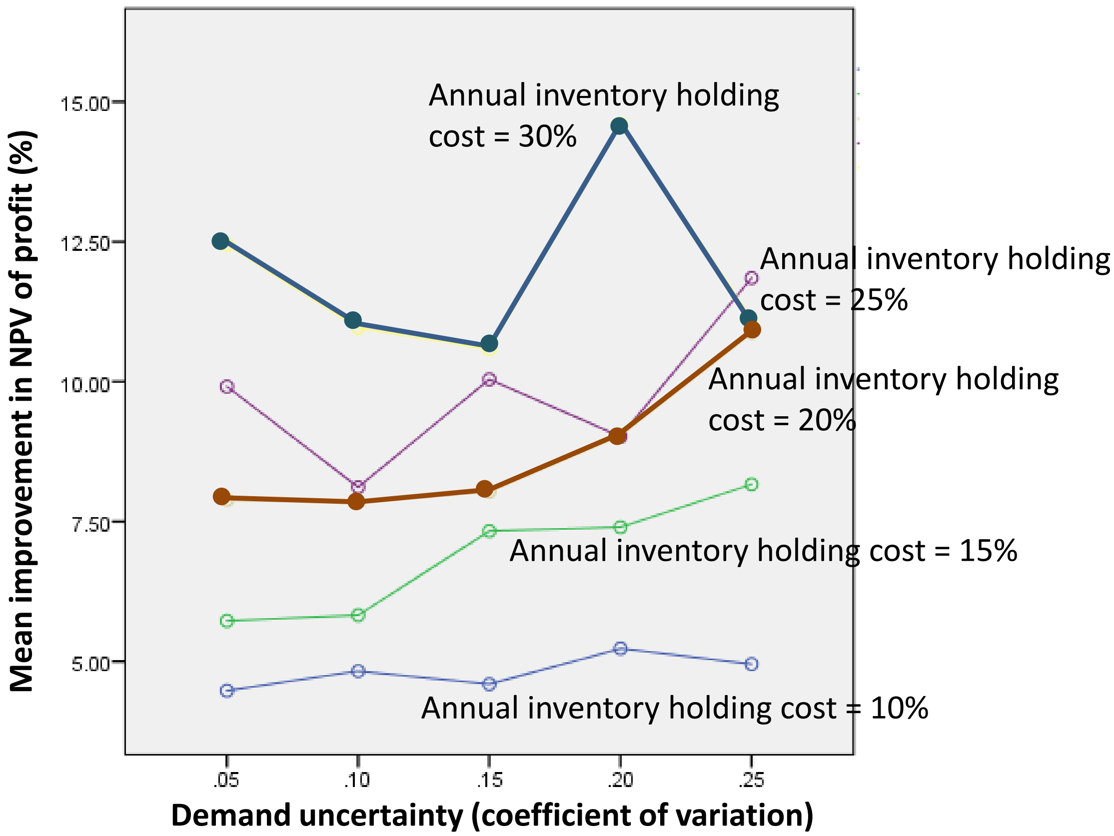


*Table 1.26: Test of between-subjects effects: individual and interactive effects of demand uncertainty and inventory holding cost*

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 6435721.510[a] | 24 | 268155.063 | 525.936 | 0.000 | .014 |
| Intercept | 64017412.714 | 1 | 64017412.714 | 125558.279 | 0.000 | .122 |
| Inv. Holding Cost % | 5349668.183 | 4 | 1337417.046 | 2623.095 | 0.000 | .012 |
| Demand Variation | 426738.334 | 4 | 106684.584 | 209.242 | .000 | .001 |
| Inv. Holding Cost % * Demand Variation | 659314.993 | 16 | 41207.187 | 80.820 | .000 | 0.14% |
| Error | 458863179.223 | 899975 | 509.862 | | | |
| Total | 529316313.447 | 900000 | | | | |
| Corrected Total | 465298900.733 | 899999 | | | | |

a. R Squared = .014 (Adjusted R Squared = .014)

## 5.3 Regression Results

Tests of between-subjects effects show that all five parameters and their interactions are statistically important ($p < .001$). Lead time presents the strongest effect on the model performance, followed by demand pattern, inventory holding cost, and interest rate. Demand uncertainty shows the least impact, accounting for less than 1% of the variation in the model performance. Among all the interaction terms, the interaction between lead time and demand pattern has the greatest effect, followed by the interaction between lead time and inventory holding cost, the interaction between lead time and interest rate, and the interaction between demand pattern and inventory holding cost. All the other interactions explain less than 1% of the variation, respectively.

Model 1 in Table 1.27 includes two control variables, unit cost and gross margin. They both are statistically important, but the former can only explain less than .03% of the variance. Model 2 therefore consists of gross margin only plus all the categorical variables and their interactions. Both models have an adjusted R square of .371.

*Table 1.27: Partial Eta Squared Statistics of multiple moderated regressions*

| | Dependent variable: profit improvement % | |
|---|---|---|
| **Independent Variables** | *Model 1* | *Model 2* |
| *Control variables* | | |
| Unit cost | .000 | |
| Gross margin % | .017 | .017 |
| *Direct effects* | | |
| Lead time | .199 | .199 |
| Demand pattern | .079 | .080 |
| Inventory holding cost % | .020 | .020 |
| Interest rate | .011 | .011 |
| Demand uncertainty | .002 | .002 |
| *Interactions* | | |
| Lead time X Demand pattern | .127 | .127 |
| Lead time X Inventory holding cost % | .024 | .023 |
| Lead Time X Interest rate | .013 | .013 |
| Demand pattern X Inventory holding cost % | .011 | .011 |
| Demand pattern X Interest rate | .008 | .008 |
| Interest rate X demand uncertainty | .002 | .002 |
| Demand pattern X demand uncertainty | .002 | .002 |
| Inventory holding cost % X Demand uncertainty | .002 | .002 |
| Lead time X Demand uncertainty | .001 | .001 |
| Inventory holding cost % X interest rate | .001 | .001 |
| Adjusted $R^2$ | .371 | .371 |

## 6 Conclusions

The purposes of this chapter are threefold: a) to construct a multi-period inventory optimization model that explicitly addresses nonstationary, stochastic demand, and allows SKU-specific lead time, arbitrary review period and inventory holding cost, with the objective to maximize the NPV of expected profit for the inventory under a periodic-review, order-up-to inventory policy; b) to investigate the performance of the optimization model in comparison with the multi-criteria inventory classification approach through a real world case; and c) to examine the impact of key inventory parameters (including lead time, interest rate, inventory holding cost, demand pattern, and demand uncertainty) on the profit improvement of the optimization model over the commonly used ABC approach.

The proposed optimization model leverages the NETFORM framework to determine the CSL for each SKU in each time period. The arrival time of a replenishment order for a particular SKU is derived from its lead time and review/order frequency, and is predetermined in the optimization model. The output of the model consists of SKU and time specific CSLs, replenishment order quantities, expected safety stocks, and expected satisfied demands.

## 6.1 Findings and Managerial Implications

The evaluation of the MPIC-NSD model relative to the MCIC approach is conducted for a real-life company. The results show an improvement of 2.75% in the NPV of profit, with 13% less inventory capital required on average. Experiments that compare the model to the optimized ABC approach display an average improvement of 7.53% for the dataset consisting of 900 SKUs. The analysis of different parameters shows that the benefit of the optimization model increases significantly with the cost of capital, lead times, and inventory holding costs. A substantial improvement in a company's bottom line can be expected by transforming from the ABC approach to an integrated inventory management process if the average lead time is more than one time unit. The impacts on profit of demand patterns and demand uncertainties, however, require investigation on a case-by-case basis.

The proposed multi-period optimization model provides companies with a decision-support tool to manage inventory for profit maximization. In addition to achieving a superior financial performance relative to the ABC and MCIC approach, practitioners may also benefit from this research in the following ways. First, the MPIC-NSD model provides management with the advanced knowledge about the target inventory levels, possible safety stocks, and replenishment order quantities in each future time period at the beginning of the planning horizon. Such information about the future would greatly facilitate the negotiation process with suppliers and support production planning. Second, companies may use the anticipated CSLs to communicate with customers and form contracts. If a customer requires a higher CSL than what the model suggests, management could explore the possibility to negotiate for higher margins or may demand more predictable, stable order quantities for specific SKUs. The CSLs of an SKU are determined mainly upon its profitability, demand pattern and uncertainty, and its holding cost percentage, relative to the other SKUs. Finally, management could utilize the model to plan for the inventory capital required in each time period, and could also evaluate the impacts of different levels of inventory investments and costs of capital on the NPV of profit.

### 6.2 Contributions

This research makes several contributions to the inventory management literature. First, it simultaneously addresses inventory classification and policy setting under nonstationary demand with a linear optimization model. It advances the fieldwork of bridging two distinct streams of inventory research, inventory classification and inventory optimization, by offering a practical inventory modeling framework. This study substantially extends the previous work done by Teunter et al. (2010), Mohammaditabar et al. (2012), Millstein et al. (2014), etc., by explicitly accounting for various practical factors, including; nonstationary demands, arbitrary review periods, different replenishment lead times, diverse inventory holding costs, and the time value of money.

Second, this research offers an alternative method to measure inventory performance, the NPV of profit, instead of the more commonly used cost metric. The prevailing use of the cost metric in the inventory literature can be partially attributed to the separation of inventory classification from inventory policy setting, with the former deemed to set product target service levels (though arbitrarily) and the latter determining inventory policies under predetermined service levels. Only the simultaneous consideration of the two allows a company to maximize its profit through balancing service levels and costs. Cost minimization models typically do not consider an SKU's contribution to the company's bottom line and often mislead management to set the lowest inventory budget possible subject to the service constraint. In addition, classical inventory research assumes a constant value of money throughout the planning horizon, though in reality capital always has an opportunity cost. For industries which have relatively high weighted-average cost of capital, it would be beneficial for the company to consider the time value of money in their inventory modeling.

Third, the research provides insights into under what situations an ABC approach may be sufficient and what circumstances an integrated model is strongly recommended. In general, if the replenishment lead time is short (e.g., one time unit), the ABC approach is acceptable as the multi-period optimization model may only provide marginal improvement. However, the company should combine the ABC approach with a dynamic inventory policy setting when facing nonstationary demand, and the service level of a class needs to be optimized for profit maximization. If the lead time is longer than one time unit, the effects of the cost of capital and inventory holding costs would be amplified and the company should consider employing the integrated optimization model to improve its NPV of profit.

## 6.3 Limitations and Future Research

Although the proposed model captures many of the real world complexities, there are a number of areas worth investigating. First, future demand often not only reflects the past demand pattern, but may also be endogenously affected by the service level offered today. As a company moves away from the current ABC approach and the corresponding service levels to implement the solution recommended by the MPIC-NSD model, the expected demands, upon which the solution is based, may have changed. Extending the model to account for the impact of different service levels on expected demands would be of value.

Second, replenishment lead times may vary systematically or stochastically through time. Future research could extend the model to include varying lead times or take into account the nonlinear relationship between lead time and order quantity. Third, in a highly competitive market or for companies that provide functional products, a stockout often results in lost sales as modelled because consumers might switch to other brands or suppliers. But there are cases, especially for industrial products, where items might be backordered if an incoming order cannot be fully fulfilled. Including backorders in the model and adding an associated penalty cost would address this situation.

Fourth, the current model allows the expected CSL of an SKU to vary over time. There might be situations where a stable service level is desired or a minimum service level is required. Modifying the assignment decision variables from time-dependent to time-independent would solve the first problem, and adding a constraint to limit the minimum service level for a specific SKU would solve the second problem. The model may be also constrained to a limited number of inventory classes. Fifth, the ordering and management cost is generally negligible when the company uses sophisticated ERP system to assist in order placements, but there could be situation where such cost cannot be ignored, especially if the company replies on physical counting of inventory to ensure the accuracy of order quantities. Quantifying the ordering and management cost associated with the number of inventory classes and including such cost in the objective function would address this situation.

Sixth, the purchasing cost of an SKU may vary by order quantity because of economies of scale in transportation, production, etc. Having multi-tiered SKU costs based on replenishment quantity and allowing the model to determine the optimal ordering frequencies and quantities would be a valuable extension. Seventh, further studies on the integration of inventory classification with policy setting based on the different demand processes would provide further insights and comparison with the results of this work.

Lastly, maximizing the joint profits of the entire supply chain and comparing the solutions with the focal-firm-focused optimization results would offer new insights to the supply chain management literature and practice.

# Chapter 2 : Integrating Inventory Classification and Control Decisions to Maximize Order Fulfillment Measures

## 1 Introduction

Customer satisfaction and loyalty are critical for businesses to sustain long-term growth and profitability. Companies carrying a large variety of items typically receive customer orders consisting of different items in different quantities. Customers will not be satisfied unless all products in an order can be delivered right from inventory or within an agreed short period of time, especially in today's growing online retailing environment where delayed order fulfillment is seen to cause the decline of future order frequency and order size from the same customer (Rao et al., 2011). Order fulfillment performance is a major factor in customer satisfaction.

The traditional operational metric, item fill rate, defined as the percentage of total demand volume of SKUs that can be filled immediately from stock, is most commonly used in both practice and the inventory literature to gauge service level. Since the standard inventory models only consider item fill rate (IFR) in determining inventory parameters such as safety stock level, companies often use IFR to evaluate customer satisfaction with the order. However, it is important to realize that item fill rate is a producer/distributor-focused measure assessing internal operations (Zinn et al., 2002; Anupindi and Tayur, 1998) rather than a customer-centric indicator measuring the service received by customers. When a firm adopts an average item fill rate of 98%, for instance, does it mean 98% of the orders are completely filled, or on average 98% of units ordered in an order can be delivered off-the-shelf? An important assumption of using IFR as an order fulfillment measure is that the demands for each item are independent of each other (Hausman et al., 1998); e.g., a customer order contains one type of item only so that no interactions between items need to be taken into account, which obviously cannot hold for most businesses. Because different SKUs are usually given different inventory policies and customer orders are comprised of different mixes of products, the relationship between an item's fill rate and order fulfillment may not be straightforward. The simulations conducted by Anupindi and Tayur (1998) demonstrate that using item fill rates as an indicator of order fill rates does not perform well in the environment where demands are correlated over items, and Song (1998) further shows that when item-based inventory performance measures are satisfactory, order-based performance can be very poor.

In spite of the inadequacy and misleading nature of using IFR as an order-based measure, the inventory literature addressing the performance of order fulfillment is limited (Larsen and Thorstenson, 2008). The correlation of the demands among items makes the evaluation and optimization of order-based fill rates significantly more difficult than that of item-based fill rates, posing a considerable computational challenge (Song, 1998) and inhibiting its wide adoption by companies as a performance measure in inventory control. The existing research that considers order-based measures typically defines the order fill rate (OFR) as the percentage of the total number of customer orders that can be met in full immediately from inventory (e.g., Lu et al., 2003; Closs et al., 2010; Bowersox et al., 2012). What this measure focuses on is a binary characterization of a customer order – whether or not an order is completely satisfied, which is highly relevant to a firm's overall service performance, but has two major drawbacks. First, it does not reflect the depth of the demand satisfied by inventory on hand. Customers in most cases accept partially filled orders. An incomplete order disappoints the customer, but a 50%-filled order and a 98%-filled order obviously have drastically different effects on customer loyalty. This research thus proposes an alternative measure of order fulfillment, customer-order fill rate (CFR), defined as the percentage of total units in a customer order that can be delivered from on-hand inventory. CFR measures to what degree individual orders are satisfied. Second, the current research around OFR assumes the equal importance of all customer orders, which is rarely true in real business settings. And since most exiting models do not provide the service levels of individual orders, it is also impossible for them to take into account the importance of a particular customer. The presented OFR optimization model in this research offers a solution to this issue.

In addition to the development of new models to optimize order fulfillment measures which are of great importance to customer satisfaction, this research also performs extensive computational experiments to examine the relationships of average fill rate across orders, traditional order fill rate, and item fill rate, how their relationships are affected by inventory capital, and their tradeoffs with profitability. In the business world, companies need to simultaneously track several different measures of the service levels to stay competitive, albeit one of them could be the primary focus. For example, the Pareto principle (i.e., 80-20 rule) is often observed in the interaction of a company's customer base and product portfolio; that is, 80% of profit comes from 20% of customers and 20% of products. When measuring the service provided to key customers, a combination of OFR and CFR is most appropriate; while in assessing the availability of key products, IFR is more applicable. Finally, profit is another critical measure of a company's health. An understanding of the implications of one performance measure (e.g., OFR) on the others (e.g., CFR, IFR, and, ultimately, profit) is necessary for companies to promptly identify any potential issues that may hinder the company's long-term growth and guide efforts to enhance competitiveness.

The remainder of this chapter is organized as follows. The next section provides a review of inventory literature involving order-based performance measures and identifies the gaps that motivate this study. Section 3 presents the formulation of the optimization model with underlying assumptions. A numerical example is also provided. In Section 4, comprehensive computational experiments are performed and the results are discussed. Finally, Section 5 summarizes the key findings of the work, the contributions to the existing inventory literature, managerial implications, and limitations and future research opportunities.

## 2 Literature Review

The literature on order-based measures can be characterized by its focus on assemble-to-order (ATO) versus a finished goods distribution system, single product versus multi-product, optimality-related issues versus approximate evaluation, discrete-time versus continuous-time, and its assumptions about the underlying demand process.

Most research concerning order fulfillment performance measures is related to ATO systems with either a single end product or multiple end products. The system is typically modeled as a queue or a set of queues with a compound Poisson demand process. An end product consists of a fixed set of components and each component is controlled by an independent base-stock policy. Demands are satisfied on a first-come-first-served basis. Song and Yao (2002), Lu et al. (2003) and Hoen et al. (2011) are some of the works using this approach. Song and Yao (2002) study a single product assembly system with focus on backorder minimization under a limited inventory capital and cost minimization under a service constraint, respectively. They develop lower and upper bounds for the performance measures and examine the tradeoff between inventory holding cost and order fill rate. Lu et al. (2003) extend the research to a multi-product ATO system with independent and identically distributed replenishment lead times for components. They derive a joint queue-length distribution of anticipated orders, from which approximations and bounds are developed to estimate the response-time-based OFR. Both works assume that components are backordered if they are not readily available and the end-products are not assembled and shipped until all components required are ready. Hoen et al. (2011) present an approximate evaluation of the OFR in lost-sales systems using a weighted average of two estimates. One estimate tends to underestimate the OFR by assuming highly coupled demands for components, while the other overestimates the OFR by completely ignoring the dependency among components. The structure of their approximation is similar to the approximation of the waiting time in a G/G/1 queue. They assume deterministic lead times for components and demonstrate that the OFR is insensitive to the distribution of the component lead times.

Customer orders of finished goods in the distribution system setting that this research addresses are similar to the final products in an ATO system in the sense that they both consist of a set of SKUs and a customer order type could be considered as a type of final product. But the two differ in two major aspects. First, a final assembled product has a fixed bill of materials (i.e., a fixed set of components with fixed quantity), whereas a customer order in a distribution system may have a fixed set of SKUs but the demand for each SKU typically varies over time following a probability distribution. Second, the components required in an assembled product are closely related and all have to be available before a product can be assembled and shipped. That is comparable to the situation where a customer order is either filled completely or rejected as whole. The traditional measure of order fulfillment, OFR, thus works well in the ATO system. In the distribution system however, the products requested by a customer in a typical order are not closely tied to each other and partial fulfillment is common in the business-to-business environment. A binary measure of order fulfillment (OFR) is not sufficient to gain insight on the service that individual customers receive.

Song (1998) explores the cases of single-unit demand of all components in an end product, multiple-units demand of a fixed kit of items in an assembly system, and random demand of a fixed set of products in a distribution system, under a continuous-review base-stock inventory policy. The author assumes a continuous-time compound Poisson demand process with backlogging and deterministic lead times. Heuristic bounds are developed to estimate the OFR using item-based information, but the effectiveness of bounds varies by the structure of the order and the size of the order relative to the other order types. Each order type is considered independent of the other order types, which is also the assumption of this research, but the demand process is modeled differently and this study assumes discrete time. Hausman et al. (1998) discuss a multi-item inventory system controlled by a periodic-review, order-up-to policy with identical review periods for all items. The demand for an item is treated as the sum of all customer orders for that item, and then the aggregated item demands are assumed to follow a multivariate normal distribution. The item demands are correlated in a period, but are independent across time periods. Optimization models are developed to maximize the probability of filling all demands in a period within a pre-specified time limit subject to an inventory capital constraint. But since they use the aggregated demand for an item, their research only provides bounds on the fulfillment probability of individual orders.

Anupindi and Tayur (1998) compare item-based performance measures with order-based performance measures including order response times and order fill rate in given lead times for a multiproduct cyclic production system. They demonstrate that the item fill rate is not a good indicator for the order fill rate. A simulation procedure is presented to obtain base-stock policies for different performance measures. Closs et al. (2010) use simulation modeling to examine how item and order fill rates behave differently under different settings of configuration capacity, inventory level and product complexity. They find that a relatively high item fill rate does not necessarily lead to a high order fill rate unless the item fill rate is over 99%. Shao and Dong (2012) develop optimization models to maximize order fill rate subject to an inventory capital constraint, and compare the order fulfillment performance of a made-to-order system with a made-to-stock system. Their research focuses on a single final product comprised of multiple-components, which simplifies the OFR model considerably in the made-to-stock system. Larsen and Thorstenson (2014) investigate how the order-based and item-based performance measures are related in magnitude for a single-item, single-stage system with backordering and constant lead times controlled by a continuous review, base-stock policy. This is the only research that has been found to include the measure of average fill rate across orders or average customer-order fill rate (average CFR). But their work is based on a single-item system, and their observation about the relationship between CFR and IFR is not evidenced in this study.

It appears that in the existing literature none of the research has so far studied multi-product, finished goods distribution systems with the objective to maximize order fulfillment performance under a limited inventory capital subject to a profit constraint. The research that is most related to the work in this chapter is Zinn et al. (2002). The authors study the product availability level experienced by individual customers and propose four order-based measures. One of them, the probability that a particular customer will not face a stockout with the next item purchased, has some similarity with the CFR because they both focus on the item availability in an order. Another measure, the probability that a particular customer will not face a stockout with the next order placed, is similar to the OFR as they all focus on the order as a whole. The focus of their research is on measuring the probability of no stockout while the partial availability of products is ignored. In contrast, this study evaluates the percentage of fulfillment in an order and the percentage of all the orders completely filled, which is in a sense similar to the difference between cycle service level and unit fill rate in the inventory management literature. The authors assume that each order profile contains a fixed set of items and the demand for an item follows a normal distribution, which corresponds to the setting of this study. As the authors have observed, in the business-to-business environment it is common that retailers or industrial customers repeatedly purchase the same variety of products from manufacturers or distributors.

The modeling approach in Zinn et al. (2002) is very different from the models proposed in this study. Their optimization model minimizes the safety stock investment subject to a service constraint. It has a nonlinear objective function which the authors linearize through a piecewise linear approximation. Since their model is designed to execute separately for each customer order and the total safety stock required for an item is calculated afterwards as a function of the joint standard deviation of the lead time demands, they cannot address the inventory capital constraint which a company typically faces nor can they investigate the tradeoff between the overall service level and the cost. In contrast, the mixed integer linear programming (MILP) models developed in this study allocate the limited inventory capital to maximize the order fulfillment performance while meeting the minimum profit requirement, in response to the fact that companies constantly need to balance profit, customer satisfaction, and investment. Table 2.1 summarizes the characteristics of this study relative to that of Zinn et al. (2002) and Larsen and Thorstenson (2014).

*Table 2.1: Comparison of different research works regarding order fulfillment measures*

| | Larsen & Thorstenson (2014) | Zinn et al. (2002) | This Research |
|---|---|---|---|
| Finished Goods Distritubution System | Yes | Yes | Yes |
| Multiple Products | - | Yes | Yes |
| Both CFR and OFR | Yes | Partially | Yes |
| Optimization | - | Yes | Yes |
| Linear Model | - | - | Yes |
| Inventory Capital Limit | - | - | Yes |
| Profit Expectation | - | - | Yes |

# 3 Inventory Classification Models with Order-based Performance Measures

Optimization models are developed to address the problem of determining inventory levels of individual SKUs such that order fulfillment can be maximized subject to inventory capital and profit constraints. This section starts with a formal description of the addressed optimization problem and key model assumptions, followed by the MILP formulations and a numerical example.

## 3.1 Problem Description and Model Assumptions

Let $N$ be the set of SKUs in the inventory, and $K$ represent the set of order types. The set of items required by order type $k$ is denoted by $N^k$. The demand for SKU $i$ in order $k$ is a random variable assumed to follow a normal probability distribution with a forecast mean demand $d_{ik}$ and standard deviation $\sigma_{ik}$. Since business customers tend to repurchase the same items over regular intervals, the demand pattern of an item requested in a particular order can be estimated from the customer's purchasing history. The unit profit $\pi_{ik}$ is order-SKU based, given that the company may quote different customers different prices based on the order quantity or the strategic position of the customer. The unit cost of an item, denoted as $c_i$, includes material cost and inbound shipping cost. Inventory holding cost $h_i$ is an SKU-based parameter, consisting of general factors such as capital cost and warehousing cost (e.g., rent, equipment, labor and utilities costs), and also SKU-specific factors such as shelf-life. All else being equal, the shorter the shelf-life, the higher the inventory holding cost percentage a product has. Inventory is managed under a periodic-review, order-up-to inventory policy with a common review period of one time unit.

The fulfillment performance of a customer order is affected by the allocated safety stock levels of individual SKUs in the order, which is a function of the item cycle service level (CSL). To overcome the complexity of nonlinear relationships among SKU cycle service level, safety stock level, and order fulfillment, CSL is discretized into a finite number of inventory classes (Millstein et al., 2014), denoted by $J$. Each inventory class $j$ has a corresponding service level ($\alpha_j$), a z-value ($z_j$) to determine safety stock, and a value from the standard loss function ($e_j$) to calculate the expected lost sales. $\alpha_j$ ranges from 1% to 99.99%, with an increment of 1% between 1% and 99%, an increment of .1% between 99.1% and 99.9%, and a highest level of 99.99% (109 levels in total). This structure allows the model to be constructed as an MILP model, to which optimal solutions can be found and optimality can be assured through the branch-and-bound procedure in MILP. An SKU in an order can only be assigned to at most one inventory class. The objective is to maximize the order fulfillment (i.e., either OFR or average CFR) while ensuring the total profit is no less than the required threshold $p$ through optimizing the assignment decision of SKUs to inventory classes (i.e., CSL).

The proposed model takes into account the limited inventory capital $\omega$ which sets the maximum dollar value of the inventory that can be carried during any given time period. The inclusion of an inventory capital constraint is for the practical application of the model because the amount that a company can borrow is typically limited and in most cases is in proportion to its own capital. The inventory capital is the decision of the company's management, and it also implicitly considers the storage space limit, if any, in relevant facilities.

The fulfillment performance measure for a customer, CFR, is the number of units filled from on-hand inventory as a fraction of total units required in an order. An order is considered to be completely filled if the expected CFR is at 99.99% or higher. The percentage of the customer orders that are satisfied in full immediately from inventory provides another performance measure, OFR. The following summarizes the additional assumptions of the models.

First, it is assumed that a customer purchases the same mix of products continually, reflecting the purchasing practice common in the business-to-business environment.

Second, orders are considered independent of each other, and an order type is assumed to occur once per time unit, given that manufacturers/distributors typically replenish the warehouses of downstream customers once per period (e.g. weekly) in order to take advantage of the economies of scale in transportation. In the cases where multiple orders occur regularly within a time period from the same customer, they can be treated as different orders in the model.

Third, if an SKU required by an order is in shortage, partially filling the order is acceptable and unmet demands are treated as lost sales. The model assumes no backordering.

Fourth, replenishment lead times of SKUs are assumed to be one time unit. Although the order-up-to inventory level is commonly calculated as the sum of the expected demand and the safety stock over the lead time, in reality a company never needs to hold an inventory level beyond its review cycle. When inventory is reviewed and replenished every time unit, the maximum inventory level on-hand is the sum of the regular stock and the safety stock of a period, regardless of lead times. In the case where demand is stationary, what the lead time affects is when to place an order rather than how much to order.

### 3.2 MILP Formulation

This section first presents the model formulation for maximizing the CFR, and then the formulation for the OFR maximization.

*CFR maximization*

*Sets*

$N$: set of inventory items (SKUs)

$K$: set of orders in a time period (during lead time)

$J$: set of possible inventory classes

$N^k$: set of SKUs in order $k$, $\forall k \in K$

*Parameters*

$d_{ik}$: expected demand of SKU $i$ in order $k$, $\forall i \in N, k \in K$.

$\sigma_{ik}$: standard deviation of the forecast demand $d_{ik}$, $\forall i \in N, k \in K$.

$\pi_{ik}$: unit gross profit of SKU $i$ in order $k$, $\forall i \in N, k \in K$.

$c_i$: cost of goods per unit, $\forall i \in N$.

$h_i$: inventory holding cost (% of the cost of SKU $i$), $\forall i \in N$.

$\alpha_j$: cycle service level associated with class $j$, $\forall j \in J$.

$z_j$: z-value associated with cycle service level $\alpha_j$, $\forall j \in J$.

$e_j$: value in standard loss function (corresponding to CSL $\alpha_j$), $\forall j \in J$.

$\omega$: inventory capital (the total inventory value must not exceed the inventory capital).

$p$: minimum profit required (which must be no higher than the maximum profit for a given inventory capital) .

*Decision Variables*

$x_{ikj} = 1$, if and only if SKU $i$ in order $k$ is assigned to inventory class $j$, 0 otherwise. $\forall i \in N^k,\ k \in K,\ \ j \in J$.

$I_i \geq 0$, average on-hand inventory of SKU $i$, $\forall i \in N$.

$F_k \geq 0$, denotes the expected fill rate of customer order $k$, $\forall k \in K$.

$v_i$, is the pooled safety stock factor (pooled $z$-value) for SKU $i$, $\forall i \in N$.

*Objective Function*

$$\text{Max} \quad \textstyle\sum_{k=1}^{K} F_k/|K| \tag{1}$$

*Constraints*

$$\textstyle\sum_{j \in J} x_{ikj} \leq 1 \qquad \forall i \in N^k,\ k \in K \tag{2}$$

$$F_k = [\sum_{i \in N^k} \sum_{j \in J} x_{ikj} \times (d_{ik} - \sigma_{ik} e_j)] / \sum_{i \in N^k} d_{ik} \qquad \forall k \in K \qquad (3)$$

$$v_i = (\sum_{k \in K} \sum_{j \in J} x_{ikj} z_j \times \sigma_{ik}) / \sum_{k \in K} \sigma_{ik} \qquad \forall i \in N \qquad (4)$$

$$\sum_{i \in N} (\sum_{k \in K} \sum_{j \in J} x_{ikj} d_{ik} + v_i \sqrt{\sum_{k \in K} \sigma_{ik}^2}) c_i \leq \omega \qquad (5)$$

$$\sum_{i \in N} [\sum_{k \in K} \sum_{j \in J} x_{ikj} (d_{ik} - \sigma_{ik} e_j) \pi_{ik} - I_i c_i h_i] \geq p \qquad (6)$$

$$\sum_{j \in J} x_{ikj} (d_{ik} + \sigma_{ik} z_j) \geq 0 \qquad \forall i \in N^k,\ k \in K \qquad (7)$$

$$I_i \geq \sum_{k \in K} \sum_{j \in J} x_{ikj} d_{ik} / 2 + v_i \sqrt{\sum_{k \in K} \sigma_{ik}^2} \qquad \forall i \in N \qquad (8)$$

$$I_i \geq (\sum_{k \in K} \sum_{j \in J} x_{ikj} d_{ik} + v_i \sqrt{\sum_{k' \in K} \sigma_{ik'}^2}) / 2 \qquad \forall i \in N \qquad (9)$$

$$I_i \geq 0 \qquad \forall i \in N \qquad (10)$$

$$x_{ikj} \in \{0, 1\} \qquad \forall i \in N^k,\ k \in K,\ j \in J \qquad (11)$$

$$v_i \in \mathbb{R} \qquad \forall i \in N \qquad (12)$$

$$F_k \geq 0 \qquad \forall i \in N \qquad (13)$$

$$F_k \leq 1 \qquad \forall i \in N \qquad (14)$$

The objective function (1) maximizes the average fill rate of $|K|$ customer orders (i.e., average CFR), which is derived from the expected satisfied demand of each order as defined in Constraint (3).

Constraint (2) restricts that an SKU in an order cannot be assigned into more than one inventory class. If an SKU in an order is not assigned with any CSL (i.e., $\sum_{j=1}^{J} x_{ikj} = 0$), it implies that there is no inventory to supply the demand of this SKU required by this particular order.

Constraint (3) computes $F_k$, the percentage of total units required in order $k$ that can be filled immediately from the available inventory. $\sigma_{ik}e_j$ is the expected lost sales in order $k$, which is determined by the inventory class to which SKU $i$ in order $k$ is assigned.

Constraint (4) calculates $v_i$ as the average of z-values across all the orders, weighted by the standard deviation of demand of SKU $i$ in each order. This definition corresponds to the pooled $z$-value defined by Zinn et al. (2002), which determines the SKU-level safety stock and may be negative.

Constraint (5) ensures that the total capital required for holding regular and safety stock does not exceed the agreed inventory investment. The first term in the left-hand side of the constraint calculates the regular stock, while the second term computes the total safety stock required for item $i$.

Constraint (6) makes sure that the net profit meets the minimum profit requirement. The first term in the left-hand side of the equation is the expected gross profit, while the second term is the inventory holding cost. The subtraction of the two gives the expected net profit.

Constraint (7) ensures that the anticipated inventory level stays nonnegative for any SKU in any order. When the required CSL is less than 50% because of either the high inventory holding cost or the low profit margin, or both, a negative safety stock occurs. And if the demand is also highly volatile (i.e., the coefficient of variation is high), the net inventory level could be negative, which obviously does not make any practical sense. Constraint (7) makes sure that if that happens, $\sum_{j=1}^{J} x_{ikj}$ is set as zero, indicating that no inventory is kept for the particular SKU required in an order.

Constraints (8) and (9) determine the average on-hand inventory level in order to compute inventory holding cost. Positive safety stock activates Constraint (8), whereas negative safety stock makes Constraint (9) effective. If there is no safety stock required, Constraints (8) and (9) converge. Constraint (10) enforces that the joint inventory level for any SKU is nonnegative.

*OFR Maximization*

For OFR maximization, in addition to the aforementioned decision variables, the following binary variable is defined:

$y_k = 1$, if and only if order $k$ is completely filled; that is, at least 99.99% of the units required in order $k$ are satisfied immediately from on-hand inventory, $\forall k \in K$.

The objective function is:

$$\text{Max} \quad \sum_{k \in K} y_k \, / |K| \tag{15}$$

subject to:

constraints (2) to (14)

$$y_k - 1 \leq F_k - 0.9999 \qquad \forall k \in K \tag{16}$$

$$y_k \in \{0, 1\} \qquad \forall k \in K \tag{17}$$

The objective function (15) maximizes the OFR; i.e., the percentage of customer orders that are expected to be filled completely from available inventory. Constraint (16) ensures that when the expected fill rate for customer order $k$ is less than 99.99%, $y_k$ must be 0; when the fill rate is equal to or higher than 99.99%, the maximization process of the objective function will set $y_k$ to be 1.

This model formulation not only provides the companies with the maximum OFR that they can possibly achieve given all the constraints, but also reveals the expected service level of individual orders – which orders are expected to be fully satisfied, and what the service level is if an order is partially filled - as presented in the next section. This has rarely been provided, if at all, in the existing literature.

As discussed in Chapter 1, the MILP model is NP-hard and there is no known polynomial algorithm to solve it to optimality. The proof of NP-hardness is established by transforming the formulation into an uncapacitated facility location problem (UFLP, cf. Drezner and Hamacher 2004).

## 3.3 Numerical Example

To illustrate the order fulfillment problem and the implementation of the optimization models, consider a distributor that carries 5 different styles of washers sold to customer A, B, and C respectively. Customer A purchases 3 styles of washers regularly: 3.5 cubic feet Top Load, Duet 4.2 cubic feet Front Load, and Duet 4.5 cubic feet Front Load. Customer B also purchases 3 models: 4.3 cubic feet Top Load, Duet 4.2 cubic feet Front Load, and Duet 4.3 cubic feet Front Load. Customer C purchases both styles of top load washers and also two front load washers: Duet 4.5 cubic feet Front Load, and Duet 4.3 cubic feet Front Load. Table 2.2 lists the unit cost of each product, the expected demand for each product from each customer, standard deviation of the demand, and unit prices quoted to individual customers. The inventory holding cost percentage is 35%, the same for all washers. The distributor has a limited inventory capital of $150,000 for washers, and the minimum profit expectation is $12,000.

*Table 2.2: Basic parameters of numerical example*

| SKU# | Product Description (Washer) | Unit Cost | Order A | | | Order B | | | Order C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Demand | Std. Dev. | Unit Price | Demand | Std. Dev. | Unit Price | Demand | Std. Dev. | Unit Price |
| SKU#1001 | 3.5 cu. ft. HE Top Load | $ 237.38 | 15 | 9 | $282.38 | | | | 28 | 21 | $ 277.38 |
| SKU#1002 | 4.3 cu. ft. HE Top Load | $ 290.46 | | | | 19 | 9 | $318.46 | 4 | 3 | $ 323.46 |
| SKU#1003 | Duet 4.2 cu. ft. HE Front Load | $ 421.51 | 26 | 6 | $533.51 | 15 | 7 | $538.51 | | | |
| SKU#1004 | Duet 4.5 cu. ft. HE Front Load | $ 486.61 | 31 | 23 | $606.61 | | | | 6 | 4 | $ 621.61 |
| SKU#1005 | Duet 4.3 cu. ft. HE Front Load | $ 465.26 | | | | 24 | 13 | $567.26 | 16 | 12 | $ 567.26 |
| Inventory budget = $150,000; Minimum profit = $12,000 | | | | | | | | | | | |

The distributor may choose to maximize either the number of completely filled orders (i.e., OFR) or the average fill rate of customer orders (i.e., CFR). Table 2.3, 2.4 and Figure 2.1 demonstrate the performance and impacts of these two approaches from various perspectives. When maximizing the OFR, customer B and C are expected to enjoy near perfect order fulfilment, but at the cost of customer A which is given a much lower service level of 96.19%, as shown in Figure 2.1. The average fill rate across all orders is 98.73%. If customer B and C are the most important customers to the firm relative to customer A, this approach may be sound for the company.

When maximizing the average CFR, the three customers are given similar levels of fill rate at 99.86%, 99.92% and 99.90% respectively, with an average of 99.89% and none receiving perfect fulfillment. Maximizing the average CFR results in slightly higher inventory holding cost, but the increase is justified by the additional revenue and ultimately the higher profit. If due to the competition, the minimum service level has to be 99.85% in order to retain customers, the CFR approach will help the company to sustain its customer base for long term growth. Both approaches exceed the required minimum profit.

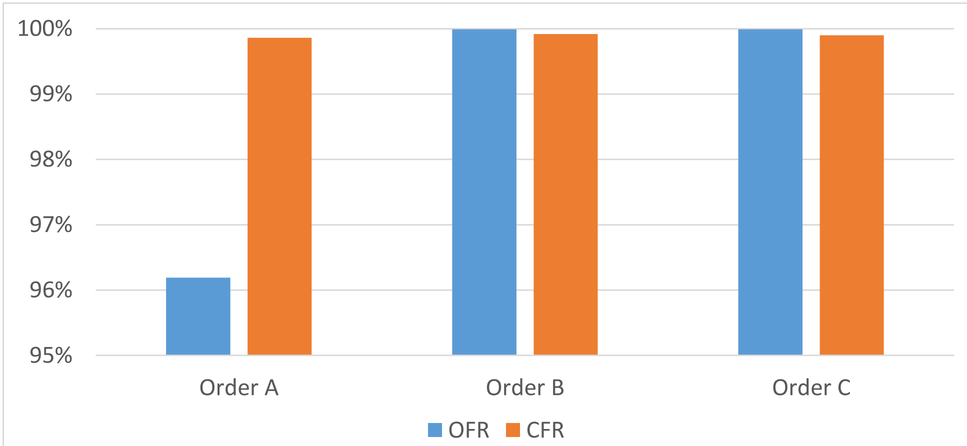
*Table 2.3: Performance comparison: OFR maximization vs. average CFR maximization*

| | **OFR** | **CFR** |
|---|---|---|
| # of orders completely filled | 2 | 0 |
| Average fill rate across orders | 98.73% | 99.89% |
| Inventory holding cost | 3,332 | 3,343 |
| Profit | 12,119 | 12,374 |

*Table 2.4: SKU Fill Rate in Order A, B and C: OFR maximization vs. average CFR maximization*

| | | **OFR** | | | **CFR** | | |
|---|---|---|---|---|---|---|---|
| | | Order A | Order B | Order C | Order A | Order B | Order C |
| **Item FR** | SKU#1001 | 96.37% | | 100.00% | 99.92% | | 99.93% |
| | SKU#1002 | | 100.00% | 100.00% | | 99.94% | 99.90% |
| | SKU#1003 | 97.73% | 99.97% | | 99.96% | 99.92% | |
| | SKU#1004 | 94.81% | | 100.00% | 99.75% | | 99.85% |
| | SKU#1005 | | 100.00% | 99.98% | | 99.89% | 99.85% |

*Figure 2.1: Fill rate of individual orders: OFR maximization vs. average CFR maximization*



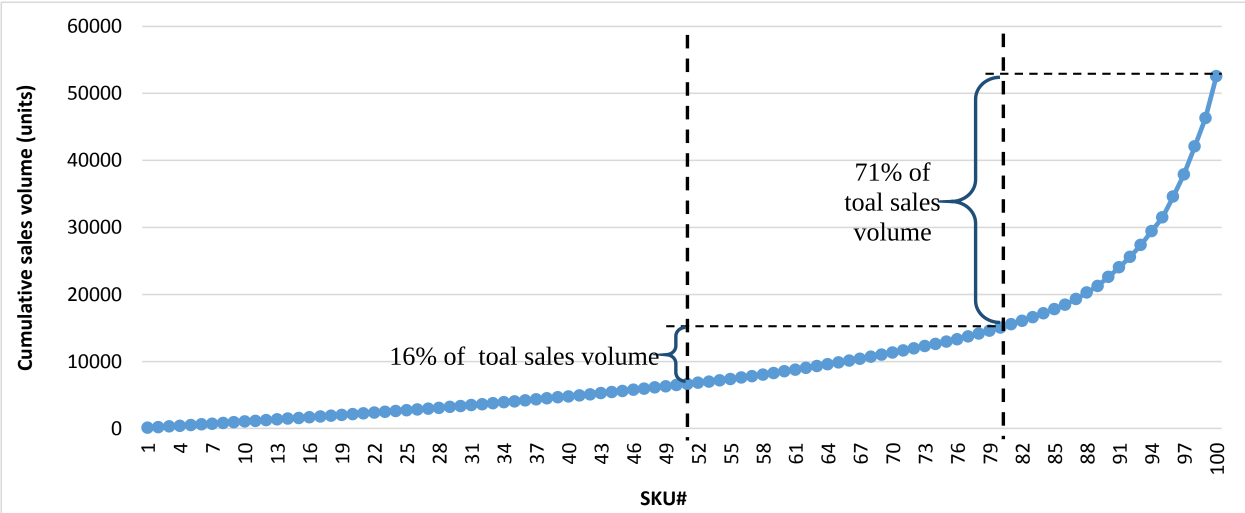

## 4 Computational Experiments

Computational experiments are performed to examine the behavior and output of the MILP models under different parameter settings. Specifically, the goals of the experiments are: a) to identify approximate Pareto efficiency frontiers (PEFs) for describing the tradeoffs between profit and order fulfillment measures; b) to gain insights on the impacts of different order fulfillment measures; c) to understand the effect of the inventory capital; and d) to compare and contrast order fulfillment measures with the commonly used item fill rate measure. Both optimization models are solved with the MILP branch-and-cut method using GAMS/CPLEX 12.4 on a PC that has an 8GB RAM and runs on Intel Core i7-4500 CPU with a maximum frequency of 2.4GHz. The relative optimality tolerance is set at $10^{-12}$ for the CFR maximization and $10^{-4}$ for the OFR maximization so that the models can be solved in a reasonable time, while generating sufficiently good results.

## 4.1 Experimental Design

The dataset used to conduct the computational study has 100 SKUs purchased by 20 customers (i.e., orders), which is large enough to capture the relationships between various performance measures and the impacts of the inventory capital, while at the same time permitting the models to be solved quickly. The total mean demand for each SKU across all the orders is generated using the Pareto distribution (Arnold, 2008) to best reflect what a company typically experiences in the real world; that is, about 20% of SKUs contribute to roughly 70-80% of the total sales volume, 30% account for 15-20% of the sales volume, and the remaining 50% represent 5-10% of the sales. In the experiments, the minimum threshold of the Pareto distribution is set at 100 and the shape parameter is 2. Figure 2.2 shows the cumulative demand of SKUs, in which SKUs are organized by sales volume (from the smallest to the largest).

*Figure 2.2: Cumulative sales volume of SKUs*



SKUs are randomly assigned to the 20 customer orders by first randomly generating a data point using a uniform distribution of $U(0, 1)$ for every SKU in every order, and then assigning "1" to the SKU in the order if its data value is less than 0.25, and "0" otherwise. The value "1" indicates that the SKU is included in the order, and "0" means that the SKU is not part of the order. In this experiment, the largest order (in terms of the number of SKUs) has 35 SKUs, the smallest has 21, and the average number of SKUs per order is about 27. Table 2.5 summarizes all the controlled parameters. The expected demand for each SKU in an order is randomly generated in a manner such that the total demand for the SKU is equal to the total mean demand generated through the Pareto distribution, and the coefficient of variation of the demand of an SKU in an order is randomly generated following a uniform distribution of $U(0.05, 0.85)$, upon which the standard deviations are derived. Inventory holding cost is set at 35% for all SKUs, albeit the models allow the holding cost percentage to vary by SKUs. The SKU unit cost is assumed to follow a normal distribution of $N(\$300, \$100)$. The SKU profit margin follows a uniform distribution of $U(0.10, 0.30)$ and is assumed to stay consistent for all orders in the experiment to ease the analysis.

*Table 2.5: Summary of controlled parameters*

| **Controlled Parameters** | **Value / Probability Distribution** |
|---|---|
| SKU Cost | Normal ($300, $100) |
| SKU Profit Margin | Uniform (0.10, 0.30) |
| SKU mean demand | Pareto (100, 2) |
| Order-SKU: mean demand | [value of Uniform(0, 1) / sum of all orders] X value of Pareto (100, 2) |
| Order-SKU: CV of mean demand | Uniform (0.05, 0.85) |
| Inventory holding cost | 35% |

The inventory capital varies from $10 million to $29 million, in increments of $1 million. $29 million is the maximum inventory investment required to completely satisfy all the orders from inventory and reach an average CFR of 99.99%. Under each level of inventory capital, 6 levels of profit requirements are examined including the minimum profit when maximizing order fulfillment measures, the maximum profit when performing profit maximization, and another 4 levels of profits equally distributed in between. Experiments are performed on a total of 162 instances. To understand the tradeoff of CFR and profit, the CFR maximization model is executed on 120 instances (20 inventory capital levels X 6 profit levels). To understand the relationship of OFR and item fill rate, the OFR maximization model is executed on 20 different levels of inventory capital. And to understand the tradeoffs of OFR, CFR, and profit, 22 instances are examined (11 levels of OFR under a given inventory capital X 2 different objective functions).

### 4.2 Computational Results

The tradeoffs of profit with CFR and OFR, the impact of the inventory capital, and the relationships among CFR, OFR, and IFR are the focuses of computational experiments.

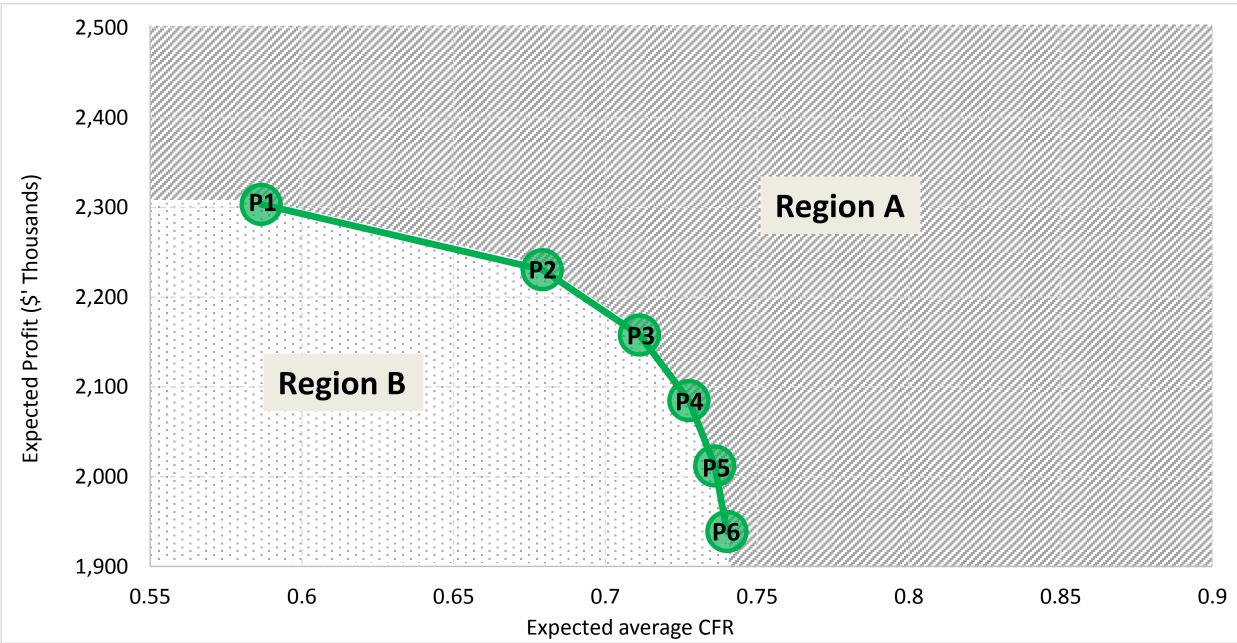
*Tradeoff between Profit and CFR*

Figure 2.3 displays the tradeoff between expected profit and average fill rate across all the orders under an inventory capital of $12 million. The green line approximates a set of choices that are considered Pareto efficient; that is, the optimal tradeoffs between short-term profits and customer services given the inventory capital constraint. If the company’s performance currently falls in region B, they have the opportunity to simultaneously improve both their bottom line and service level without any additional inventory investment. With the current product mix, cost structure and ordering status however, it will not be possible to go beyond the green line to region A without further investment in inventory.

Point P1 on the green line indicates the maximum expected profit that a company can possibly achieve with a $12 million inventory capital, and the minimum expected service level that should be targeted. Point P6, on the other hand, shows the highest expected service level that could be reached and the minimum expected profit that a company should aim for. At point P6, the expected average CFR is 74.0% with individual order expectations ranging from 6.7% to 97.7%. None of the orders are expected to be completely filled when maximizing the average fill rate of customer orders. The intervals between adjacent points have the same amount of expected profit increase/decrease; that is, moving from P1 to P2 or moving from P4 to P5 reduces expected profit by about $73,000, but as shown in Figure 2.3, the improvements in expected service level are drastically different. At the cost of $73,000 profit, the company can improve its expected service level from 58.7% to 67.9% (from P1 to P2), but to further improve to 74.0%, the company has to be prepared to lose four times more in expected profit. While it may not be surprising that the increase of service level demands an economic tradeoff when a company is already operating on an efficient frontier, a key question is whether the tradeoff can be justified in the long run. The aforementioned findings lead to insight 1:

*Insight 1: There exists a tradeoff between the expected profit and the expected average customer-order fill rate for a given inventory capital. And the higher the current expected service level, the less room for improvement and the greater impact on expected profit for every single percentage-point improvement.*

*Figure 2.3: An approximate Pareto efficiency frontier of expected average CFR and expected profit at a $12 million inventory investment*




Plotting the tradeoffs of expected profits and expected average CFRs under various levels of inventory capital results in the graph in Figure 2.4, where the red line on the top represents maximum expected profits that can be achieved under different levels of inventory investments and the blue line at the bottom shows the maximum expected average CFRs for the same levels of inventory capital. Point P1 at the end of the red line shows the highest expected profit that the company can possibly accomplish given the current revenue and cost structure, which requires maintaining an inventory level close to

$22.2 million with an expected average fill rate of 98.4%. More inventory investment will not help improve the expected profit further assuming the company operates under the status quo. In other words, region K at the top is presently out of reach of the company no matter how much more inventory capital is invested or how further the service level is scaled down.

Point P3 on the left end of the red line represents an expected profit of nearly $2 million by keeping $10 million stock (i.e., $10 million inventory capital). If the company can invest another $1 million in inventory, the expected profit has a chance to grow by nearly 8% or over $156,000 (moving up to P5 on the line for $11 million inventory capital). The expected return on this investment is over 15.6%. When the inventory grows to $15 million, adding another $1 million of inventory can only boost the expected profit by about $102,000; and when the inventory level is as high as $21 million, an additional $1 million investment in inventory can only add about $10,000 in expected profit, which means the expected return on investment is merely 1%. Clearly, there are diminishing returns associated with increasing inventory capital with regard to the growth in expected profit, as suggested by Figure 2.6.

The lines connecting the red and blue lines in Figure 2.4 approximate Pareto frontiers under different levels of inventory investment (from $10 million to $22 million). Regions A to K show combinations of expected profit and average CFR that are not Pareto optimal for given inventory budgets. As the available inventory capital increases, the tradeoff between expected profit and average CFR becomes smaller, which implies that there is little room for further improvement on either customer service or profit if the company already operates on the Pareto frontier. However, there is a higher chance that the company may not operate optimally at present, because if the company is holding $19 million of inventory, for example, then operating anywhere in the region A to J means a Pareto improvement is possible.

When the inventory capital is limited (e.g., $10 million), focusing on the average fill rate could result in 21% less profit in comparison to a more profit-driven strategy (P3 vs. P4). At the capital of $23 million, profit maximization and the average fill rate maximization produce similar results: the difference in expected profit is less than 0.34% and the expected service level is 98.39% versus 99.23% (P1 vs. P2). Moving beyond the expected service level of 99.23% (point P2) on the blue line requires further inventory investment, and the inventory level becomes so high that the additional sales cannot compensate for the increased cost, eventually causing a reduced profit, as shown in the enlarged chart in Figure 2.5.

To the left of point P2 in Figure 2.4, as the average fill rate grows from 66.11% (under a $10 million inventory investment) to 99.23% (under a $23 million inventory investment), the expected profit grows accordingly. A faster growth close to $150,000 on average is observed for every $1 million increase in inventory capital before it reaches $20 million, after which the increases in both expected profit and expected service level slow down significantly for the same amount of inventory investment. In summary, Figure 2.4 reveals the following insights:

*Insight 2: Expected profit increases with the inventory investment until it reaches its maximum value; and there is a diminishing return on the inventory investment for* *profit improvement.*

*Insight 3: The tradeoff between the expected average CFR and expected profit becomes smaller as more inventory capital becomes available.*

*Insight 4: Along with the increase of the inventory capital, the expected average CFR and expected profit exhibit a curvilinear, convex upward relationship with a threshold value below which the average CFR and profit are positively related and above which the two are negatively related.*

*Insight 5: The expected average CFR increases with the inventory capital. A much higher investment is required for making any further improvement on service level as the average CFR becomes higher.*

*Figure 2.4: Approximate Pareto efficiency frontiers of expected average CFR and expected profit under various inventory capital investments*

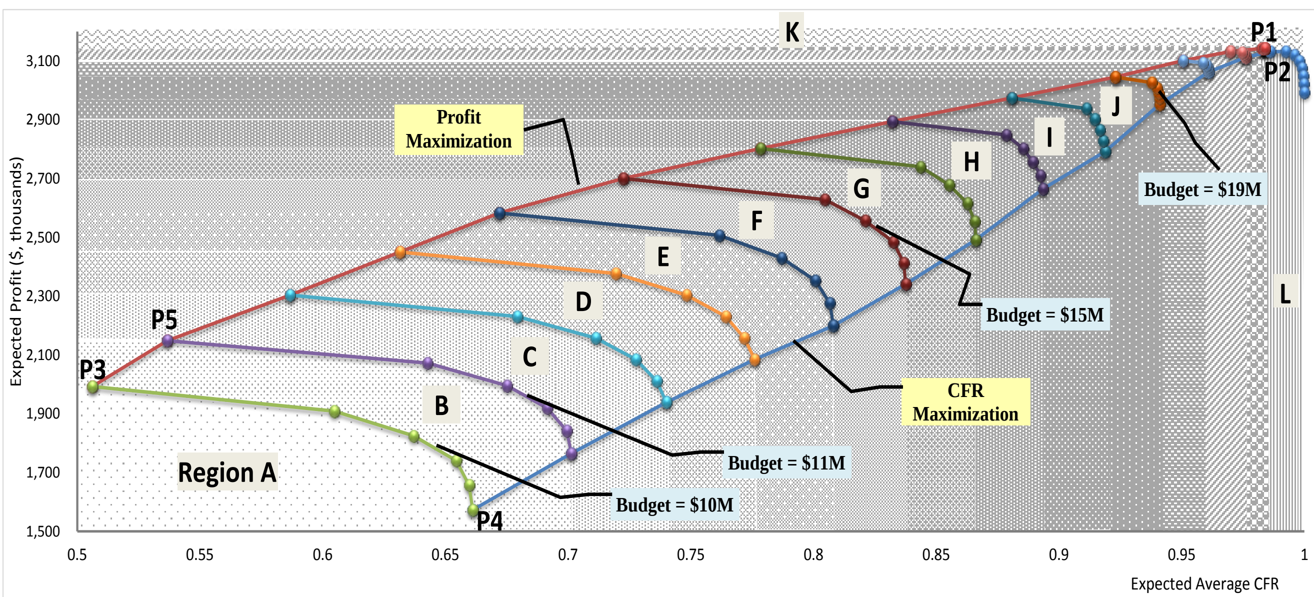


*Figure 2.5: Expected average CFR and expected profit when maximizing CFR under different inventory capital investments*

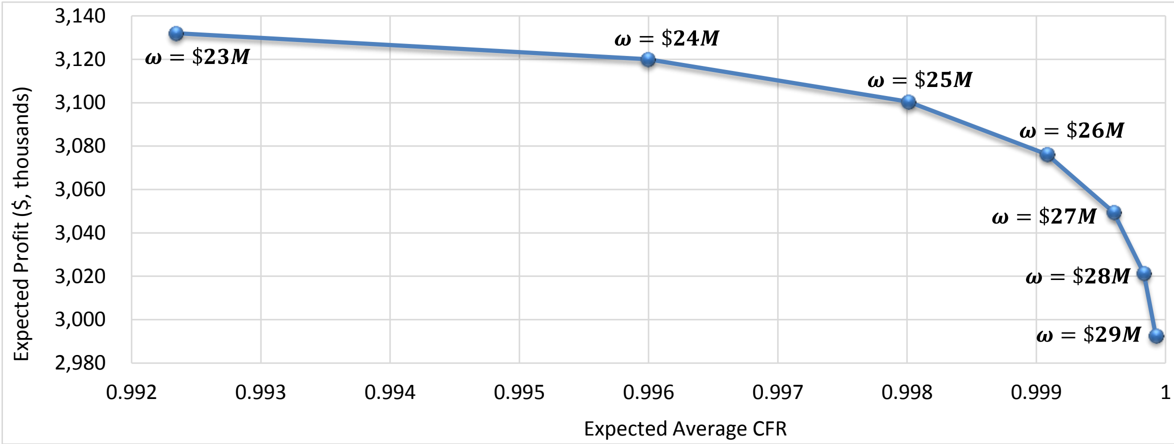
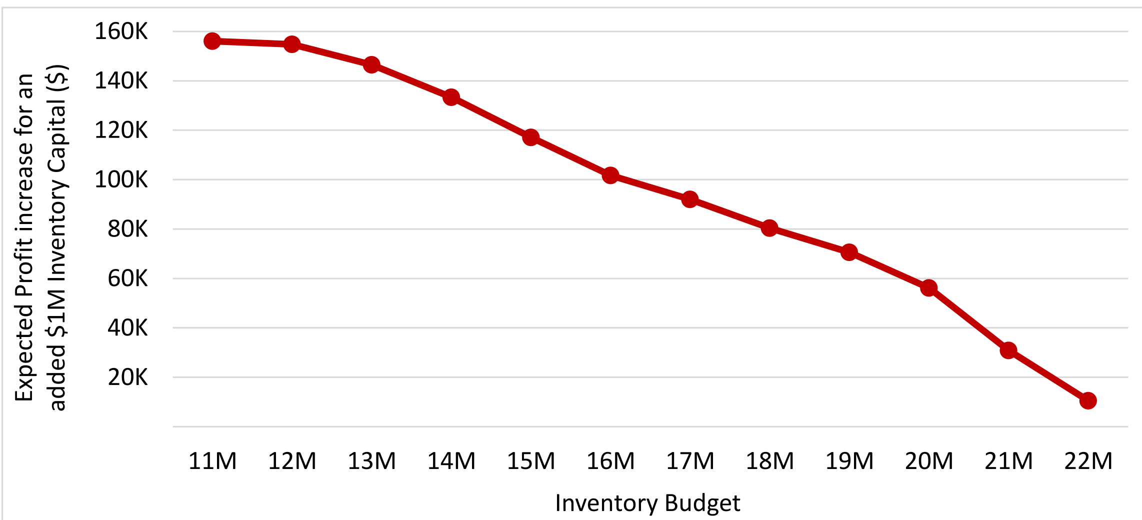



*Figure 2.6: Expected profit increase resulting from an added $1M inventory capital*

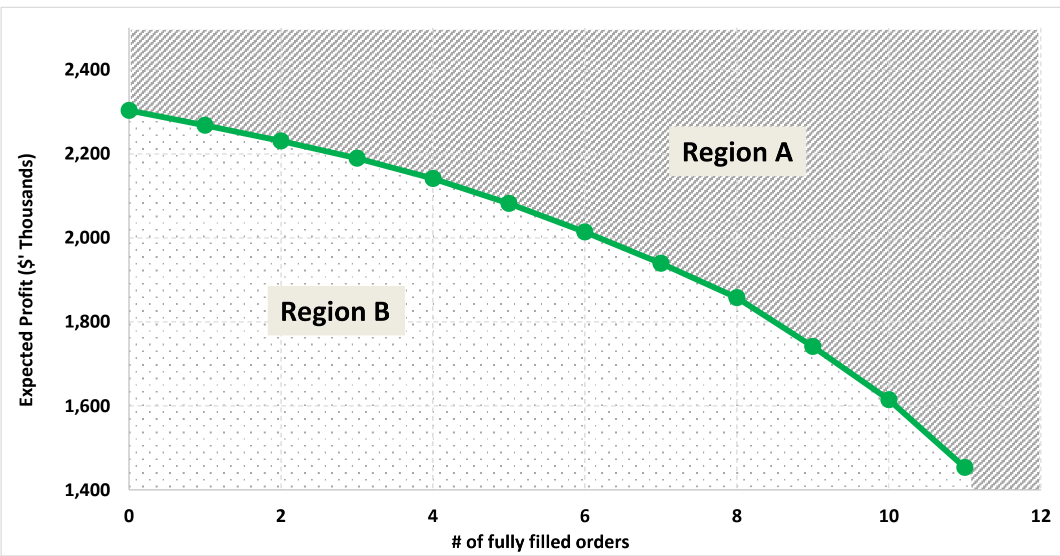



*Tradeoffs of Profit, OFR, and average CFR*

With an inventory investment of $12 million, a maximum of 11 orders (out of 20) are expected to be fully filled, generating over $1.4 million expected profit; while if there is no requirement on the order fill rate, expected profit can be 58.5% higher, reaching $2.3 million. Figure 2.7 shows the tradeoff between OFR and profit. If all customers are equally important to the company, a performance residing on the green line is most desired. Region A is beyond the reach with the existing investment, while region B may imply an inferior inventory performance unless certain customers must take higher priority for perfect order delivery. For a company that has already operated on the green line, the financial tradeoff (i.e. profit) grows larger with the OFR. For example, increasing the number of fully filled orders from 0 to 1 reduces profit by less than \$36,000, while growing from 10 perfect orders to 11 costs 4.5 times more or over \$160,000. This suggests insight 6:

*Insight 6: Under a given inventory capital, expected profit decreases as the expected OFR grows, and the economic tradeoff becomes larger as the expected OFR gets higher.*

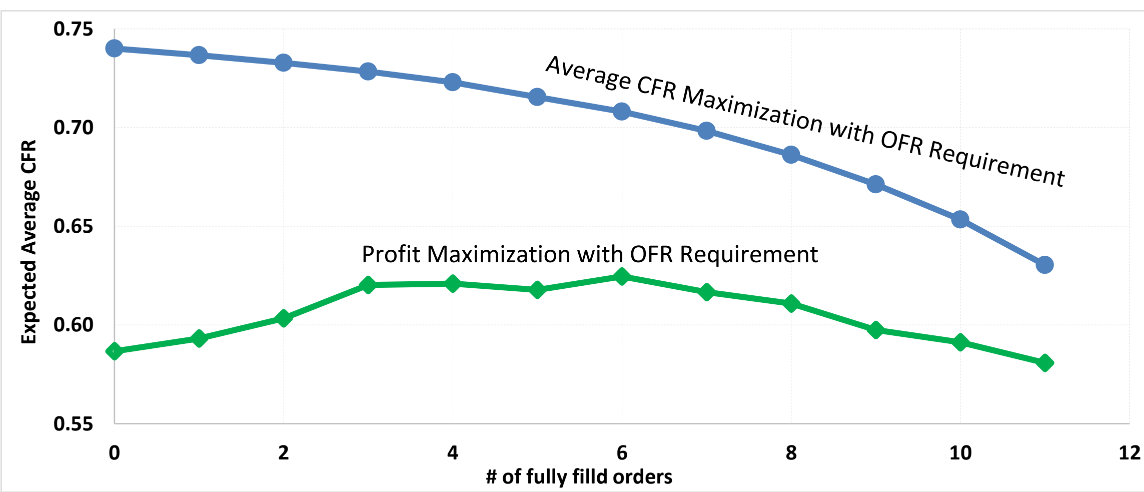
*Figure 2.7: Tradeoff of expected OFR and profit at a \$12 million inventory investment*




When maximizing expected profit subject to a minimum OFR requirement, the expected average CFR fluctuates between 58.1% and 62.5% as evidenced by the green line in Figure 2.8. If taking an alternative approach of maximizing the expected average CFR constrained by the minimum OFR, the expected average fill rate across orders appears to be much higher as presented by the blue line, but at the expense of a lower profit range of $1.3 million to $1.9 million (versus $1.45 million to $2.3 million as shown in Figure 2.7.). From the blue line in Figure 2.8, as the expected number of fully filled orders increases, the expected average CFR declines. The company could achieve an expected average CFR of 74% when there is no requirement on the number of fully filled orders, but could only reach an expected average CFR of 63% when 11 orders are expected to be perfectly filled. And as the expected OFR grows higher, a faster decline is expected in the average CFR. This provides insight 7:

*Insight 7: For the given inventory capital, the expected average CFR decreases as the expected OFR increases, and the decline of the expected average CFR is accelerated as the expected OFR gets higher.*

*Figure 2.8: Expected OFR vs. expected average CFR under different approaches for an inventory investment of $12 million*



*Measure of OFR vs. measure of average CFR*

Maximizing CFR and OFR have different implications for expected profit. The green line in Figure 2.9 provides a benchmark of the maximum expected profits that a company can generate under different levels of inventory capital. The highest profit level, $3.14 million, is achieved at a capital of $23 million, beyond which (point P1 in Figure 2.9) there is no further profit improvement. Maximizing OFR generally yields much less expected profit than the CFR approach. When the capital is less than $22 million, a CFR-focused strategy generates 20% to 33% more profit than an OFR-focused strategy. The two approaches start to converge only when the inventory capital is large enough ($28 million in this experiment) to support perfect fulfillment for almost all orders.

The data values next to the blue line in Figure 2.9 represent the maximum number of orders that are expected to be fully satisfied under each inventory capital. While the overall relationship between the expected OFR and expected profit presents an upward trend, an increase in the expected OFR resulting from the increased inventory investment may not always lead to a higher expected profit. For example, when increasing the capital from $19 million to $20 million, the company can expect to have one more order completely served, but the expected profit suffers a marginal decrease of 1%, which is probably because the resources are moved away from more profitable but unlikely to be fully filled order to maximize the overall OFR. However, if the company can secure another $1 million investment, a profit increase of 8.6% can be expected assuming the cost of capital is consistent. Another example is that when the OFR reaches 95% (19 out of 20 orders are completely filled), further improvement in the OFR also leads to reduced profit, but this is likely for a different reason of the excessive inventory required to support a perfect order fill rate. From Figure 2.9, insights 8 and 9 are derived:

*Insight 8: With a limited inventory capital, a strategy focusing on the number of perfect orders is less profitable than focusing on the average fill rate across orders. The two approaches converge only when the inventory capital is large enough to enable almost all the orders to be fully filled.*

*Insight 9: Profit and the number of perfect orders do not always go in the same direction as the inventory capital increases. Impact of additional investment depends on where the company stands presently.*

*Figure 2.9: Effect of different approaches on expected profit*

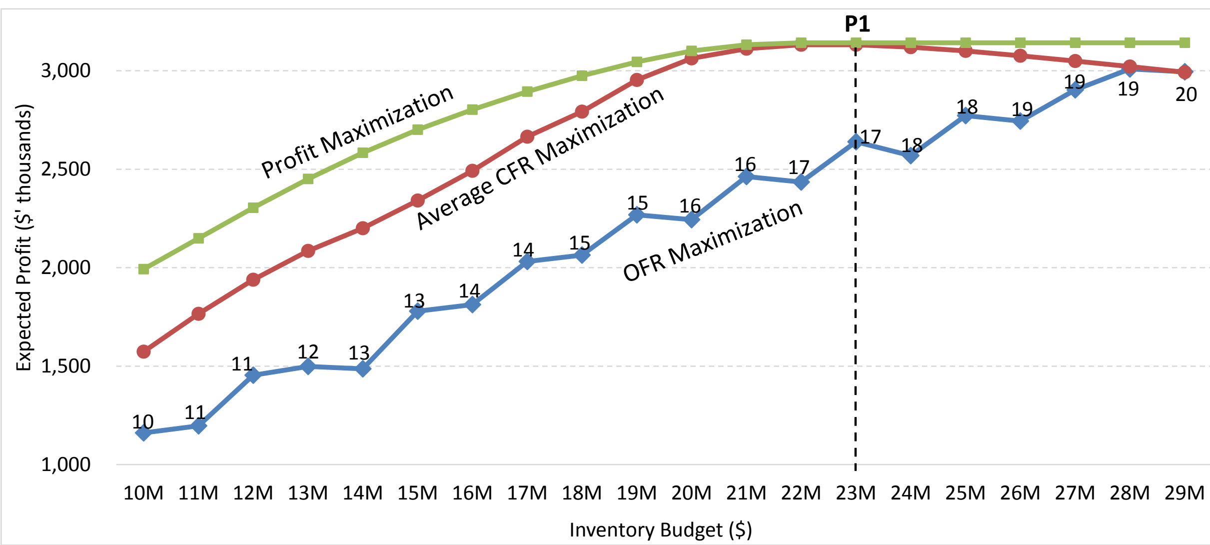
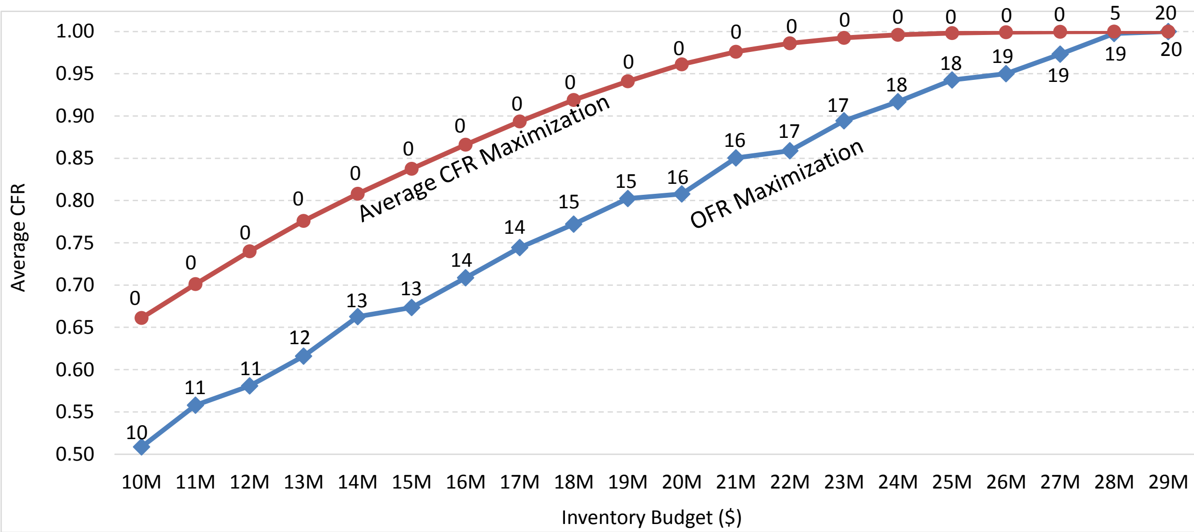



Figure 2.10 displays the effects of two approaches on customer service. The data values next to the red line represent the number of fully filled orders under different levels of inventory capital when maximizing the expected average CFR, which are significantly different from the approach of maximizing the expected OFR. When the

focus of the company is on the average fill rate of customer orders, resources are more evenly distributed and none of the orders are suggested to be perfectly filled unless the inventory is high enough. In comparison, an OFR-centered approach focuses resources on a few customers and the rest may have to experience below-average service, which in general produces a much lower average fill rate across orders. When the inventory capital is relatively tight (equal to or less than $22 million), maximizing the expected OFR reduces the expected average CFR by 12.5% to 16.5% (compared to the CFR maximization), but the gap will be narrowed as more inventory capital becomes available. Insight 10 is thus stated as below:

*Insight 10: CFR and OFR are entirely different customer service measures with different inventory implications, which a company should carefully consider when adopting one metric or the other based on the nature of its business, especially when the inventory capital is tight.*

*Figure 2.10: Comparison of expected OFR vs. average CFR*



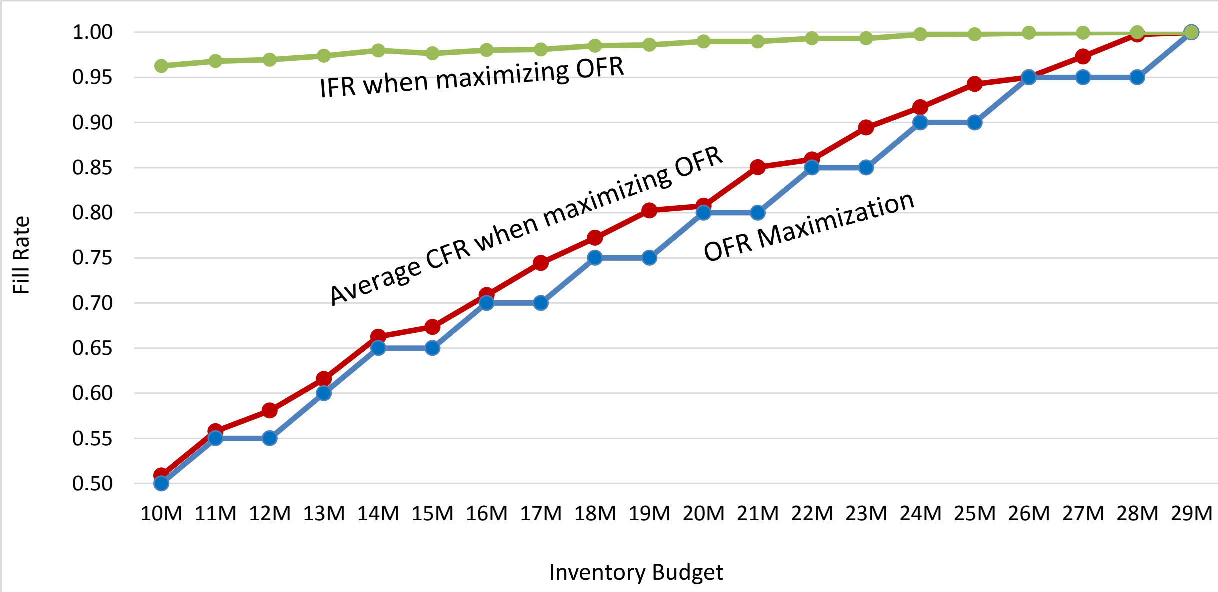
*Relationships of OFR, average CFR, and IFR under OFR-focused strategy*

If the interest of the company is to deliver as many perfect orders as possible, OFR is obviously the most appropriate performance measure. Maximizing the expected OFR under different levels of inventory capital produces the chart in Figure 2.11. It's interesting to note that the expected average CFR is very similar to the expected OFR, suggesting the two may be interchangeable in measuring the effectiveness of an OFR-focused strategy. Item fill rate, on the other hand, remains high (96.26%) even when the OFR is expected to be only 50%, supporting the findings by Anupindi and Tayur (1998) and Song (1998) that IFR is a product-based measure and cannot well reflect service at the order level. But IFR and OFR do appear to be positively related; higher IFR improves OFR. When the item fill rate is higher than 99.9%, 95% of orders are expected to be completely filled. However, the two measures won't become identical until the capital reaches $29 million where a perfect OFR can be expected. These observations suggest insights 11 and 12:

*Insight 11: The expected average CFR improves with the inventory investment and the expected number of perfectly filled orders when OFR is optimized.*

*Insight 12: IFR tends to be significantly greater than OFR and should not be used to measure order fulfillment, especially when the inventory capital is tight.*

*Figure 2.11: Expected OFR, average CFR, and IFR under OFR-focused measure*

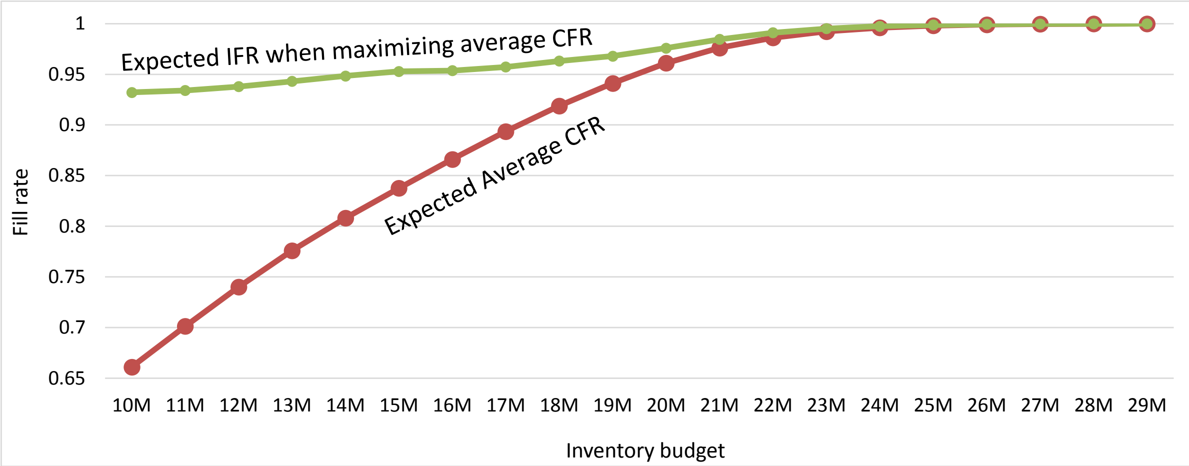



*<u>Relationship of average CFR and IFR under CFR-focused strategy</u>*

CFR measures the distance between shipment and order quantity in an order, which applies to the company that is concerned with the depth of the order fulfillment and focuses on improving the average fill rate across orders. With a tight inventory capital, the expected average CFR differs significantly from the IFR as showed in Figure 2.12; at a $10 million capital, the expected item fill rate is over 93%, while the average fill rate across orders is only 66%. As the inventory capital increases, the two get closer. When the inventory investment is beyond $22 million, the difference between the two is less than 0.3% and IFR provides a good indication on the performance of the average CFR. This leads to insight 13:

*Insight 13: IFR significantly exceeds the average CFR under a tight inventory capital, but may provide an alternative service measure when the investment is sufficient* *and inventory parameters are optimized for the maximum average CFR.*

*Figure 2.12: Expected average CFR and IFR under CFR-focused measure*




## 5 Conclusions

This chapter focuses on the customer service aspect of inventory management, specifically the order fulfillment performance. Two order-based measures are studied: the OFR and the average CFR. Two MILP optimization models are constructed to maximize the order fulfillment performance measures by endogenously selecting the appropriate inventory classes to which SKUs are assigned, subject to the inventory capital constraint and the minimum profit requirement. An extensive sensitivity analysis is performed over inventory capital and profits, as well as order fill rates, to gain insights into the tradeoff between order-based measures and profit, the relationship between the two order fulfillment measures, and their relationships with item fill rate.

### 5.1 Findings and Managerial Implications

Research results demonstrate that approximate Pareto efficiency frontiers can be delineated between profit and order fill rate, between profit and average fill rate across orders, and between the two order-based measures. Under a given inventory capital, an increase of customer service level causes a reduction in profitability, and the closer the current service level is to the maximum inventory-capital-constrained point, the more resources are required to make a single percentage-point improvement. Similarly, from the inventory investment perspective, a diminishing return on investment is observed for both profit and the average fill rate of customer orders as the inventory level becomes higher. In addition, the numerical experiments show that item fill rate, order fill rate and the average customer-order fill rate are very different measures. When the item fill rate is high (e.g., at 93%), the average customer-order fill rate may not be acceptable (merely at 66%) and the order fill rate could be very poor (e.g., less than 50%). Under a tight inventory capital, a strategy focusing on the number of perfect orders may cause a decline (e.g., over 10%) in the average customer-order fill rate, whereas a strategy focusing on the average customer-order fill rate may result in none of the orders being perfectly delivered. Furthermore, without additional inventory investment, the improvement of one performance measure may cause the decrease of the other(s). These findings have the following managerial implications for manufacturers and distributors managing finished goods inventory.

First, it is imperative for the firm to correctly match its performance measure with its performance measurement needs in inventory control. If the availability of a few key products is most important to the company, item fill rate (IFR) would be an appropriate measure. An example is the PC manufacturer that supplies different PC models and sells directly to end customers. A few most popular models are the major profit generators, and an end-customer in most cases would not buy multiple models. Thus, the IFR may be the appropriate measure. If the company's major clients are distributors and retailers who typically buy multiple products, then IFR may not be proper and the average CFR could be an effective measure. Most finished goods distributors are in this category. If the complete fulfillment of a customer order is most crucial, then OFR could be the appropriate measure. An example is a distributor that supplies multiple items to manufacturers whose end products are dependent upon those items. In this case, the complete fulfillment of an order could be vital to the distributor. In a nutshell, it is important to note that the three inventory-related measures, IFR, OFR, and the average CFR, in most cases are not interchangeable nor can each act as a sufficient indicator of the other(s). To create superior performance, the correct performance measure has to be in place. Inappropriate measures may cloud the real problems, lead the company to take incorrect actions, and eventually result in a growing number of unhappy customers.

Second, to improve performance, the company should use the proposed models to evaluate whether its current performance is Pareto optimal, especially if it operates under a tight inventory capital. If the firm is currently not on the Pareto curve, there is an opportunity for it to improve profitability and service level concurrently without making any additional investment. If the company is already operating optimally given the current resources, any improvement in service level will demand a sacrifice of profitability, which therefore should be carefully evaluated.

Third, selecting appropriate inventory capital is important for OFR improvement. A sensitivity analysis on inventory capital reveals that the relationships between the OFR and inventory capital and profit are not linear and do not even always move in the same directions. There are occasions where a slight decrease of inventory investment may not affect the OFR nor may it impact much on profit, and there are also instances where a slight increase of inventory capital may boost the OFR to the next level or significantly improve profitability. Without performing a sensitivity analysis, the firm may miss the opportunity to maximize the return on its inventory capital.

Fourth, the inventory capital required for the “perfect” performance of order fulfillment tends to be much higher than the inventory capital required for profit maximization, and a “perfect” OFR or average CFR also implies less profitability. Companies therefore need to balance short-term financial targets and long-term customer satisfaction. Securing a capital level that can at least maximize profit is recommended because that is where the least economic tradeoff is required to maintain a relatively high level of CFR.

Lastly, in some cases, the company may wish take into account different service performance measures simultaneously, with one acting as the primary objective. The proposed models already considered the minimum profit requirement. The IFR and OFR can also be included when maximizing the average CFR, and vice versa. For example, if some of the items and/or a few particular customers are strategically important to the company, minimum fill rates for the concerned items and/or customers can be posed as an additional constraint(s) in the optimization model. However, it should be noted that there is always a tradeoff among the average CFR, OFR, and profit no matter at what level of inventory investment unless the company is not operating optimally at present.

Additional constraints will affect the performance of the targeted measure, but it may be essential for the firm to have a balanced performance in various dimensions.

## 5.2 Contributions

This study contributes to the inventory literature in the following ways. First, it proposes an alternative measure, the customer-order fill rate, to gauge the fulfillment performance of the individual customer orders in a multi-product, finished goods inventory system. A major issue with the traditional order fill rate measure is its binary assessment on the order performance, which cannot offer any insight on how well a customer is served.

Second, linear optimization models are proposed with the consideration of minimum profit requirement and inventory capital restriction. The models can be easily extended to include additional constraints such as the item fill rate. And the model formulation for the order fill rate, in particular, gives managers the visibility on how well each order is handled and also allows them to “require” perfect order fill rates for the key customers. Further, companies may utilize the models to quantify the level of inventory needed for each SKU to improve the service level for individual customers or get a sense on the total resources required to increase the overall fill rate.

Finally, this research adds new evidence to the understanding of how item-based service measure may relate to order-based service measures. The results from the numerical experiments support the finding by Anupindi and Tayur (1998) and Song (1998) that the item fill rate often significantly overstates the order fill rate, but contradict the finding by Larsen and Thorstenson (2014) that the average customer-order fill rate always exceeds the item fill rate. Research results of this study demonstrate that the item fill rate is always the highest among the three measures even though the inventory is organized for the maximization of order-based performance, but the item fill rate may adequately indicate the average customer-order fill rate when the inventory capital is sufficiently large.

## 5.3 Limitations and Future Research

This study has a number of limitations that may be addressed in future research. First, similar to many of the inventory models, the performance of the proposed models depend on the quality of sales forecasting, especially order-level forecasting. Many manufacturers and distributors today still use shipment data for demand forecasting, which can never truly capture the real demand unless the ordered quantity is 100% shipped. Research addressing the issue of effectively tracking customer orders and forecasting at order-SKU level would represent a significant contribution to both the marketing literature and the inventory literature. Second, further studies on the relationship between the item fill rate and the average customer-order fill rate based on the different demand processes would provide further insights and comparison with the results of this work. Third, customer orders might be correlated due to the macroeconomic factors. Extending the models to include coupling factors that reflect the connections of orders would address this situation. Fourth, there are cases where items in shortage may be put on backorder. Extending the models to include backorder through a penalty cost would be useful. Fifth, a more sophisticated approach may consider the demand pattern to be endogenously determined as a function of service level. Sixth, there might be occasions where SKUs offered by the company are substitutable with each other.

Extending the models to take into account directly competitive or substitutable relationships among SKUs would be of value. Finally, this study shows that there are tradeoffs among the two order-based measures, item fill rate, and profit. Research that offers guidelines on the situation where the company should put profit as the first priority, the situation where a particular tradeoff of profit and order fulfillment performance is most beneficial, the situation where a certain level of inventory capital should be secured, etc. would be of great value to the practitioners.

# Chapter 3 Summary

This dissertation addresses the financial and customer service aspects of inventory management, in an effort to fill the gaps in the inventory literature and provide business with practical decision-support tools. Chapter 1 presents a multi-period inventory optimization model that is built upon the NETFORM framework and explicitly addresses nonstationary demand, arbitrary review periods, SKU-specific lead times, and diverse inventory holding costs, with the objective of maximizing the net present value of profit. The model is evaluated against an "advanced" multi-criteria inventory classification scheme through a real-world case, and is also compared with the more commonly used ABC approach through extensive computational experiments. Results show that the optimal dynamic inventory classification and control decisions obtained from the model significantly reduce both safety stock and base stock levels compared to the aforementioned two approaches, and demonstrate a superior financial performance. Experiments conducted around key inventory parameters show that the benefit of the optimization model increases significantly with the cost of capital, lead times, and inventory holding costs. With the proposed multi-period optimization model, management may obtain an advanced knowledge about the target inventory levels, possible safety stocks, and replenishment order quantities in each future time period at the beginning of the planning horizon, which would greatly facilitate the negotiation process with suppliers and support production planning.

Chapter 2 presents two MILP optimization models that are developed to maximize the order fulfillment performance measures by endogenously selecting the appropriate inventory classes to which SKUs are assigned, subject to the inventory capital constraint and the minimum profit requirement. In this chapter, one model maximizes the traditional order fill rate which assesses whether or not an order is completely fulfilled, and the other maximizes the average customer-order fill rate which gauges the extent to which an order is satisfied. Computational experiments show that there exist tradeoffs between profit and order fill rate, between profit and average customer-order fill rate, and between the two order-based measures. Under a given inventory capital, an increase of customer service level causes a reduction in profitability, and the closer the current service level is to the maximum inventory-capital-constrained point, the more resources are required to make a single percentage-point improvement. Further, results show that item fill rate, order fill rate and the average customer-order fill rate are in most cases are not interchangeable nor can each act as a sufficient indicator of the other(s). With limited inventory capital, an OFR-focused strategy is less profitable than an average CFR-focused strategy. And the IFR in most cases is significantly greater than the OFR unless inventory capital is sufficiently large. Companies should correctly match its performance measure with its performance measurement needs in inventory control.

The modeling frameworks developed in this research offer a variety of future research opportunities. For example, the current research is for a three-tier supply chain network. It will be interesting to develop new models to address system-wide supply chain inventory costs in more general networks. Another promising research direction is to apply stochastic optimization and integrated simulation-optimization methodologies to address random demand with general probability distributions. In addition, the models assume demand is exogenously given. A promising Supply Chain – Marketing interdisciplinary research is to consider demand being endogenously determined by inventory control policies and service level. Further, the purchasing cost of an SKU may vary by order quantity due to economies of scale. A valuable extension is to set multi-tiered SKU costs based upon replenishment quantity and extend the model to determine optimal ordering frequencies and quantities.

## References

Applied Maintenance Supplies & Solutions. (n.d.). About Applied Maintenance Supplies & Solutions. Retrieved from http://www.appliedmss.com/aboutus.html on Jan. 21, 2015.

Ammar, M. A. (2010). Optimization of project time-cost trade-off problem with discounted cash flows. *Journal of Construction Engineering and Management*, 137(1), 65-71.

Anupindi, R. and Tayur, S. (1998). Managing stochastic multiproduct systems: model, measures, and analysis. *Operations Research*, *46*(3), S98-S111.

Arnold, B. (2008). Pareto and Generalized Pareto Distributions. *In Modeling Income Distributions and Lorenz Curves*; Chotikapanich, D., Ed.; Springer: New York, 2008; Vol. 5, pp 119– 145.

Ballou, R. H. (2003). *Business Logistics Management*, 5th edition, Prentice-Hall.

Boute, R.N. and Lambrecht, M. (2005). Exploring the bullwhip effect by means of spreadsheet simulation, *FETEW Research report*, KBF_0706, Kuleuren, Vol. 28, http://www.econ .kuleuven.be/fetew/pdf_publicaties/KBI_0706.pdf.

Bowersox, D.J., Closs, D.J., and Cooper, M.B. (2012). *Supply chain logistics management*, 4th edition, McGraw-Hill, New York, NY.

Brightstar Intelligence (2014). 2014 Mobile Trends Report. Retrieved from http://www.brightstar.com/assets/uploads/docs/2014_Mobile_Trends.pdf on April 6, 2016.

Chase, R. B., Aquilano, N.J., and Jacobs, F.R. (2001). *Operations management for competitive advantage*, 9th edition, Irwin-McGraw Hill.

Chen, F. and Song, J.S. (2001). Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research*, *49*(2), 226-234.

Chen, Y., Li, K.W., Kilgour, D.M., and Hipel, K.W. (2008). A case-based distance model for multiple criteria ABC analysis. *Computers & Operations Research*, *35*(3), 776-796.

Chwif, L., Barretto, M.R.P. and Saliby, E. (2002). Supply chain analysis: spreadsheet or simulation, *Proceedings of the 2002 Winter Simulation Conference*, pp.59–65.

Closs, D.J., Nyaga, G.N., and Voss, M. D. (2010). The differential impact of product complexity, inventory level, and configuration capacity on unit and order fill rate performance. *Journal of Operations Management*, *28*(1), 47-57.

Cormier, G. and Gunn, E. A. (1996). On coordinating warehouse sizing, leasing and inventory policy. *IIE transactions*, *28*(2), 149-154.

Damodaran, A. (2016). Cost of Capital by Sector (US). Retrieved from http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/wacc.htm on December 10, 2015.

Drezner, Z. and Hamacher, H.W. (2004). *Facility Location: Applications and Theory*, New York, Springer.

Ernst, R. and Cohen, M.A. (1990). Operations related groups (ORGs): a clustering procedure for production/inventory systems. *Journal of Operations Management*, 9(4), 574-98.

Ettl, M., Feigin, G.E., Lin, G.Y., and Yao, D.D. (2000). A supply network model with base-stock control and service requirements. *Operations Research*, *48*(2), 216-232.

Flores, B.E., Olson, D.L. and Dorai, V.K. (1992). Management of multicriteria inventory classification. *Mathematical and Computer Modelling*, 16(12), 71-82.

First Research (2014). *Industrial Supply Wholesalers - Quarterly Update 11/17/2014*. Austin: Hoover's Inc.

Glover, F., Klingman, D. and Phillips, N. V. (1992). *Network models in optimization and their applications in practice*, John Wiley & Sons, New York.

Goic, M., Bosch, M. and Castro, J.P. (2015). Detecting inefficiently managed categories in a retail store. *The Journal of the Operational Research Society*, 66(1), 160.

Graves, S.C. (1999). A single-item inventory model for a nonstationary demand process. *Manufacturing & Service Operations Management*, *1*(1), 50-61.

Graves, S.C. and Willems, S. P. (2008). Strategic inventory placement in supply chains: Nonstationary demand. *Manufacturing & Service Operations Management*, *10*(2), 278-287.

Hadi-Vencheh, A. (2010). An improvement to multiple criteria ABC inventory classification. *European Journal of Operational Research*, *201*(3), 962-965.

Hatefi, S. M., Torabi, S. A., & Bagheri, P. (2014). Multi-criteria ABC inventory classification with mixed quantitative and qualitative criteria. *International Journal of Production Research*, *52*(3), 776-786.

Hausman, W.H., Lee, H.L., and Zhang, A. X. (1998). Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, *109*(3), 646-659.

Hoen, K.M., Güllü, R., Van Houtum, G.J., and Vliegen, I.M. (2011). A simple and accurate approximation for the order fill rates in lost-sales assemble-to-order systems. *International Journal of Production Economics*, *133*(1), 95-104.

Jung, D.Y., Han, S.H., Im, K.S. and Ryu, C.K. (2007). Modeling an inventory management in construction operations involving on-site fabrication of raw materials, *Proceedings IGLC-15*, Michigan, USA, July, pp.367–379.

Karlin, S. (1960). Dynamic inventory policy with varying stochastic demands. *Management Science*, *6*(3), 231-258.

Knod, E.M. and Schonberger, R. (2001). *Operations management: meeting customers' demand*, 7th edition, McGraw-Hill/Irwin, New York.

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

Lajili, I., Babai, M. Z., & Ladhari, T. (2012). Inventory performance of multi-criteria classification methods: An empirical investigation. *Proceedings of MOSIM'12*.

Larsen, C., & Thorstenson, A. (2008). A comparison between the order and the volume fill rate for a base-stock inventory control system under a compound renewal demand process. *Journal of the Operational Research Society*, *59*(6), 798-804.

Larsen, C. and Thorstenson, A. (2014). The order and volume fill rates in inventory control systems. *International Journal of Production Economics*, *147*, 13-19.

Lovejoy, W.S. (1990). Myopic policies for some inventory models with uncertain demand distributions. *Management Science*, *36*(6), 724-738.

Lu, Y., Song, J.S., and Yao, D. D. (2003). Order fill rate, leadtime variability, and advance demand information in an assemble-to-order system. *Operations Research*, *51*(2), 292-308.

Mahamani, A., Prahlada Rao, K., & Pandurangadu, V. (2008). The development of a simulation-based approach to optimise the inventory policy in a single-echelon supply chain: a case study. *International Journal of Data Analysis Techniques and Strategies*, 1(2), 173-192.

Millstein, M. A., Yang, L., & Li, H. (2014). Optimizing ABC inventory grouping decisions. *International Journal of Production Economics*, *148*, 71-80.

Mohammaditabar, D., Ghodsypour, S.H. and O'Brien, C. (2012). Inventory control system design by integrating inventory classification and policy selection. *International Journal of Production Economics*, 140(2), 655-659.

Molenaers, A., Baets, H., Pintelon, L., & Waeyenbergh, G. (2012). Criticality classification of spare parts: A case study. *International Journal of Production Economics*, *140*(2), 570-578.

Neale, J.J. and Willems, S.P. (2009). Managing inventory in supply chains with nonstationary demand. *Interfaces*, *39*(5), 388-399.

Partovi, F.Y. and Anandarajan, M. (2002). Classifying inventory using an artificial neural network approach. *Computers & Industrial Engineering*, *41*(4), 389-404.

Porras, E. and Dekker, R. (2008). An inventory control system for spare parts at a refinery: an empirical comparison of different re-order point methods. *European Journal of Operational Research*, 184(1), 101-32.

Ramanathan, R. (2006). ABC inventory classification with multiple-criteria using weighted linear optimization. *Computers & Operations Research*, 33(3), 695-700.

Rao, S., Griffis, S. E., and Goldsby, T. J. (2011). Failure to deliver? Linking online order fulfillment glitches with future purchase behavior. *Journal of Operations Management*, *29*(7), 692-703.

Scarf, H. (1959). The optimality of (s, S) policies in the dynamic inventory problem, Chapter 13 in Arrow K.J., Karlin S. and Suppes P. (Eds). *Mathematical methods in the social science*, Stanford University Press, Standford, CA.

Sezen, B. and Kitapci, H. (2007). Spreadsheet simulation for the supply chain inventory problem, *Journal of Production Planning and Control*, January, 18(1), pp.9–15.

Shang, K. H. (2012). Single-stage approximations for optimal policies in serial inventory systems with nonstationary demand. *Manufacturing & Service Operations Management*, *14*(3), 414-422.

Shao, X. F. and Dong, M. (2012). Comparison of order-fulfilment performance in MTO and MTS systems with an inventory cost capital constraint. *International Journal of Production Research*, *50*(7), 1917-1931.

Silver, E. A. (2008). Inventory management: An overview, Canadian publications, practical applications and suggestions for future research. *INFOR: Information Systems and Operational Research*, *46*(1), 15-28.

Song, J.S. (1998). On the order fill rate in a multi-item, base-stock inventory system. *Operations Research*, *46*(6), 831-845.

Song, J. S. and Yao, D. D. (2002). Performance analysis and optimization of assemble-to-order systems with random lead times. *Operations Research*, *50*(5), 889-903.

Soylu, B., & Akyol, B. (2014). Multi-criteria inventory classification with reference items. *Computers & Industrial Engineering*, *69*, 12-20.

Stanford, R. E., & Martin, W. (2007). Towards a normative model for inventory cost management in a generalized ABC classification system. *Journal of the Operational Research Society*, *58*(7), 922-928.

Stock, J.R. and Lambert, D.M. (2001). *Strategic logistics management*, 4th edition, McGraw-Hill/Irwin, New York.

Syntetos, A.A., Keyes, M. and Babai, M.Z. (2009). Demand Categorisation in a European Spare Parts Logistics Network. *International Journal of Operations and Production Management*, 29(3), 292-316.

Tarim, S.A. and Kingsman, B.G. (2006). Modelling and computing (R n, S n) policies for inventory systems with non-stationary stochastic demand. *European Journal of Operational Research*, *174*(1), 581-599.

Teunter, R. H., Babai, M. Z., and Syntetos, A. A. (2010). ABC classification: service levels and inventory costs. *Production and Operations Management*, *19*(3), 343-352.

Treharne, J.T. and Sox, C.R. (2002). Adaptive inventory control for nonstationary demand and partial information. *Management Science*, *48*(5), 607-624.

Tunc, H., Kilic, O. A., Tarim, S. A., & Eksioglu, B. (2011). The cost of using stationary inventory policies when demand is non-stationary. *Omega*, *39*(4), 410-415.

van Kampen, T. J., Akkerman, R., & van Donk, D. P. (2012). SKU classification: A literature review and conceptual framework. *International Journal of Operations & Production Management*, *32*(7), 850-876.

Viswanathan, S. and Bhatnagar, R. (2005). The application of ABC analysis in production and logistics: an explanation for the apparent contradiction. *International Journal of Services and Operations Management*, *1*(3), 257-267.

Zinn, W., Mentzer, J.T., and Croxton, K.L. (2002). Customer-based measures of inventory availability. *Journal of Business Logistics*, *23*(2), 19-43.