An Efficient Heuristic Optimization Algorithm for a Two-Echelon (*R*, *Q*) Inventory System

Authors and Affiliations:

Mohammad H. Al-Rifai, Ph.D., Propak Corporation

Manuel D. Rossetti[1], Ph. D., University of Arkansas

Abstract

This paper presents a two-echelon non-repairable spare parts inventory system that consists of one warehouse and *m* identical retailers and implements the reorder point, order quantity (*R*, *Q*) inventory policy. We formulate the policy decision problem in order to minimize the total annual inventory investment subject to average annual ordering frequency and expected number of backorder constraints. In order to solve the problem, we decompose the system by echelon and location, derive expressions for the inventory policy parameters, and develop an iterative heuristic optimization algorithm. Experimentation showed that our optimization algorithm is an efficient and effective method for setting the policy parameters in large-scale inventory systems.

Keywords: Inventory Optimization, Multi-Echelon, Heuristics

[1] 4207 Bell Engineering Center, Department of Industrial Engineering, Fayetteville, AR 72701, USA, Phone: (479) 575-6756, Fax: (479) 575-8431, Email: rossetti@uark.edu

1. Introduction

Large multi-echelon, multi-item inventory systems usually consist of hundreds of thousands of stock keep units (SKUs). These SKUs can be classified into two main categories: consumables and repairables. Calculating the optimal inventory policy parameters for each SKU is a computational burden that necessitates the need for efficient policy setting techniques that reduce the computational time, and at the same time, improve the ability of inventory managers to more effectively manage the supply chain. Multi-echelon inventory systems are important to large corporations and to the military to support their operations.

In large supply networks like Wal-Mart, and the US-Navy, thousands of SKUs are stocked at different inventory holding points (IHPs). These holding points might be at different echelons where the higher echelons supply the lower echelons. Each of these IHPs might follow different stocking policies resulting in decentralized control of the supply network. This case is most likely to occur when each of the locations that constitute the supply network are owned by different owners who are not willing to give control of their inventories to external parties. Under this case, each location might not take into consideration interactions with the other locations that might have a significant effect on the efficiency of the whole supply network as well as on each single location. On the other hand, if all of these locations are owned or managed by a centralized management system, a single inventory control system might be implemented. Previous research shows that tremendous improvements are attainable if a centralized inventory management system is considered for the entire supply network. This motivated building inventory models that consider the entire supply network and the interactions between their constituent IHPs. Most of these models have their own assumptions and characteristics. Some of these models, as we will see in the next section, are built for a special class of supply networks such as slow moving and expensive spare part supply networks. Other models are built for a particular structure of a supply network that might not be applicable to other supply networks. Hence, modeling multi-echelon inventory systems is still a rich area for research.

In this research, we model a two echelon inventory system that implements (R, Q) policies at each IHP at each echelon. We consider a centralized inventory management system under which interactions between IHPs at different echelons are allowed. Calculating optimal inventory policies for each item at each location in a multi-echelon inventory system requires efficient solution procedures that can handle *large scale* inventory systems, reduce the associated computational time, and reduce modeling complexity due to the dependency between echelons. In a multi-echelon inventory system that implements (R, Q) policies, modeling complexity arises when modeling the effect of the delay at the replenishment source due to stockout on the lead times of the lower echelons and modeling the lead time demand process at the higher echelons.

We formulate the policy setting problem in order to minimize the total annual inventory investment subject to average annual order frequency and expected number of backorder constraints. Due to the complexity of the inventory modeling, we derived expressions for the policy parameters at each location at each echelon under different lead time assumptions such as deterministic lead times and stochastic lead times (due to stockout at the warehouse). In order to calculate inventory policy parameters and incorporate the effect of the delay at the warehouse due to stockout, we developed a two-echelon heuristic optimization algorithm that implements these expressions.

The rest of this paper is organized as follows. In Section 2, we provide a literature review of important multi-echelon inventory models that have been developed and implemented. In

Section 3, we present our problem definition and model formulation. In Section 4, we present and discuss our solution procedure. We implement, validate and experiment with the optimization algorithm in Section 5. Finally, in Section 6 we conclude and provide extensions for future work.

2. Literature Review

One of the most important multi-echelon, multi-item inventory models for spare parts management is METRIC. METRIC is the Multi-Echelon Technique for Recoverable Items Control, developed by Sherbrooke (1968) and it is used for setting repairable items inventory control policies using the base stock model. The base stock model is a special case of the reorder point, order quantity inventory policy, where the reorder quantity Q=1 and it is usually used with expensive, slow moving items, and when the holding and backorder costs dominate. The objective function in METRIC is minimizing the expected number of backorders at the base level, subject to budget constraints while setting optimal inventory policy parameters. In the case of low or medium cost items with medium to high demand rates, the (R, Q) policy may be more appropriate.

Many inventory models have been developed for expensive, low demand, and repairable spare parts (e.g. Sherbrooke, 1968; Graves, 1985; Diaz and Fu, 1997; Caglar *et al*., 2004) where the base stock model is implemented at least at one echelon of the supply network. First indenture spare parts are only considered for repair, where the repair operations are performed at each facility at the first echelon or at the distribution center. In other research, multi-indenture repairable spare parts have been considered where lower indentures are modeled (e.g. Muckstadt, 1973). The base stock model is also implemented in systems that support consumable spare parts (e.g. Axsäter, 1990; Hopp *et al.*, 1999).

Deuermeyer and Schwarz (1981) presented an analytical model for estimating the expected performance measures of a one-warehouse, *m* identical retailers, and non-repairable spare parts inventory system. They examined a system that involves *m* identical retailers facing stationary Poisson demand and operating under (*R*, *Q*) replenishment policies. In their research, the main challenge was to model the demand process at the warehouse which is a superposition of the retailer's ordering processes. Since they implemented (*R*, *Q*) policies at the retailers, they ordered in batches of units of items. In this case, the demand process at the warehouse is not a superposition of simple Poisson processes. Instead, it is a superposition of the retailer's ordering processes. Since the demand rate at each retailer for each item is λ, and the retailer's order batch size is *Q*, the demand process at the warehouse is a superposition of renewal processes with *Q* stages and rate λ (Deuermeyer and Schwarz, 1981). Unfortunately, the renewal property is not preserved under superposition (Torab and Kamen, 2001). More precisely, except for Poisson sources, the inter-arrival times in the superposition process are statistically dependent, a property that cannot be captured by a renewal model (Torab and Kamen, 2001). Hence, Deuermeyer and Schwarz (1981) approximated the demand process at the warehouse that is generated by identical retailers by a renewal process and derived expressions that approximate the mean and variance of the warehouse lead time demand.

Svoronos and Zipkin (1988) proposed a refinement of the Deuermeyer and Schwarz model. At the warehouse level, they estimated differently the mean and variance of the warehouse lead time demand. They approximated the warehouse lead time demand using a mixture of two translated Poisson distributions (MTP). Using the MTP, they estimated the performance measures at the warehouse such as the expected number of backorders, which they used later to calculate the delay at the warehouse due to stockout.

Hopp *et al.* (1997) considered a single location that utilizes (R, Q) policies and presented three heuristics that approximate the inventory policy parameters. Using some approximations and the theory of Lagrange multipliers, they derived simple expressions for the inventory policy parameters. Hopp *et al.* (1999) considered a two-echelon spare parts stocking and distribution system with an objective function of minimizing total average inventory investment in the entire system subject to constraints on average annual order frequency and total average delay at each facility due to stockout. At the warehouse, they implemented an (R, Q) policy while at each retailer they implemented a base stock model and assumed the demand process is a Poisson process. Therefore, the demand process at the warehouse is a superposition of Poisson processes which is also a Poisson process. Since they incorporated the effect of delay at the warehouse, the service measures at each retailer depend on the delay at the warehouse due to stockout. The average number of backorders at the warehouse is a function of the inventory policy parameters at the warehouse. In order to derive expressions that estimate the policy parameters at both echelons, they decomposed the system by level and by facility. First, they modeled the warehouse and then they modeled each facility. Hopp and Spearman (2001) presented a multi-product (R, Q) backorder model with an optimization algorithm that estimate the inventory policy parameters at a single facility that is faced with Poisson demands and assume fixed lead time.

Cohen *et al.* (1990) developed a multi-echelon inventory model for the IBM network in the United States. IBM’s network is a large multi-echelon system that consists of four main echelons with over 15 million part-location combinations and over 50,000 product-location combinations.

They developed and implemented a system called *Optimizer* that determines stocking policies for each part at each location. Their objective was to determine the stocking policies for each part at each location. They considered emergency shipments, holding costs, replenishment costs (includes transportation, handling, and ordering costs). In order to solve the problem, they decomposed the model development into three stages; a one-part, one location model, a multi-product, one location model, and a multi-product, multi-echelon model. In developing the one-part, one-location problem they developed a periodic review, stochastic model.

In the inventory systems under consideration, the stocking policies at any given facility depend directly on the stocking policies of the facility's supplier. The effective lead time at any facility at any echelon is mainly a function of two components, the transportation times (including ordering, receiving, and handling the order, etc) and the random delay at the supplier due to stockout. Under decomposition, each facility is modeled under the assumption of ample supply at its supplier. Hence, the effective lead times are only a function of the transportation times which are assumed to be constant in many situations. Cohen *et al*. (1990) assumed deterministic lead times, and treated each echelon independent of the other echelons, i.e. there is always ample supply at the replenishment source. According to Cohen *et al*. (1990) such a solution procedure is likely close to optimality in cases where service requirements at all sites are high. Their methodology for decomposing the system by level and assuming constant lead time is an efficient one, in which, the system is simplified. As we can see, decomposing the systems is widely used and has been shown to be efficient in solving such complicated systems.

3. Problem Definition and Model Formulation

We build on the previous research by modeling a two echelon inventory system that implements (R, Q) policies at each location. Figure 1 shows a two-echelon inventory system that consists of an external supplier that can supply any item with a given lead time and a single warehouse that supplies any number of independent identical retailers.

FIGURE 1 ABOUT HERE

Under this system, the retailers are faced with demands that are generated by random failures of the spare parts at the customer's sites according to a Poisson process. Since the demand process at each retailer for each item is a Poisson process, the demand process at any warehouse is a superposition of the retailer's ordering processes. Specifically, it is a superposition of renewal processes each with an Erlang interrenewal processes time with Q_{ri} stages and rate per state λ_{ri} (Svoronos and Zipkin, 1988).

The above two-echelon (R, Q) inventory system operates as follows. When a retailer is faced with a demand, the demand is satisfied from shelves if the amount demanded is less or equal to the number of units available. Otherwise, the demand is backordered. Under a (R, Q) policy, item i's inventory position at retailer r is checked continuously, if it drops to or below its reorder point R_{ri}, a replenishment order of size Q_{ri} is placed at the warehouse. The inventory position is defined as the on-hand inventory plus the on-order inventory minus the number of outstanding backorders. After placing an order with the warehouse, an effective lead time ℓ_{ri} elapses between placing the order and receiving it. After receiving the replenishment order, the outstanding backorders at the retailer are immediately satisfied according to a First-In-First-Out (FIFO) policy.

Since the same policy is followed at the warehouse, the retailer replenishment orders placed at the warehouse are satisfied if the on-hand inventory at the warehouse is greater or equal to the retailer's replenishment order size. In other words, a partial filling of an order at the warehouse is not allowed. This is a plausible policy not uncommon in practice, especially when there is a fixed cost connected to each transport (Anderson and Marklund, 2000). The warehouse inventory position for each item is checked continuously. If it drops to or below the reorder point R_{wi}, a replenishment order Q_{wi} is placed with the supplier, where a deterministic lead time L_{wi} elapses between placing the order and receiving it. After receiving the replenishment order, the outstanding backorders at the warehouse are immediately satisfied according to a FIFO policy.

Before proceeding in developing the model, we state our assumptions as follows. We model a two echelon inventory system, where each retailer is replenished by only one warehouse. The demand process at each retailer occurs according to a Poisson process. All orders that are not satisfied from on hand inventory are backordered (i.e. lost sales are not considered). The warehouse's supplier has infinite capacity with a fixed lead time, the warehouse has limited supply, and no lateral shipments are permitted between the retailers. We do not model the delivery process from the retailer to the end customer.

The following is a list of the notation that we will use throughout the paper:

w = Warehouse index

r = Retailer index

i = Item index

m = Number of retailers

N = Number of items

F_r = Target order frequency at retailer r (orders per year)

F_w = Target order frequency at the warehouse (orders per year)

B_r = Target number of backorders at retailer r

B_w = Target number of backorders at the warehouse

λ_{ri} = Item *i* demand rate at retailer *r* (units/year)

λ_{wi} = Item *i* demand rate at the warehouse (in units of item *i* batch size at the retailer per year)

L_{ri} = Item *i* lead time (ordering and transportation) at retailer *r* (years)

L_{wi} = Item *i* lead time (ordering and transportation) at the warehouse (years)

ℓ_{ri} = Item *i* effective lead time at retailer *r* (years)

C = Total inventory investment at both echelons ($)

c_i = Item *i* unit cost ($)

c = Superscript that represents the current value.

p = Superscript that represents the previous value.

e = Tolerance.

Q_{ri} = Item *i* replenishment batch size at retailer *r* (units)

R_{ri} = Item *i* reorder point at retailer *r* (units)

Q_{wi} = Item *i* replenishment batch size at the warehouse (in units of Q_{ri})

R_{wi} = Item *i* reorder point at the warehouse (in units of Q_{ri})

$\bar{I}_{ri}(R_{ri}, Q_{ri})$ = Item *i* expected on-hand inventory at retailer *r* (units)

$\bar{I}_{wi}(R_{wi}, Q_{wi})$ = Item *i* expected on-hand inventory at the warehouse (in units of Q_{ri})

$\overline{B}_{ri}(R_{ri}, Q_{ri})$ = Item *i* expected number of backorders at retailer *r* (units). Also, B_{ri}

$\overline{B}_{wi}(R_{wi}, Q_{wi})$ = Item *i* expected number of backorders at the warehouse (in units of Q_{ri}). Also, B_{wi}

$\phi(x)$ = The pdf of the standard normal distribution function

Ö(x) = The cdf of the standard normal distribution function

Ö$^{-1}(x)$ = The inverse of the cdf of the standard normal distribution function

F_{ri} = Item *i* average order frequency at retailer *r*

F_{wi} = Item *i* average order frequency at the warehouse

ς_r, α_r, & ς_w = Lagrange multipliers that represents the ordering costs at the retailers and the warehouse

k_r, δ_r & k_w = Lagrange multipliers that represents the backordering costs at the retailers and the warehouse

We assume identical retailers and formulate the two-echelon (*R*, *Q*) policy problem in order to minimize the total annual inventory investment at both echelons subject to the following average annual order frequency and average number of backorder constraints:

$$\text{Average annual order frequency at each retailer} \leq F_r \tag{1}$$

$$\text{Average annual order frequency at the warehouse} \leq F_w \tag{2}$$

$$\text{Total expected number of backorders at each retailer} \leq B_r \tag{3}$$

$$\text{Total expected number of backorders at the warehouse} \leq B_w \tag{4}$$

We represent the above model mathematically as follows:

$$\textit{Minimize}\ \ C = m\sum_{i=1}^{N} c_i \bar{I}_{ri}\left(R_{ri}, Q_{ri}\right) + \sum_{i=1}^{N} c_i Q_{ri} \bar{I}_{wi}\left(R_{wi}, Q_{wi}\right) \tag{5}$$

Subject to

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{ri}}{Q_{ri}} \leq F_r \tag{6}$$

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{wi}}{Q_{wi}} \leq F_w \tag{7}$$

$$\sum_{i=1}^{N} \overline{B}_{ri}\left(R_{ri}, Q_{ri}\right) \leq \mathrm{B}_r \tag{8}$$

$$\sum_{i=1}^{N} \overline{B}_{wi}\left(R_{wi}, Q_{wi}\right) \leq \mathrm{B}_w \tag{9}$$

$$R_{ri} \geq -Q_{ri}, i = 1, 2 \ldots N \tag{10}$$

$$R_{wi} \geq -Q_{wi}, i = 1, 2 \ldots N \tag{11}$$

$$Q_{ri} \geq 1, i = 1, 2, \ldots N \tag{12}$$

$$Q_{wi} \geq 1, i = 1, 2 \ldots N \tag{13}$$

$$Q_{ri}, R_{ri}, Q_{wi}, \& R_{wi} : \text{Integers}, i = 1, 2, \ldots N \tag{14}$$

Constraints 10 and 11 are used to make sure that the outstanding backorders are satisfied when a replenishment order is received. This means that, customer orders will be satisfied from the retailer's replenishment order that has been placed when the customer placed the order or from orders that have been placed with the retailer prior to the customer's order. Also, the retailer's order will be satisfied from the warehouse replenishment order that has been placed when the retailer placed its order or from orders that have been previously placed with the supplier. Constraints 12 and 13 are used to make sure that the minimum allowable replenishment order size is one. Constraint 14 is necessary, since in real life no partial parts are ordered. Later on, in order to simplify the problem, constraint 14 will be relaxed to allow for continuous values. Under an (R, Q) policy the expected on-hand inventory for item i at any location when the demand during lead time is modeled using a discrete distribution (under which the inventory level declines in discrete steps) is defined as follows (Hadley & Whitin, 1963):

$$\bar{I}_i = \overline{B}_i\left(R_i, Q_i\right) + R_i + \frac{Q_i + 1}{2} - E[D_i] \tag{15}$$

Where, $E[D_i]$ is item i expected lead time demand and $\overline{B}_i(R_i,Q_i)$ is item i expected number of backorders at any time. Since almost all real-world systems involve discrete inventory, it generally makes sense to use the discrete inventory formula (Eq. 15) even when a continuous model is used to compute the policy parameters (Hopp and Spearman, 2001). Hence, we evaluate the inventory level using Eq. 15. Since the demand process for item *i* at retailer *r* is a simple Poisson process with an annual rate $\ddot{e}_{ri}$, item *i*'s expected lead time demand at retailer *r* is:

$$E[D_{ri}] = \ddot{e}_{ri} \times \ell_{ri} \tag{16}$$

$$\ell_{ri} = L_{ri} + d_{ri} \tag{17}$$

The first part of Eq. 17, specifically L_{ri}, represents item *i*'s transportation time from the warehouse to retailer *r*. Since non-repairable spare parts are considered, no parts are shipped back to the warehouse. Hence, no explicit assumption is made on the transportation time from any retailer to the warehouse. Also, ordering times are assumed to be negligible and transportation times are assumed to be deterministic. The second part of Eq. 17, specifically d_{ri}, is the delay at the warehouse due to stockout and it is given as follows (Svoronos and Zipkin, 1988 and Deuermeyer and Schwarz, 1981):

$$d_{ri} = \frac{\overline{B}_{wi}(R_{wi}, Q_{wi})}{\lambda_{wi}} \tag{18}$$

Since the demand process at each retailer is a Poisson process and an (*R*, *Q*) policy is implemented at each retailer, the demand process at the warehouse is a superposition of all the retailers ordering processes. Specifically, it is a superposition of independent renewal processes each with an Erlang inter-renewal time with Q_{ri} stages and rate per state λ_{ri} (Svoronos and Zipkin, 1988). Dividing the demand rate (λ_{ri}) for item *i* at retailer *r* during a given period of time by its reorder batch size (Q_{ri}) yields the number of replenishment orders during that period, i.e. the order frequency. Thus, item *i*'s order frequency at retailer *r* is:

$$\mathrm{F}_{ri} = \frac{\ddot{e}_{ri}}{Q_{ri}} \tag{19}$$

Under the assumption of identical retailers item *i*'s demand rate at the warehouse (λ_{wi}) is:

$$\ddot{e}_{wi} = m\mathrm{F}_{ri} = \frac{m\ddot{e}_{ri}}{Q_{ri}} \tag{20}$$

Svoronos and Zipkin (1988) derived the following expressions for the mean and variance of the warehouse lead time demand under the assumption of identical independent retailers:

$$E[D_{wi}] = \frac{m\ddot{e}_{ri}L_{wi}}{Q_{ri}} \tag{21}$$

$$V[D_{wi}] = \frac{\lambda_{ri}L_{wi}m}{Q_{ri}^2} + \frac{m}{Q_{ri}^2}\sum_{k=1}^{Q_{ri}-1}\left(\frac{[1-\exp(-\alpha_k\lambda_{ri}L_{wi})\cos(\beta_k\lambda_{ri}L_{wi})]}{\alpha_k}\right) \tag{22}$$

$$\text{Where } \alpha_k = 1-\cos(2\pi k/Q_{ri}) \tag{23}$$

$$\beta_k = \sin(2\pi k/Q_{ri}) \tag{24}$$

We use the normal approximation to the Poisson distribution to approximate the distribution of the retailer's lead time demand. In addition, we approximate the distribution of the warehouse lead time demand using a normal distribution with mean and variance as given by Eq. 21 and Eq. 22. Backorders occur at any point in time at which the demand exceeds the available inventory. Under an (*R*, *Q*) policy, item *i*'s expected number of backorders is (see Hopp and Spearman, 2001):

$$\overline{B}_i(R_i, Q_i) = \frac{1}{Q_i}[\beta(R_i) - \beta(R_i + Q_i)] \tag{25}$$

$$\beta(x) = \frac{\sigma^2}{2}\left\{\left(z^2+1\right)\left[1-\Phi(z)\right] - z\phi(z)\right\} \tag{26}$$

$$z = \frac{(x-\theta)}{\sigma} \tag{27}$$

Where θ and σ are the mean and standard deviation of the demand during replenishment lead time, respectively. Equation 26 is the continuous analogue of the second-order loss function $\beta(x)$ (Hopp and Spearman, 2001). The second-order loss function represents the time-weighted backorders arising from lead time demand in excess of *x* (Hopp *et al.*, 1997).

4. Solution Procedure

The above two-echelon, (*R*, *Q*) optimization model is a large scale, non-linear, integer optimization problem. Under the above assumptions, modeling each echelon independent of the other echelons is not attainable due to the dependency between them. In order to model the warehouse, the retailer's order batch size must be known a priori. On the other hand, in order to model a retailer, its effective lead time must be known. The retailer's effective lead time is a function of the warehouse's expected number of backorders, which is function of the warehouse's policy parameters. This indicates that both echelons must be modeled and solved simultaneously. To solve the above two-echelon inventory system, we assumed identical retailers and decomposed the problem into two levels; the retailer (Model 1) and the warehouse (Model 2) as follows:

Model 1: The retailer**:** Since minimizing total inventory investment across the retailers is the same as minimizing the inventory investment at a single retailer under the assumption of identical retailers we formulate the optimization problem at the retailer level as minimizing total inventory investment subject to the order frequency and backorder constraints as follows:

$$Minimize\ \ C_r = \sum_{i=1}^{N} c_i \bar{I}_{ri}\left(R_{ri}, Q_{ri}\right) \tag{28}$$

Subject to

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{ri}}{Q_{ri}} \leq \mathrm{F}_r \tag{29}$$

$$\sum_{i=1}^{N} \bar{B}_{ri}\left(R_{ri}, Q_{ri}\right) \leq \mathrm{B}_r \tag{30}$$

$$R_{ri} \geq -Q_{ri}, i = 1, 2 \ldots N \tag{31}$$

$$Q_{ri} \geq 1, i = 1, 2, \ldots N \tag{32}$$

$$Q_{ri}\ \&\ R_{ri}\ :\text{Integers}, i = 1, 2, \ldots N \tag{33}$$

Model 2: The warehouse: We formulate the optimization problem at the warehouse as minimizing total inventory investment subject to the order frequency and backorder constraints as follows.

$$Minimize\ \ C_w = \sum_{i=1}^{N} c_i Q_{ri} \bar{I}_{wi}\left(R_{wi}, Q_{wi}\right) \tag{34}$$

Subject to

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{wi}}{Q_{wi}} \leq \mathrm{F}_w \tag{35}$$

$$\sum_{i=1}^{N} \bar{B}_{wi}\left(R_{wi}, Q_{wi}\right) \leq \mathrm{B}_w \tag{36}$$

$$R_{wi} \geq -Q_{wi}, i = 1, 2 \ldots N \tag{37}$$

$$Q_{wi} \geq 1, i = 1, 2, \ldots N \tag{38}$$

$$Q_{wi}\ \&\ R_{wi}\ :\text{Integers}, i = 1, 2, \ldots N \tag{39}$$

Decomposition has been used widely in many areas such as inventory management and queuing systems (e.g. Cohen *et al.*, 1990 and Hopp *et al.*, 1999). By treating the echelons one at a time, we use the assumption that the replenishment lead time is constant, that is; there is always an ample supply of parts at the replenishment sources (Cohen *et al*., 1990). Under this assumption, the retailer's effective lead time is equal to its fixed lead time. In other words, the second component that is due to the delay at the warehouse due to a stockout is assumed to be equal to zero. This implies that, the retailers can be modeled independent of the warehouse. This enables us to calculate the warehouse lead time demand which is function of the retailer's replenishment batch size.

By decomposing the system into two levels, the warehouse and the retailers are modeled as different problems. In the case of identical retailers there are only two problems to solve, one for the warehouse (Model 2) and one for the retailers (Model 1). The level-by-level decomposition does not, in general, give truly optimal solutions to the multi-echelon problem (Cohen *et al*., 1990). Therefore, we are seeking procedures that eliminate the effect of decomposing the system by level on the quality of the final solutions. Hence, we are seeking to derive simple formulas that approximate the policy parameters under different assumptions such as fixed and stochastic lead times, and then to develop an optimization algorithm that implements these expressions within the multi-echelon context.

4.1 The Retailer Heuristics

Hopp *et al*. (1997) presented heuristics for approximating policy parameters at a single location that implements an (R, Q) policy under the assumption of fixed lead time and Poisson demands. They approximated the expected number of backorders during lead time using a base stock model. Under the base stock model, the expected number of backorders is only a function of the reorder point which results in simple formulas for the policy parameters as we will see in the next section.

4.1.1 The Retailer Under the Assumption of Fixed Lead Time Heuristic (H1)

The following policy parameters at the retailer under the assumption of fixed lead times are derived as follows (for more details refer to Hopp and Spearman, 2001):

- Assume ample supply at the warehouse, i.e. fixed lead times
- Approximate the expected number of backorders at the retailer using a base stock model
- Assume continuous decision variables
- Move the order frequency and backorder constraints at the retailer into the objective function in model 1
- Derive the resulting version of the Lagrange objective function with respect to Q_{ri} which results in the following expression for the retailers batch size:

$$Q_{ri} = \sqrt{\frac{2\varsigma_r \ddot{e}_{ri}}{Nc_i}}\text{, } i\text{=1, 2 …}N \tag{40}$$

- Derive the resulting version of the Lagrange objective function with respect to R_{ri} which results in the following expression for the retailer's batch size:

$$R_{ri} = \sqrt{\ddot{e}_{ri} L_{ri}}\,\Phi^{-1}\left(1 - \frac{c_i}{(c_i + \kappa_r)}\right) + \ddot{e}_{ri} L_{ri} \tag{41}$$

Eq. 40 and Eq. 41 are simple expressions that approximate the stocking policies at the retailer under the assumption of fixed replenishment lead time. Each one of these expressions is a function of only one Lagrange multiplier. Hopp and Spearman (2001) presented an optimization algorithm to search for these Lagrange multipliers under which the search is guided towards the target order frequency and the backorder values. This is due to the convexity of these constraints. The average on-hand inventory, expected number of backorders, and the average order frequency constraints are convex functions of R and Q (for more details see Zipkin, 2000, page 217). Instead, we derived expression for the Lagrange multiplier ς_r which replaces the first four steps of Hopp and Spearman's optimization algorithm, by substituting Eq. 40 into Eq. 29 after replacing the less or equal sign in Eq. 29 by an equal sign as follows:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{ri}}{\sqrt{\frac{2\varsigma_r \ddot{e}_{ri}}{Nc_i}}} = \mathrm{F}_r \tag{42}$$

Solving Eq. 42 with respect to ς_r results in the following expression:

$$\varsigma_r = \left(\frac{a}{\mathrm{F}_r N}\right)^2 \tag{43}$$

$$a = \sum_{i=1}^{N}\frac{\lambda_{ri}}{\sqrt{\frac{2\ddot{e}_{ri}}{Nc_i}}} \tag{44}$$

Unfortunately, the backorder constraint, Eq. 30, is too complicated to be solved in exact form for κ_r. The bisection technique is used to search for the Lagrange multiplier (κ_r) that results in a reorder point, as given by Eq. 41, that satisfies the following backorder constraint:

$$\sum_{i=1}^{N}\overline{B}_{ri}\left(R_{ri}, Q_{ri}\right) = \mathrm{B}_r \tag{45}$$

4.1.2 The Retailer Under the Assumption of Stochastic Lead Time Heuristic (H2)

In order to model the effect of the delay at the warehouse due to stockout we relax the assumption of fixed lead time at the retailer by assuming limited supply at the warehouse. In order to derive simple expressions for the policy parameters at the retailer under the assumption of limited supply at the warehouse we assume that the policy parameters at the warehouse are known a priori and approximate the expected number of backorders at the retailer using the base stock model. Also, we relax the assumption of integer decision variables to allow for continuous decision variables. After incorporating these assumptions, we moved the retailer's average order frequency and the expected number of backorder constraints into the objective function in Model 1 using the theory of Lagrange multipliers which results in the following Lagrange version of Model 1's objective function:

$$
\begin{aligned}
Min\ L_r = &\sum_{i=1}^{N} c_i\left(\overline{B}_{ri}\left(R_{ri}\right) + R_{ri} + \frac{Q_{ri}+1}{2} - \ddot{e}_{ri}\left(L_{ri} + \frac{Q_{ri}\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right)}{m\ddot{e}_{ri}}\right)\right) + \\
&\alpha_r\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\ddot{e}_{ri}}{Q_{ri}} - F_r\right) + \delta_r\left(\sum_{i=1}^{N}\overline{B}_{ri}\left(R_{ri}\right) - B_r\right)
\end{aligned}
\tag{46}
$$

Taking the partial derivative of Eq. 46 first with respect to (Q_{ri}) and then with respect to (R_{ri}) results in the following simple expressions for the policy parameters at the retailer:

$$
Q_{ri} = \begin{cases} \sqrt{\dfrac{\alpha_r \ddot{e}_{ri}}{N\left(\dfrac{c_i}{2} - \dfrac{\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right)}{m}\right)}}, & \text{if } \dfrac{c_i}{2} - \dfrac{\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right)}{m} > 0.0 \\ \sqrt{\dfrac{\alpha_r \ddot{e}_{ri}}{N}}, & \text{Otherwise} \end{cases} \quad , i = 1, 2 \ldots N \tag{47}
$$

$$
R_{ri} = \sqrt{\ddot{e}_{ri}\ell_{ri}}\,\Phi^{-1}\left(1 - \frac{c_i}{\left(c_i + \delta_r\right)}\right) + \ddot{e}_{ri}\ell_{ri} \quad i = 1, 2 \ldots N \tag{48}
$$

We derived the following expression for the Lagrange multiplier (α_r) that appears in Eq. 47 by substituting Eq. 47 into Eq. 29 after replacing the less or equal sign in Eq. 29 by an equal sign and solving the resulting expression with respect to α_r:

$$\alpha_r = \left(\frac{b}{\mathrm{F}_r N} \right)^2 \tag{49}$$

$$b = \begin{cases} \displaystyle\sum_{i=1}^{N} \frac{\lambda_{ri}}{\sqrt{\dfrac{\ddot{e}_{ri}}{N\left(\dfrac{c_i}{2} - \dfrac{\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right)}{m} \right)}}}, & \text{if } \dfrac{c_i}{2} - \dfrac{\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right)}{m} > 0.0 \\ \displaystyle\sum_{i=1}^{N} \frac{\lambda_{ri}}{\sqrt{\dfrac{\ddot{e}_{ri}}{N}}}, & \text{Otherwise.} \end{cases} \tag{50}$$

Finally, the bisection technique can be used to search for the Lagrange multiplier (δ_r) that appears in Eq. 48 such that it results in a reorder point, as given by Eq. 48, that satisfies the backorder constraint, as given by Eq. 45.

4.2 The Warehouse Heuristic (H3)

Under a two-echelon (*R*, *Q*) inventory system, the demand process at the warehouse is a superposition of the retailer's ordering processes. Hence, in order to approximate the expected demand at the warehouse, the retailer replenishment order size (Q_{ri}) for each item must be known a priori. In order to derive simple expressions for the policy parameters at the warehouse we assumed that the policy parameters at the retailers are known a priori and approximated the expected number of backorders at the warehouse using a base stock model. Also, we relaxed the assumption of integer decision variables to allow for continuous decision variables. After incorporating these assumptions we moved the warehouse average order frequency and the expected number of backorder constraints into the objective function in Model 2 using the theory of Lagrange multipliers which results in the following Lagrange version of Model 2's objective function:

$$Min.\ \ L_w = \sum_{i=1}^{N} c_i Q_{ri}\left(\overline{B}_{wi}\left(R_{wi}\right) + R_{wi} + \frac{Q_{wi}+1}{2} - \frac{\ddot{e}_{ri}L_{wi}m}{Q_{ri}}\right) + \varsigma_w\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\ddot{e}_{wi}}{Q_{wi}} - F_w\right) + \hat{e}_w\left(\sum_{i=1}^{N}\overline{B}_{wi}\left(R_{wi}\right) - B_w\right) \tag{51}$$

Taking the partial derivative of Eq. 51 first with respect to (Q_{wi}) and then with respect to (R_{wi}) results in the following simple expressions for the policy parameters at the warehouse:

$$Q_{wi} = \sqrt{\frac{2m\varsigma_w \ddot{e}_{ri}}{Nc_i Q_{ri}^2}},\ i = 1, 2 \ldots N \tag{52}$$

$$R_{wi} = \sqrt{V[D_{wi}]}\Phi^{-1}\left(1 - \frac{c_i Q_{ri}}{\left(c_i Q_{ri} + \kappa_w\right)}\right) + \frac{\ddot{e}_{ri}L_{wi}m}{Q_{ri}},\ i = 1, 2 \ldots N \tag{53}$$

We derived the following expression for the Lagrange multiplier (ς_w) that appears in Eq. 52 by substituting Eq. 52 into Eq. 35 after replacing the less or equal sign in Eq. 35 by an equal sign and solving the resulting expression with respect to ς_w:

$$\varsigma_w = \left(\frac{c}{\mathrm{F}_w N}\right)^2 \tag{54}$$

Where,

$$c = \sum_{i=1}^{N}\frac{\lambda_{wi}}{\sqrt{\frac{2m\ddot{e}_{ri}}{Nc_i Q_{ri}^2}}} \tag{55}$$

Finally, the bisection technique can be used to search for the Lagrange multiplier (κ_w) that appears in Eq. 53 such that it results in a reorder point, as given by Eq. 53, that satisfies the following backorder constraint:

$$\sum_{i=1}^{N}\overline{B}_{wi}\left(R_{wi}, Q_{wi}\right) = \mathrm{B}_w \tag{56}$$

4.3 Two-Echelon (*R, Q*) Optimization Algorithm

The above heuristics are based on modeling each echelon by assuming that the policy parameters at the other echelons are known a priori. In the above problem, neither the warehouse's nor the retailer's policy parameters are known. These policy parameters are decision variables to be determined by the optimization model. Hence, the above heuristics can not be used independently to set the policy parameters for the system under consideration. Heuristics H2 and H3 can not be used to solve the problem directly without first knowing the warehouse's and the retailer's stocking policies, respectively. Also, the use of H1 in conjunction with H3 will not incorporate the effect of the delay at the warehouse due to stockout. Hence, in order to arrive at an approximate solution for the stocking policy parameters, we developed and implemented the above heuristics in the following iterative heuristic optimization algorithm (*IHOA*):

<u>Algorithm IHOA:</u>

Step 1. Set $\ell_{ri} = L_{ri}$, $i = 1, 2 \ldots N$.

Step 2. Model the retailer:

1. Calculate ς_r using Eq. 43.
2. Calculate Q_{ri} for each item using Eq. 40.
3. Use the bisection technique to search for the Lagrange multiplier (κ_r) that appears in Eq. 41 such that it results in a reorder point, as given by Eq. 41, that satisfies the expected number of backorder constraint at the retailer, as given by Eq. 45.

Step 3. Model the warehouse:

1. Calculate the expected lead time demand at the warehouse using Eq. 21.

2. Calculate ς_w using Eq. 54.

3. Calculate Q_{wi} for each item using Eq. 52.

4. Use the bisection technique to search for the Lagrange multiplier (κ_w) that appears in Eq. 53 such that it results in a reorder point, as given by Eq. 53, that satisfies the expected number of backorder constraint at the warehouse, as given by Eq. 56.

Step 4. Calculate the expected number of backorders at the warehouse using Eq. 25.

Step 5. Calculate the retailer effective lead time using Eq. 17.

Step 6. Refine the policy parameters at the retailer:

1. Calculate α_r using Eq. 49.

2. Calculate Q_{ri} for each item using Eq. 47.

3. Use the bisection technique to search for the Lagrange multiplier (δ_r) that appears in Eq. 48 such that it results in a reorder point, as given by Eq. 48, that satisfies the expected number of backorder constraint at the retailer, as given by Eq. 45.

Step 7. If $\left|Q_{ri}^{c} - Q_{ri}^{p}\right| \le e, \quad i = 1,...,N$

$\left|R_{ri}^{c} - R_{ri}^{p}\right| \le e, \quad i = 1,...,N$

$\left|Q_{wi}^{c} - Q_{wi}^{p}\right| \le e, \quad i = 1,...,N$

$\left|R_{wi}^{c} - R_{wi}^{p}\right| \le e, \quad i = 1,...,N$

Stop

Else, Go to Step 3

5. Experimentation and Analysis

In order to assess the quality of the solutions obtained via the above heuristic optimization algorithm we compared the solutions obtained using *Algorithm IHOA* with the solutions obtained using *OptQuest* for Java search engine. After testing the solutions obtained using *Algorithm IHOA* for a small set of problems with the solutions obtained using *OptQuest*, *Algorithm IHOA* is used to set the inventory policy parameters for large scale inventory systems. Within these experiments, we monitored the associated computation times and the percentage differences in the estimated inventory investment. For the sake of experimentation, we set the following target values of the order frequency and the expected number of backorder constraints at the retailer and the warehouse ($\mathrm{F}_r = 24$, $\mathrm{F}_w = 12$, $\mathrm{B}_r = 1.0 \times N$, $\mathrm{B}_w = 0.2 \times N$). Also, we set the number of retailers equals to four and the tolerance value *e* equal to 0.01. *Algorithm IHOA*, the bisection technique, the inventory policy parameters and the Lagrange expressions, and the above inventory model were coded in the Java programming language. The following experiments were executed on a Pentium 4 computer with a 3.06 GH processor and 512 Cache memory.

A meta-heuristic is a family of optimization approaches that includes scatter search, genetic algorithms, simulated annealing, Tabu search, etc. and their hybrids. The *OptQuest* engine combines Tabu search, scatter search, integer programming, and neural networks into a single, composite search algorithm. For more details about *OptQuest*, we refer the reader to Rogers (2002).

5.1. *Algorithm IHOA* versus OptQuest

Algorithm IHOA takes advantage of the structure of the problem under which the search is guided towards the target values of the average order frequency and the expected number of backorder constraints. *Algorithm IHOA* requires no bounds on the decision variables and does not require any stopping criteria except for the tolerances associated with the bisection search technique. On the other hand, the *OptQuest* search engine requires the user to set lower and upper bounds on the decision variables and to specify at least one stopping criterion. The number of iterations and/or the optimization times can be used as the stopping criteria in *OptQuest*. The quality of the solutions obtained using *OptQuest* depends heavily on the decision variable lower and upper bounds, number of decision variables, and the stopping criteria. Since we do not know the regions where the optimal solutions might be, *OptQuest* might not be able to find any feasible solutions at all if the specified solution space does not contain any feasible solutions. Therefore, we must supply *OptQuest* with the proper lower and upper bounds in order to arrive at acceptable solutions. Hence, we set the policy parameters using *Algorithm IHOA* and then we set the lower and upper bounds around these estimated solutions to be used as the bounds on the decision variables in *OptQuest*. As we can see *OptQuest* relies on *Algorithm IHOA* to specify the decision variable's lower and upper bounds. Therefore, completely independent comparison between the two methods is not attainable since we do not have an idea about the regions where the optimal solutions or near optimal solutions might be before using *Algorithm IHOA*.

In order to arrive at a reasonable comparison between *Algorithm IHOA* and *OptQuest*, we set the time in *OptQuest* as the optimization stopping criterion. We ran *Algorithm IHOA* and recorded the associated optimization times. Then, we set lower and upper bounds on the estimated solutions. Next, we set the time in *OptQuest* equal to the *Algorithm IHOA* optimization times. Finally, we ran *OptQuest* for the specified times and recorded the solutions found and the number of iterations executed within that time. Since the quality of the solutions obtained using *OptQuest* deteriorates with an increase in the number of decision variables, we limit our initial experiments to systems that consists of a maximum of 25 items. Table 1 shows the systems under consideration.

TABLE 1 ABOUT HERE

The policy parameters for the above six systems were estimated using *Algorithm IHOA* and *OptQuest* where the time is used as the stopping criterion. Table 2 shows the results of these experiments, where the inventory investment and the percentage differences in inventory investment between *Algorithm IHOA* and *OptQuest* are recorded. As we can see from Table 2, *Algorithm IHOA* optimization times are less than one second. During these experiments, the *OptQuest* engine could not find any feasible solutions in times less than a second. Hence, we set the time in *OptQuest* equal to one second.

TABLE 2 ABOUT HERE

The percentage difference is calculated using the following formula:

$$\%\text{Difference} = \frac{\text{InventoryInvestment}_{Algorithm\ IHOA} - \text{InventoryInvestment}_{OptQuest}}{\text{InventoryInvestment}_{Algorithm\ IHOA}} \times 100\% \qquad (57)$$

Table 2 shows that the percentage differences in the inventory investment between *Algorithm IHOA* and *OptQuest* are high in most of the cases and they do not follow any regular pattern. We concluded that one second of optimization time is not sufficient to arrive at acceptable results using *OptQuest*. In order to obtain better solutions using *OptQuest*, the number of iterations or the optimization times must be increased. With the increase in the number of iterations the computational times increase in *OptQuest*. *OptQuest* relies on comparing any new feasible solution with the old feasible solutions stored in its database. Hence, the database size increases with the number of iterations which results in an increase in each iteration's associated computation times.

The above experiments were repeated in *OptQuest* where the number of iterations is set fixed at 40,000 iterations. Table 3 shows the inventory investment obtained using *Algorithm IHOA* and *OptQuest* (at 40,000 iterations) and the percentage difference between them. *OptQuest* engine found better solutions for the single item system compared to *Algorithm IHOA.* The inventory investment obtained using *OptQuest* for the single item system is less than the inventory investment obtained using *Algorithm IHOA* by less than 4%. On the other hand, *Algorithm IHOA* found better solutions than *OptQuest* for the rest of the systems.

TABLE 3 ABOUT HERE

OptQuest computational times increase dramatically with the increase in the number of items, when the number of iterations is fixed. This is because the solution space and the time required to execute each iteration increases with the increase in the number of items. Hence, *OptQuest* managed to find better solutions for the small systems in shorter time than for the larger systems.

Figure 2 is a plot of the inventory investment versus the number of items shown in Table 3. From Figure 2, it is clear that the quality of the solutions obtained using *OptQuest* when fixing the number of iterations deteriorates with the increase in the number of items when compared to *Algorithm IHOA.*

FIGURE 2 ABOUT HERE

5.2. Experiments on Large Scale Systems

Now, we apply *Algorithm IHOA* on large scale inventory systems where the computation times and the average inventory investment are recorded. Two factors are varied over the experiments, the number of items and the effect of the delay at the warehouse due to stockout. Table 4 shows the large scale inventory systems under consideration:

TABLE 4 ABOUT HERE

Decomposing the system by echelon eliminates the effect of the delay at the warehouse due to stockout on the retailer's policy parameters. In order to test the effect of the delay at the warehouse we modeled the above systems under the assumption of fixed effective lead times at the retailers and the warehouse where only Steps 1-3 of *Algorithm IHOA* are executed. Also, we modeled the above systems using *Algorithm IHOA* where the effect of the delay at the warehouse is incorporated. Table 5 shows the results of these experiments.

TABLE 5 ABOUT HERE

As we can see from Table 5, *Algorithm IHOA* managed to set the inventory policy parameters for a 40,000 item system in 26.14 seconds. *Algorithm IHOA* is a fast optimization algorithm that can handle large systems in negligible times. We report in Table 5 the average inventory investment for the above six systems when the delay at the warehouse is ignored and when it is incorporated. The percentage differences is calculated similar to Eq. 57 except that we replaced the term that represents the inventory investment obtained using *OptQuest* by the inventory investment obtained using Steps 1-3 of *Algorithm IHOA*. The average percentage difference in the inventory investment is 0.5325%, the minimum is 0.5236%, and the maximum is 0.5567%. The effect of increasing the number of items on the percentage difference is almost negligible. Incorporating the effect of the delay at the warehouse due to stockout increases the inventory investment on average by 0.5325%. This result is natural since the retailer's effective lead time is expected to increase when modeling the delay at the warehouse. Hence, the inventory levels at the retailer are expected to increase with the increase in its effective lead times. So, *Algorithm IHOA* results in more accurate results than in the case of fixed retailer effective lead times where the delay at the warehouse is ignored. A 0.5325% increase on average in the inventory investment might be significant in large scale systems as we can see from Table 5. Under a 40,000 item system, the difference in the inventory investment is more than 13 million dollars.

5.3. Simulation Analysis

In this section, the quality of the solutions obtained via *Algorithm IHOA* will be tested against a simulation optimization model where the objective functions and performance measures are evaluated using a simulation model. The motivation behind this simulation investigation is to compare the solution of *Algorithm IHOA* which is based on using analytical inventory models to approximate the objective function and performance measures with the solution of *OptQuest* where it uses a simulation model to estimate the objective function and performance measures values. In Section 5.1, we compared *Algorithm IHOA* and *OptQuest* where both of them used the analytical inventory model to estimate the objective function and performance measures. Since in both algorithms the objective function and performance measures are evaluated using the same analytical model, both solutions are feasible and satisfy the model constraints.

Due to the complexity of the problem and its mathematical modeling assumptions, we should expect the values of the objective function and performance measures of the analytical model to be different than the values obtained using a simulation model. The analytical formulation and solution procedure must make assumptions and approximations that the simulation model does not have to make in order to estimate the performance measures. Thus, there will be natural differences between analytical performance and simulated performance. For more on this issue, the interested reader is referred to Tee and Rossetti (2002). Also, it is important to point out that simulation is a statistical experiment and contains sampling error. This comparison is meant to provide an insight into how the analytical formulation approximates the underlying problem with the caveat that we know it is only an approximation. Three test cases with number of items, unit cost, demand rate, and lead times as shown in Table 6 are considered in this section.

Tee and Rossetti (2002) in an extensive simulation study for multi-echelon inventory systems developed a simulation model for a single item two-echelon (R, Q) inventory system where they studied the robustness of two-echelon (R, Q) analytical inventory models developed by Deuermeyer and Schwarz (1981), Svoronos and Zipkin (1988), and Axsäter (2000). The main objective of their study was to examine the analytical models via simulation when the model's basic assumptions are violated.

They built the simulation model in Arena 5.0 Simulation language. In this simulation optimization study, we rebuilt the simulation model developed by Tee and Rossetti (2002) in Arena 9.0. Also, since we are considering multi-item systems, we extended the simulation model for the multi-item case. For more details about the simulation model development and simulation study we refer the reader to Tee and Rossetti (2002). An optimization model was built for each of the above three test cases in *OptQuest for Arena* where *OptQuest* used the Arena simulation model to evaluate the objective function and performance measures.

TABLE 6 ABOUT HERE

Because simulation optimization is so computationally intensive, we used the optimal policy parameters suggested by *Algorithm IHOA* to initialize the search. The range of each decision variable was specified around the initial starting values as was done in Section 5.1. The simulation optimization model was allowed to run for 20,000 iterations where the total simulation optimization time, inventory investment, policy parameters, expected number of backorders, and average order frequency were recorded as shown in Tables 8-13.

Table 7 summaries the results of the experiments. As shown in Table 7, *Algorithm IHOA* underestimated the simulated inventory investment for the last two cases and overestimated it for the first case. Table 7 shows that the simulation optimization times are high compared to the optimization time of *Algorithm IHOA* and increase as the number of items increases.

TABLE 7 ABOUT HERE

Tables 8, 10, and 12 show the solutions obtained using *Algorithm IHOA*. For the three test cases, the objective functions and performance measures are evaluated using the analytical and simulation models based on the policy parameters obtained using *Algorithm IHOA.* Tables 8, 10, and 12 shows that the solutions obtained using *Algorithm IHOA* are feasible for the analytical model and satisfy all the constraints. On the other hand, for all the three test cases the solutions of *Algorithm IHOA* are not feasible based on the simulated results. Again, we caution the reader to understand that this is a little like comparing “apples” and “oranges” since the simulation can capture complexity that the analytical model cannot capture.

TABLE 8 ABOUT HERE

Tables 9, 11, and 13 show the solutions obtained using the simulation optimization model. For the three test cases the objective functions and performance measures are evaluated using the analytical and simulation models based on the policy parameters obtained using the simulation optimization model. Tables 9, 11, and 13 shows that the solutions obtained using the simulation optimization model are feasible for the simulation model and satisfy all the constraints. On the other hand, for all the three test cases the solutions of the simulation optimization model are not feasible for the analytical model.

TABLES 9-13 ABOUT HERE

Tables 8-13 (last row of each table) show that, the inventory investment and performance measures of the analytical model are off the true values for the same policy parameter for all the test cases under consideration. As we can see, overestimating or underestimating the inventory performance measures impacts directly the behavior of the optimization algorithm. Naturally, if the approximations are more accurate, the resulting policy parameter values should better reflect results from the simulation model.

Based on our knowledge, we have not seen any analytical inventory models that correctly estimate all of the inventory performance measures for this problem context over a wide range of conditions and values for the control variables. As expected, this simulation study shows that the results from our algorithm are different from the results of a simulation optimization approach. However, the results are close enough to show that *Algorithm IHOA* has clear potential for setting reasonably good policy parameter values for *large scale* problems. Ultimately, that was our goal. It should be clear that the simulation optimization approach is impractical for any realistically sized problems and in that context *Algorithm IHOA* is clearly a good alternative. These results also show that future work is still needed to get better approximations of system performance for large scale problems. A simulation study similar to that done by Tee and Rossetti (2002) should be considered for this multi-echelon case to better understand where the approximations begin to break down.

6. Conclusions and Future Work

We modeled a two-echelon inventory system that implements (R, Q) policies at each facility. In order to solve the two-echelon inventory system we decomposed it by echelon. We derived expressions for the inventory policy parameters at each facility under different assumptions and expressions for the Lagrange multipliers that appear in the replenishment batch size expressions. We developed an efficient multi-echelon optimization algorithm that implements these expressions. Our experiments showed that *Algorithm IHOA* is an efficient algorithm that can set the inventory policy parameters for a two-echelon inventory system in negligible times. *Algorithm IHOA* managed to set the policy parameters for a 40,000 items inventory system in 26.14 seconds. *Algorithm IHOA* is more efficient than *OptQuest* for Java from the computation times and quality of the solutions points of view in most of the cases examined for small inventory systems. The effect of the delay at the warehouse due to stock out has a significant impact on the inventory investment when modeling large scale systems. The average percentage difference in the inventory investment due to incorporating the effect of the delay at the warehouse due to stockout is 0.5325%. The percentage differences in the inventory investment do not vary significantly with the number of items. We also showed via simulation that additional research is needed to better approximate the system performance measures of this system. Improved approximations would have a definite effect on the quality of the results from IHOA. This paper considered only the identical retailer case; however, future work is under way to consider the non-identical retailer case..

Acknowledgment

We would like to gratefully acknowledge the support of the Naval Systems Supply Command through the University of Arkansas' Center for Engineering Logistics and Distribution (CELDi). This material is based upon work supported by the National Science Foundation under Grant No. 0214478 in cooperation with the Naval Systems Supply Command. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the U.S. Navy.

7. References

1. Albin, S. L., February 1982. On Poisson approximation for superposition arrival processes in queues, *Management Science*, 28(2), 126-137.
2. Al-Rifai, M. H.; Rossetti, M. D.; and Desai, V. L., 2005. An iterative heuristic optimization model for multi-echelon (R, Q) inventory systems, *Industrial Engineering Research Conference (IERC)*, Atlanta, Georgia.
3. Anderson, J. and Marklund, Johan, 2000. Decentralized inventory control in a two-level distribution system, *European Journal of Operational Research*, 127, 483-506.
4. Axsäter, S., January-February 1990. Simple solution procedures for a class of two-echelon inventory problems, *Operations Research*, 38(1), 64-69.
5. Axsäter, S., July-August 1993. Exact and approximate evaluation of batch-ordering policies for two-level inventory systems, *Operations Research*, 41(4), 777-785.
6. Axsäter, S., 1997. Simple evaluation of echelon stock (R, Q) policies for two-level inventory systems, *IIE Transactions*, 2, 661-669.

7. Axsäter, S., 1998. Evaluation of installation stock based (R, Q) policies for two-level inventory systems with Poisson demand, *Operations Research*, 46(3), S135-S145.

8. Axsäter, S., September-October 2000. Exact analysis of continuous review (R, Q) policies in two-echelon inventory systems with compound Poisson demand, *Operations Research*, 48(5), 686-696.

9. Axsäter, S., 2001. Scaling down multi-echelon inventory problems, *International Journal of Production Economics*, 71, 255-261.

10. Caglar, D.; Li, C-L; and Simchi-Levi, D., 2004. Two-echelon spare parts inventory system subject to a service constraint, *IIE Transactions*, 36, 655-666.

11. Cohen, M. A.; Kamesam, P. V.; Kleindorfer, P.; Lee, H.; and Tekerian, A., January-February 1990. Optimizer: IBM's multi-echelon inventory system for managing service logistics, *Interfaces*, 20(1), 65-82.

12. Cohen, M. A. and Kleindorfer, P. R., January-February 1989. Near-optimal service constrained stocking policies for spare parts, *Operations Research*, 37(1), 104-117.

13. Deshpande, V.; Cohen, M.A.; and Donohue, K., 2003. An empirical study of service differentiation for weapon system service parts, *Operations Research*, 51(4), 518-530.

14. Deuermeyer, B. L. and Schwarz, L. B., 1981. A model for the analysis of system service level in warehouse-retailer distribution systems: the identical retailer case, *TIMS Studies in the Management Sciences*, 16, 163-193.

15. Dìaz, A. and Michael, C. F., 1997. Models for Multi-echelon repairable item inventory systems with limited repair capacity, *European Journal of Operational Research*, 97, 480-492.

16. Graves, S. C., 1985. A multi-echelon inventory model for a repairable item with one-for-one

replenishment, *Management Science*, 31(10), 1247-1256.

17. Hadley, G. and Whitin, T. M., 1963. *Analysis of inventory systems*. Prentice-Hall, Inc.
18. Hopp, W. J.; Spearman, M. L.; and Zhang, R. Q., May-June 1997. Easily implementable inventory control policies, *Operations Research*, 45(3), 327-340.
19. Hopp, W. J.; and Spearman, M. L., 2001. *Factory Physics*, 2nd Edition. McGraw-Hill, New-York.
20. Hopp, W. J.; Zhang, R. Q.; and Spearman, M. L., 1999. An easily implementable hierarchical heuristic for a two-echelon spare parts distribution system, *IIE Transactions*, 31, 977-988.
21. Muckstadt, J. A., December 1973. A model for a multi-item, multi-echelon, multi-indenture inventory system, *Management Science*, 20(4), 472-481.
22. Rogers, P., 2002. "Optimum-seeking Simulation in the Design and Control of Manufacturing Systems: Experience with *OptQuest* for Arena", in the *Proceedings of the 2002 Winter Simulation Conference*, E. Yucesan, *et al.*, Ed.
23. Sherbrooke, C. C., 1968. METRIC: A multi-echelon technique for recoverable item control, *Operations Research*, 16, 122-141.
24. Silver, E. A.; Pyke, D. F.; and Peterson, R., 1998. *Inventory management and production planning and scheduling*, 3rd edition, John Wiley & Sons, Inc.
25. Svoronos, A and Zipkin, P., 1988. Estimating the performance of multi-level inventory systems, *Operations Research*, 36(1), 57-72.
26. Tee, Yeu-San and Rossetti, M. D., 2002. A robustness study of a multi-echelon inventory model via simulation, *International Journal of Production Economics*, 80(3), 265-277.
27. Tee, Yeu-San and Rossetti, M. D., November 2001. Using simulation to evaluate a continuous review (*R*, *Q*) two-echelon inventory model, *Proceedings of the 6th Annual*

International Conference on Industrial Engineering-Theory, Applications, and Practice, San Francisco, CA, USA, 18-20.

28. Torab, P. and Kamen, E., 2001. On approximate renewal models for the superposition of renewal processes, *IEEE*, 2901-2906.
29. Zipkin, P. H., 2002. *Foundations of Inventory Management*, McGraw-Hill Companions, Inc.

FIGURES

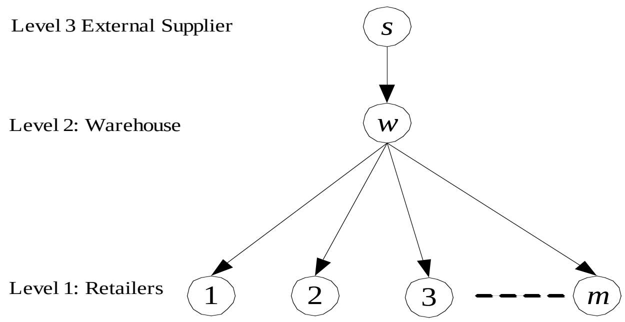

Figure 1: A typical multi-echelon inventory system

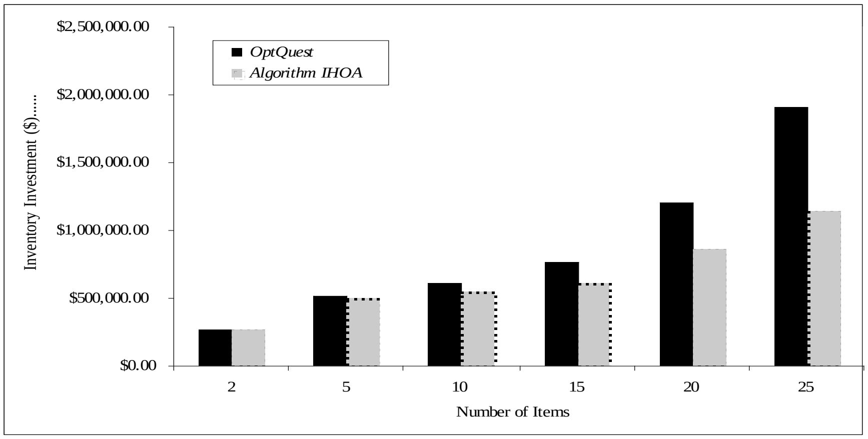

Figure 2: Inventory investment versus the number of items

TABLES

Table 1: Number of items per system

System	1	2	3	4	5	6
Number of items	2	5	10	15	20	25

Table 2: *Algorithm IHOA* versus *OptQuest*, *OptQuest* optimization time = 1 second

System	Number of Items	*Algorithm IHOA* Inventory Investment ($)	*Algorithm IHOA* Optimization Times (seconds)	*OptQuest* Inventory Investment ($)	*OptQuest* Optimization Times (seconds)	*OptQuest* Number of Iterations	Inventory Investment Difference ($)	Inventory Investment % Difference
1	**2**	$273,720.06	0.266	$342,409.77	1.000	3291	-$68,689.71	-25.09%
2	**5**	$500,344.88	0.235	$2,585,779.65	1.000	1616	-$2,085,434.77	-416.80%
3	**10**	$546,758.88	0.281	$1,802,175.16	1.000	870	-$1,255,416.28	-229.61%
4	**15**	$613,069.63	0.297	$5,815,571.80	1.000	546	-$5,202,502.17	-848.60%
5	**20**	$865,165.48	0.297	$20,629,153.19	1.000	290	-$19,763,987.71	-2284.42%
6	**25**	$1,147,666.65	0.296	$9,682,534.74	1.000	207	-$8,534,868.09	-743.67%

Table 3: Inventory investment, *Algorithm IHOA* versus *OptQuest* (40,000 iterations)

System	Number of Items	*Algorithm IHOA*		*OptQuest*			Inventory Investment Difference ($)	Inventory Investment % Difference
		Inventory Investment ($)	Optimization Times (seconds)	Inventory Investment ($)	Optimization Times (seconds)	Number of Iterations		
1	**2**	$273,720.06	0.266	$263,487.72	60	40,000	$10,232	3.74%
2	**5**	$500,344.88	0.235	$517,651.44	289	40,000	-$17,307	-3.46%
3	**10**	$546,758.88	0.281	$612,695.26	712	40,000	-$65,936	-12.06%
4	**15**	$613,069.63	0.297	$770,719.29	1143	40,000	-$157,650	-25.71%
5	**20**	$865,165.48	0.297	$1,213,210.21	1603	40,000	-$348,045	-40.23%
6	**25**	$1,147,666.65	0.296	$1,916,849.33	2089	40,000	-$769,183	-67.02%

Table 4: Large scale systems: Number of items per system

System	**1**	**2**	**3**	**4**	**5**	**6**
Number of Items	100	1000	5000	10000	20000	40000

Table 5: Inventory Investment versus *N* and the Effect of the Delay at the Warehouse

System	Number of Items	*Algorithm IHOA*		Steps 1-3 of *Algorithm IHOA*		Effect of delay at the warehouse	
		Inventory Investment ($)	Optimization Times (seconds)	Inventory Investment ($)	Optimization Times (seconds)	Inventory Investment Difference ($)	Inventory Investment % Difference
1	100	$5,148,283.88	0.69	$5,119,623.34	0.22	$28,660.54	0.5567%
2	1000	$58,084,336.82	0.77	$57,780,234.72	0.59	$304,102.09	0.5236%
3	5000	$303,844,521.66	2.98	$302,241,641.00	2.13	$1,602,880.66	0.5275%
4	10000	$611,181,403.94	5.84	$607,943,287.54	4.02	$3,238,116.40	0.5298%
5	20000	$1,237,651,996.22	11.42	$1,231,111,661.07	7.72	$6,540,335.15	0.5284%
6	40000	$2,471,328,595.61	26.14	$2,458,262,719.19	15.19	$13,065,876.42	0.5287%

Table 6: Data Set for 2-Item, 4-Item, and 8-Item Test Cases

Case	Item	Unit Cost ($)	Demand Rate (Units/Year)	*Lri* (Days)	*Lwi* (Days)
1	**1**	901.00	114.00	4.28	4.94
	2	3897.00	60.00	29.00	4.62
2	**1**	459.00	45.00	4.55	27.82
	2	7622.00	92.00	21.46	29.39
	3	722.00	431.00	24.31	4.47
	4	624.00	98.00	21.06	4.59
3	**1**	3633.00	227.00	4.06	21.31
	2	5923.00	98.00	28.42	4.13
	3	2026.00	97.00	4.47	21.88
	4	2629.00	365.00	4.89	4.84
	5	7699.00	39.00	4.25	4.02
	6	413.00	150.00	27.23	21.63
	7	2186.00	32.00	27.96	24.49
	8	1761.00	69.00	29.54	4.88

Table 7: Inventory Investment % Error, *Algorithm IHOA* vs. Simulation Optimization

Case	Items	*Algorithm IHOA*			*Simulation Optimization*			Inventory Investment % error
		Inventory Investment ($)	Time (Seconds)	No. of Runs	Inventory Investment ($)	Time (Seconds)	No. of Runs	
1	2	$67,226.73	<1.0	1	$55,089.41	37,726	20,000	22.03%
2	4	$179,897.74	<1.0	1	$254,248.00	39,740	20,000	-29.24%
3	8	$482,089.00	<1.0	1	$547,874.10	75,180	20,000	-12.01%

Table 8: 2-Items, *Algorithm IHOA*, Time < 1 Second

Item	*Algorithm IHOA*				Performance Measures: Analytical Model				Performance Measures: Simulation Model			
	Qri (Units)	*Rri* (Units)	*Qwi* (Units)	*Rwi* (Units)	F*ri* (Order/year)	F*wi* (Order/year)	B*ri* (Units)	B*wi* (Units of Qri)	F*ri* (Order/year)	F*wi* (Order/year)	B*ri* (Units)	B*wi* (Units of Qri)
1	5.958	1.157	47.668	-1.529	19.133	9.566	0.107	0.152	19.156	2.389	0.807	1.130
2	2.078	2.304	16.628	-0.511	28.867	14.434	1.893	0.248	28.554	3.570	3.050	1.046
		Constraints			F*r*	F*w*	B*r*	B*w*	F*r*	F*w*	B*r*	B*w*
		Estimated			24.000	12.000	2.000	0.400	23.855	2.979	3.857	2.176
		Target			24.000	12.000	2.000	0.400	24.000	12.000	2.000	0.400
		Constraint Satisfied			Yes	Yes	Yes	Yes	Yes	Yes	No	No
		Inventory Investment			$67,226.727				$67,042.030			

Table 9: 2-Items, Simulation Optimization, Time = 37,726 Seconds, 20000 Iterations

Item	***Simulation Optimization***				**Performance Measures: Analytical Model**				**Performance Measures: Simulation Model**			
	***Qri* (Units)**	***Rri* (Units)**	***Qwi* (Units)**	***Rwi* (Units)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**
1	4.237	0.356	42.379	0.000	26.905	10.760	0.314	0.149	26.952	2.683	0.500	0.155
2	3.000	2.960	11.000	1.976	20.000	21.818	1.146	0.097	20.111	5.492	1.500	0.209
		Constraints			**F*r***	**F*w***	**B*r***	**B*w***	**F*r***	**F*w***	**B*r***	**B*w***
		Estimated			23.453	16.289	1.459	0.245	23.532	4.087	2.000	0.365
		Target			24.000	12.000	2.000	0.400	24.000	12.000	2.000	0.400
		Constraint Satisfied			Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
		Inventory Investment			$66,484.695				$55,089.413			

Table 10: 4-Items, *Algorithm IHOA*, Time < 1 Second

Item	***Algorihtm IHOA***				**Performance Measures: Analytical Model**				**Performance Measures: Simulation Model**			
	***Qri* (Units)**	***Rri* (Units)**	***Qwi* (Units)**	***Rwi* (Units)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**
1	5.826	0.595	46.607	15.676	7.724	3.862	0.026	0.024	7.746	0.967	0.065	0.000
2	2.044	1.967	16.354	25.219	45.005	22.502	2.941	0.706	45.162	5.643	3.022	0.000
3	14.376	27.370	115.008	14.245	29.981	14.990	0.773	0.053	29.870	3.735	3.359	4.582
4	7.374	5.224	58.989	4.710	13.291	6.645	0.260	0.017	13.279	1.664	0.873	1.011
		Constraints			**F*r***	**F*w***	**B*r***	**B*w***	**F*r***	**F*w***	**B*r***	**B*w***
		Estimated			24.000	12.000	4.000	0.800	24.014	3.002	7.319	5.592
		Target			24.000	12.000	4.000	0.800	24.000	12.000	4.000	0.800
		Constraint Satisfied			Yes	Yes	Yes	Yes	No	Yes	No	No
		Inventory Investment			$179,897.740				$327,696.910			

Table 11: 4-Items, Simulation Optimization, Time = 39,740 Seconds, 20000 Iterations

Item	***Simulation Optimization***				**Performance Measures: Analytical Model**				**Performance Measures: Simulation Model**			
	Qri **(Units)**	***Rri*** **(Units)**	***Qwi*** **(Units)**	***Rwi*** **(Units)**	**F*ri*** **(Order/year)**	**F*wi*** **(Order/year)**	**B*ri*** **(Units)**	**B*wi*** **(Units of Qri)**	**F*ri*** **(Order/year)**	**F*wi*** **(Order/year)**	**B*ri*** **(Units)**	**B*wi*** **(Units of Qri)**
1	3.110	1.109	41.967	10.012	14.468	4.289	0.019	0.118	14.556	1.079	0.045	0.002
2	2.500	1.671	11.064	20.000	36.799	33.261	3.903	2.112	36.730	8.365	2.722	0.000
3	13.207	27.962	110.064	11.888	32.635	15.664	0.726	0.074	32.302	3.857	0.734	0.000
4	8.500	5.872	53.127	-0.459	11.529	7.379	0.164	0.073	11.508	1.841	0.498	0.762
		Constraints			**F*r***	**F*w***	**B*r***	**B*w***	**F*r***	**F*w***	**B*r***	**B*w***
		Estimated			23.858	15.148	4.812	2.377	23.774	3.786	3.999	0.764
		Target			24.000	12.000	4.000	0.800	24.000	12.000	4.000	0.800
		Constraint Satisfied			Yes	No	No	No	Yes	Yes	Yes	Yes
		Inventory Investment			$137,147.420				$254,248.000			

Table 12: 8-Items, *Algorithm IHOA*, Time < 1 Second

Item	***Algorihtm IHOA***				**Performance Measures: Analytical Model**				**Performance Measures: Simulation Model**			
	***Qri* (Units)**	***Rri* (Units)**	***Qwi* (Units)**	***Rwi* (Units)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**
1	5.862	0.708	46.898	39.990	38.722	19.361	0.782	0.447	38.725	4.841	0.670	0.000
2	3.017	3.250	24.133	0.236	32.487	16.243	3.239	0.182	32.362	4.041	3.506	0.021
3	5.132	0.141	41.053	16.395	18.903	9.451	0.296	0.198	19.051	2.384	0.344	0.000
4	8.738	2.620	69.908	4.191	41.769	20.885	0.766	0.269	41.660	5.206	1.622	1.301
5	1.669	-0.735	13.353	-0.372	23.365	11.683	0.599	0.161	23.102	2.889	1.028	0.138
6	14.134	10.991	113.072	25.206	10.613	5.306	0.254	0.082	10.659	1.330	0.299	0.032
7	2.837	0.771	22.700	5.489	11.278	5.639	0.929	0.173	11.360	1.422	1.177	0.072
8	4.642	3.286	37.138	32.487	14.864	7.432	1.135	0.088	14.844	1.854	2.227	1.069
		Constraints			**F*r***	**F*w***	**B*r***	**B*w***	**F*r***	**F*w***	**B*r***	**B*w***
		Estimated			24.000	12.000	8.000	1.600	23.970	2.996	10.873	2.635
		Target			24.000	12.000	8.000	1.600	24.000	12.000	8.000	1.600
		Constraint Satisfied			Yes	Yes	Yes	Yes	Yes	Yes	No	No
		Inventory Investment			$482,089.000				$664,572.720			

Table 13: 8-Items, Simulation Optimization. Time = 75,180 Seconds, 20000 Iterations

Item	***Simulation Optimization***				**Performance Measures: Analytical Model**				**Performance Measures: Simulation Model**			
	***Qri* (Units)**	***Rri* (Units)**	***Qwi* (Units)**	***Rwi* (Units)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**	**F*ri* (Order/year)**	**F*wi* (Order/year)**	**B*ri* (Units)**	**B*wi* (Units of Qri)**
1	5.500	0.449	41.239	34.000	41.273	22.018	1.359	0.946	41.524	5.540	0.819	0.000
2	3.326	4.222	19.004	0.820	29.469	20.628	2.355	0.172	29.524	5.159	2.914	0.325
3	5.137	-0.035	36.000	11.038	18.882	10.778	0.515	0.507	18.698	2.683	0.384	0.019
4	8.000	2.532	64.000	4.615	45.625	22.813	0.855	0.278	44.857	5.603	0.655	0.069
5	2.500	-1.258	8.000	-0.216	15.600	19.500	0.780	0.201	15.873	4.952	1.083	0.327
6	14.555	10.005	108.000	20.043	10.306	5.556	0.441	0.176	10.286	1.397	0.369	0.012
7	2.667	2.610	19.539	0.000	11.999	6.551	0.385	0.853	12.016	1.635	0.522	0.404
8	4.903	3.760	34.339	-0.432	14.074	8.038	0.874	0.093	14.048	1.984	1.225	0.274
		Constraints			**F*r***	**F*w***	**B*r***	**B*w***	**F*r***	**F*w***	**B*r***	**B*w***
		Estimated			23.403	14.485	7.563	3.227	23.353	3.619	7.971	1.429
		Target			24.000	12.000	8.000	1.600	24.000	12.000	8.000	1.600
		Constraint Satisfied			Yes	No	Yes	No	Yes	Yes	Yes	Yes
		Inventory Investment			$395,952.046				$547,874.100			