IEMS Vol. 8, No. 2, pp. 80-92, June 2009.

An Inventory Rationing Method in a M-Store Regional Supply Chain Operating under the Order-up-to Level System

Chumpol Monthatipkul†
Graduate School of Management and Innovation, King Mongkut's University of Technology Thonburi
126 Prachautid Road Bangmod Thrukru Bangkok 10140, THAILAND
Tel: +66 (0) 2470-9783, Fax: +66 (0) 2470-9798, E-mail: chumpol.mon@kmutt.ac.th

Received Date, April 5, 2008; Accepted Date, March 18, 2009

Abstract. This paper addresses the inventory rationing issue embedded in the regional supply chain inventory replenishment problem (RSIRP). The concerned supply chain, which was fed by the national supply chain, consisted of a single warehouse distributing a single product to multiple stores (M-stores) with independent and normally distributed customer demand. It was assumed that the supply chain operated under the order-up-to level inventory replenishment system and had only one truck at the regional warehouse. The truck could make one replenishment trip to one store per period (a round trip per period). Based on current inventories and the vehicle constraint, the warehouse must make two decisions in each period: which store in the region to replenish and what was the replenishment quantity? The objective was to position inventories so as to minimize lost sales in the region. The warehouse inventory was replenished in every fixed-interval from a source outside the region, but the store inventory could be replenished daily. The truck destination (store) in each period was selected based on its maximum expected shortage. The replenishment quantity was then determined based on the predetermined order-up-to level system. In case of insufficient warehouse inventories to fulfill all projected store demands, an inventory rationing rule must be applied. In this paper, a new inventory rationing rule named Expected Cost Minimization (ECM) was proposed based on the practical purpose. The numerical results based on real data from a selective industry show that its performance was better and more robust than the current practice and other sharing rules in the existing literature.

Keywords: Inventory Rationing Rule, Supply Chain, Order-up-to Level System, Minimization, Heuristics

1. INTRODUCTION

1.1 Overview

Nowadays, supply chains utilize a collection of different inventory replenishment systems. One such system configures regional warehouses that are supplied from a higher-level (national) supply chain. The regional warehouses then operate as an inventory distribution center for several stores or customer points within regions. In each region, the warehouse has an authority to control its downstream members. The motivation for this arrangement is that the unique demand behavior of each region is best understood and responded to at the regional level. This configuration is popular with several large and medium retailing chains and induces several important decision-making problems in supply chain management. As noted by Dotoli *et al*. (2005) the configuration of the supply chain arrangement is essential for business to pursue a competitive advantage and to meet the market demand. In this paper we study the inventory replenishment problem as it relates to regional supply chains. Specifically we solve the inventory rationing problem of a single type of products to M-stores from a single regional warehouse with a truck constraint.

Figure 1 illustrates the concerned supply chain, which is modeled from a small frozen-food supply chain in Bangkok, Thailand. A regional warehouse receives periodical inventory replenishment of frozen-food from factories, which are controlled by the national unit (Head Office). The national unit utilizes a nationwide planning system (e.g. ERP) that determines the shipping quantity and time. They are based on inventories in the national chain, the sales history and sales forecasts for each re-

† : Corresponding Author

gion. Selectively, fixed-cycle (e.g. weekly, monthly) inventory replenishments, known as the *R*, *S* system with predetermined order-up-to level is adopted. The regional warehouse then operates a limited number of trucks that transfer products to M regional stores. The stocking points are therefore M+1 and the product transfer among stores or back to the regional warehouse is not allowed. Each store experiences an uncertain customer demand, and a lost sale occurs when customer demand is not satisfied. Since the total inventory in the regional supply chain is predetermined by the national planning policy, the primary objective of the regional warehouse is to minimize the total number of lost sales within the region. The replenishment policy must therefore position the inventory to meet customer demand, while at the same time holding back inventories so as to be able to respond quickly when the demand at a store rises. Obviously, when the available inventory at the warehouse far exceeds the regional demand then the decision is trivial because each store will be overstocked. Nevertheless, when it is not sufficient the inventory rationing decision is vital because it affects system performance.

To simplify the analysis without loss of generality, the following additional conditions are made:

(i) a single type of products from the interesting industry is selected
(ii) the customer demand is independent, stationary and normally distributed
(iii) the regional warehouse operates only one truck, (which can make only one replenishment trip to only one store per period)
(iv) storage capacities at any stocking points and on the truck are sufficient
(v) inventory replenishments occur at the start of each period and are available in that period
(vi) shipment quantity from the sources to the regional warehouse is known in advance and
(vii) The store to replenish is selected based on its maximum expected shortage

The main concern in our research is the determination of the replenishment quantity after selecting the truck destination so as to minimize the total lost sales in long run. In section 2 we mathematically describe the regional supply chain replenishment problem. Section 3 proposes a new inventory rationing rule and also presents existing inventory rationing rules. Section 4 presents the solution procedure. In section 5 numerical results to evaluate the relative performance of the proposed method against the existing methods are presented. Section 6 concludes the results.

1.2 Related Literature

So far, the inventory rationing problem has received much attention from researchers. Earlier work on this issue belongs to Veinott (1965), he focused on periodic review systems. Veinott analyzed an inventory rationing with the concept of critical levels. His model was concerned with multiple demand classes, zero lead time, and backordering policy. Topkis (1968) studied a similar model, but he considered both backorder and lost sale situations. His dynamic programming showed that, under certain conditions, a base stock policy was optimal and the optimal rationing policy could be specified by a set of control limits. Kaplan (1969) made similar analysis but for two classes of customers. Nahmias and Demmy (1981) also specified critical levels for several demand classes. They first studied a single period inventory model, where demand was accounted at the end of a period then extended to a multiple-period model. Demand was satisfied only if the inventory level was above the critical level, this way it was possible to reserve stock for possible future higher-priority demand. A two-echelon divergent supply chain was first considered by Eppen and Schrage (1981). Their model contained many

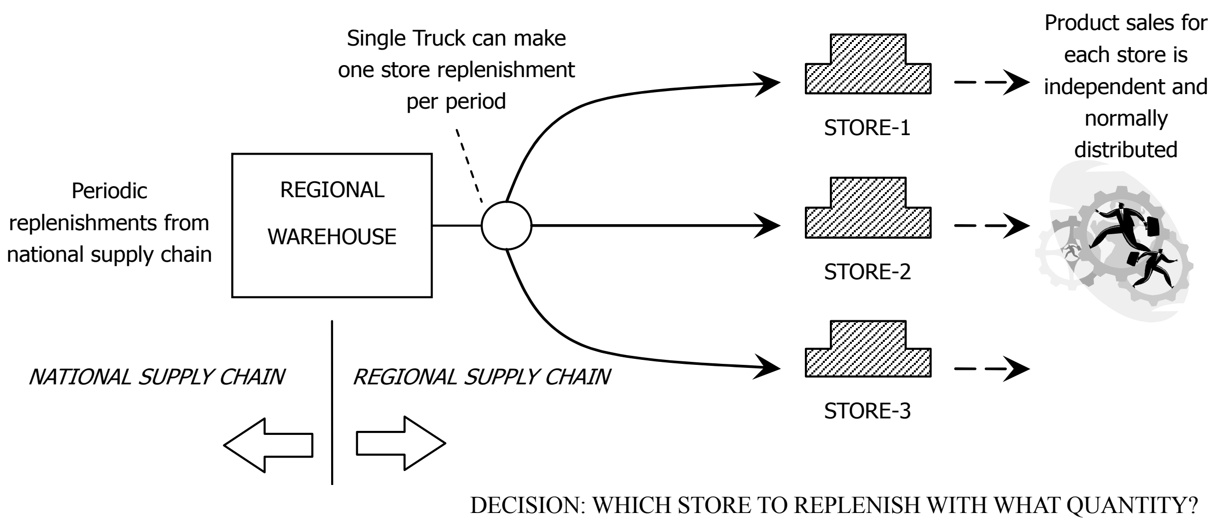

Figure 1. The single truck M-store regional supply chain replenishment model.

assumptions: such as (i) the depot did not maintain inventories (ii) the holding-penalty cost ratios at customers were identical (iii) the customer demand was normally distributed (iv) lead times of the customers were identical and (v) for each allocation, the depot received enough material from the supplier to be able to allocate the material to each customer such that an equal fractile point was achieved (presently known as balanced inventories assumption). Under these assumptions, Eppen and Schrage (1981) derived an optimal order-up-to-policy at the depot, assuming no set-up costs. In case of fixed set-up costs, an approximate optimal policy was derived based on an equal-stockout probabilities allocation rule. Federgruen and Zipkin (1984) relaxed some assumptions of Eppen and Schrage (1981). They stated the holding-penalty cost ratios could be varied. The customer demand could be other classes of demand distributions (e.g. Erlang and gamma distributions). The material allocation was obtained by solving a myopic allocation rule, which aimed to minimize the expected costs in the period that the allocation actually took effect, ignoring costs in all subsequent periods. Zipkin (1984) proposed a dynamic program to allocate inventories from a single depot to many demand points. His model was concerned with normally distributed demands with backlogging of unfulfilled demand. The total cost included inventory holding costs and backordering costs. Jackson (1988) also studied a stock allocation in a two-echelon distribution system. He showed significant effects of risk pooling between replenishments in a single-cycle model where the initial stock level was given. In his model, the replenishment cycle was divided into many equal intervals. At the start of each interval, the inventory positions of the local stockpoints were raised to their order-up-to levels if sufficient central stock was available. McGavin *et al*. (1993) also considered allocation policies in a single replenishment cycle with given initial stock. Various ways were examined to exploit the effect of risk pooling between replenishments. They concluded that benefits could be obtained using a simple 50/25 heuristic–the replenishment cycle contained two shipment opportunities, 25% of the mean replenishment cycle demand was kept at the central depot for a second shipment after 50% of the replenishment cycle had passed. Both Jackson (1988), and McGavin *et al*. (1993) did not consider the issue of controlling the total system stock by adequate replenishments to the central depot. A similar model with a different cost structure, including the costs of shipments between the central depot and local warehouses, was analyzed by GÜllÜ and Erkip (1996).

With a long history in this area an extensive review on divergent multi-echelon systems can be seen in Diks *et al*. (1996). The authors reviewed inventory rationing methods, namely, Fair Share (FS) rationing, Appropriate Share (AS) rationing, Consistent Appropriate Share (CAS) rationing, Priority rationing, and other rationing rules. The FS rationing policy rations the available material so as to maintain all end-stockpoints at a balanced position: all end-stockpoints have the same non-stock-out probability. The AS is to ensure that a prespecified target service level can be attained at a retailer. In the AS, the allocation fractions can be chosen freely. In a restricted version of AS, which is called the CAS method, the allocation fraction is chosen such that the ratio of the projected net inventory at any retailers over the system-wide projected net inventory constant at any time. The Priority rationing uses a list of retailers and rations from the available inventory so as to satisfy them in the sequence they were listed. The CAS rule is also introduced in De Kok (1990). The author analyzed a two-echelon divergent model with stockless depot and took into account the service criterion (fill rates), which was more customer-oriented and therefore practically more important. The similar models which allow the depot to maintain stock were analyzed by Seidel and De Kok (1990) and Verrijdt and De Kok (1996). The adaptation of the CAS policy to a one-depot/multi-retailer inventory system is presented in Diks and De Kok (1996). They used the Monte-Carlo simulation to validate some heuristics which were used to determine integral order-up-to-levels, parameters of allocation policy at the depot and of the rebalancing policy at the retailers so as the desired service levels were attained at the retailers for minimal expected total costs. Lagodimos (1992) proposed models to measure system service performances of two-echelon divergent production networks consisting of a central stocking point feeding a number of end stocking points. The system operated under periodic review ordering policies and used fair sharing and priority rules when the central stocking point had insufficient inventories to cover its demands. Diks and De Kok (1999) extended the study to cover a divergent multi-echelon inventory system. They proposed a decomposition-based algorithm to determine the control parameters of a near-optimal replenishment policy: the order-up-to-levels and the allocation fractions. Diks and De Kok (1999) considered both inventory holding cost and penalty cost for backorders. The extension version to cover multi-echelon systems with only end-stockpoint were allowed to hold stock was also studied by Verrijdt and De Kok (1995). A divergent multi-echelon inventory system in which every member was allowed to hold stock was analyzed in Diks and De Kok (1998). Each node was responsible for allocating materials among its successor periodical-fixed lead time. The model concerned fixed lead time, periodical orders, and backlogging. The objective function was to minimize the average costs per period in the long term, which included penalty and inventory holding costs. A more complex system named Supply Chain Operations Planning Problem (SCOP) was studied by De kok and Fransoo (2003). SCOP aimed to coordinate the release of materials and resources such that supply chain costs were minimized under the required customer service. The author proposed the so-called Synchronized Base Stock (SBS) policy and compared it against the

LP-based planning system. The SBS outperformed the LP-based system because it used a more sophisticated linear-allocation inventory rationing rule. This rule coincides with the one derived in Van der Heijden (1997), which is named Balanced Stocking (BS) rationing. Van der Heijden (1997) proved that the BS was accurate and more robust than the CAS rationing rule. The BS consists of two parameter sets: a set of rationing factors at the depot with the purpose to minimize inventory imbalance and a set of order-up-to levels at the local stockpoint with the purpose to achieve the target fill-rate. If sufficient central stock is available, the inventory position of each local stockpoint is raised to the target level and the remaining part is kept at the central depot for subsequent shipments. If the central stock is less than required, then the local inventory positions are only raised to the target level minus a rationing factor multiplied by the projected shortage. An extensive numerical experiment to prove the effectiveness of the BS method on general multi-echelon distribution systems was performed by Van der Heijden *et al*. (1997). Van der Heijden (1999) considered shipment frequencies on a two-echelon divergent supply chain consisting of a central depot and multiple non-identical local warehouses. He showed that a decomposition approach used in Van der Heijden (1997) could be used to solve this model. Van der Heijden (2000) dealt with the optimization of stock levels in general divergent networks operating under a periodic-review/order-up-to level policy. The objective was to reach target fill rates with minimum inventory holding cost. A procedure to determine the control parameters was also proposed and then tested by extensive numerical experiments.

Recently, a model with several demand classes, as first studied by Veinott (1965) was still being investigated. Melchiors *et al*. (2000), and Melchiors (2003), analyzed an (s, Q) inventory model with unit Poisson demand, several demand classes and lost sales. He proposed the so-called Restricted Time-Remembering (RTR) policy, where the critical levels were allowed to be varied depending on time. Deshpande *et al*. (2003) used a threshold inventory rationing policy for service differentiated demand classes. The customer service level was also further analyzed by Zhang (2003). He considered multi-customer service level requirements when the Priority Rationing rule was applied. He demonstrated that the optimal purchasing quantity for the two-customer problem might be found through a line search. For multi-customer problems, a formula to find a near-optimal purchasing quantity was derived. Ayanso *et al*. (2006) also used a threshold level inventory rationing policy. They considered non-perishable make-to-stock products in the context of e-commerce.

The extension of the inventory rationing issue to cover the production stage can be seen in Lee and Hong (2003), Huang and Iravani (2007), Ha (1997a, b, 2000), De Véricourt *et al*. (2002), and Axsäter *et al*. (2004). Ha (1997b) considered the problem of stock rationing in a capacitated make-to-stock production system with multi-class demand and lost sales. He showed that the threshold-type policies were optimal for both production and rationing decisions. Similar results were provided in Ha (1997a) for systems with two demand classes and backlog. De Véricourt *et al*. (2002) extended Ha's mo-del to multiple demand-class models. They demonstrated that the benefit of inventory pooling could be realized only if the stock was efficiently allocated. Axsäter *et al*. (2004) studied a similar model in a continuous review two echelon inventory model. Lee and Hong (2003) studied a production model with a (s, S) control policy. The production time to process an item was assumed to follow a 2-phase Coxian distribution. A stock rationing was set for each demand class and the demands were lost if the inventory level was lower than the predetermined rationing level. The production system was modeled as a continuous-time Markov chain, and an efficient algorithm to calculate the steady-state probability distribution was then proposed. The average operating cost under the stock rationing policy was compared with the one without stock rationing. Huang and Iravani (2007) considered a two-echelon capacitated supply chain with two non-identical retailers and information sharing. They discussed the benefits of the optimal stock rationing policy over the first come first served (FCFS) and the modified echelon-stock rationing (MESR) policies.

The literature review reveals that while several papers consider the divergent supply chain, especially the two-stage supply chain with one upstream and multiple downstream members, they are primarily focused on the inventory and distribution costs, which typically include inventory holding, inventory ordering, shipment, and stockout costs. The problem studied in this article differs in that the inventory holding and distribution costs are not a main concern. It is because the authority faces a facing problem and his or her main concern is to provide the region with as many products as reasonably possible with regard to the customer loyalty. Therefore, the main focus is on the lost sale cost, which therefore affects the replenishment problem. More justification of this setting can be seen in Monthatipkul (2006). Further, all previous studies on the inventory rationing problem never considered the vehicle constraint, while this issue is of concern in this paper. This vehicle constraint is significant because it affects the inventory replenishment system in term of some stores cannot be replenished simultaneously due to lack of trucks. Thus, benefits of rationing some stocks to these stores need investigation. Finally, the existing literature normally solves the inventory rationing problem by using a critical level or a priority list, equalizing the stockout probability or target service levels among stockpoints, solving a dynamic program, multiplying by a rationing fraction, utilizing the second shipment opportunity. In this paper, an inventory rationing rule based on expected cost minimization is proposed.

2. PROBLEM DESCRIPTION

2.1 The regional supply chain inventory replenishment problem (RSIRP)

The regional supply chain model under consideration is illustrated in figure 1. It consists of a centralized regional warehouse supplying a single type of product to multiple regional stores (M-store) with a single truck, which can make a round trip per period. The regional warehouse gathers the inventory status of the entire regional supply chain and then decides which store in the region to replenish and also determines the replenishment quantity. The following notations are used to mathematically describe the main decision problem at the regional warehouse.

Notation:

- j store index (j = 1, 2, 3, ⋯, M)
- t time period index (t = 1, 2, 3, ⋯, T)
- I_t inventories at the regional warehouse at the end of period t
- $V_{j,t}$ inventories at store j at the end of period t
- A_t shipment quantity arriving at the regional warehouse at the start of period t, it is noted that A_t is known in advance because the Head Office must place orders at the factories in advance due to the order processing and transportation lead-time.
- μ_{Dj} mean customer per period demand for the product at store j
- σ_{Dj} standard deviation of customer demand per period for the product at store j
- $d_{j,t}$ actual demand or sales at store j in period t
- $L_{j,t}$ Lost sales (stock outs) at store j in period t
- C_j Cost of a lost sale (stock out) per unit at store j

Main decision variables at the regional warehouse:

- $X_{j,t}$ a binary variable representing which store the truck replenishes in period t (if store j is supplied then $X_{j,t} = 1$ else $X_{j,t} = 0$)
- $Q_{j,t}$ replenishment quantity shipped to store j in period t

At the end of each period, the planner learns the inventory status as [I_t, $V_{j,t}$ | j = 1 to M]. At any point in time $d_{j,t}$ is known for the history periods and unknown for the future periods, but it can be described by the mean μ_{Dj} and standard deviation σ_{Dj}. Since parameters of the periodic review order-up-to level system are predetermined in a higher-level system configuration (see section 2.2), hence minimizing the inventory levels is not a primary objective. Also, the supply chain operates with only one truck, which can make only one round trip per period, the inventory positioning to the right place with the right quantity to minimize the lost sale cost ψ becomes a key performance metric. Hence, the objective function of the regional warehouse can be expressed as an equation 1. The transportation cost is excluded from the model because it is dominated by the lost sale cost. Thus, frequent replenishments to stores so as to position inventories for a minimum ψ is reasonable.

$$\text{Minimize the lost sale cost } (\psi) = \text{Min } \Sigma_{j,t}\, C_j\, L_{j,t} \qquad (1)$$

It is clear from the truck constraint that in any period at most one $X_{j,t}$ can be non-zero. Thus, the RSIRP at the regional warehouse can then be described as selecting $X_{j,t}$ and $Q_{j,t}$ so as to minimize ψ given the current state [I_t, $V_{j,t}$].

2.2 The pre-designed order-up-to level system

It is assumed that the supply chain operates under the periodic review order-up-to level system (R, S system). The warehouse inventory is replenished in every fixed-interval FI (e.g. FI = week) from a source outside the region, but the store inventory can be replenished daily by the single truck. The inventory replenishment at each store occurs at the start of each period and is available in that period. Let S_j be the pre-designed order-up-to level for a store. It can be intuitively derived by the equation 2.

$$S_j = M\mu_{Dj} + \lambda_j\, \sigma_{Dj} \sqrt{M} \quad \text{where } \lambda_j \geq 1 \qquad (2)$$

Note that in this paper the simple parameter setting of S_j is adopted (for more advanced method, see Schneider and Rinks, 1991 and Schneider *et al.*, 1995). The above equation is modified from the so-called Modified Base Stock (MBS) policy which is clearly explained in De Kok and Fransoo (2003). It projects that the next replenishment to the store will be M periods later. S_j is therefore set equal to the cumulative expected demand plus a safety factor over the next $\sqrt{M}$ periods. The safety factor at store j is represented by λ_j, and can be selected to meet a specific stockout probability. Similarly, the pre-designed order-up-to level at the warehouse S_w must be equal to (FI+TL) $\mu_W + \lambda_W\, \sigma_W \sqrt{FI+TL}$, where μ_W and σ_W are average aggregate demand and standard deviation across all downstream members, respectively, TL is the sum of order processing time and transportation lead-time from the external source to the warehouse, and λ_W is a safety stock factor at the warehouse.

2.3 Store selection and the concerned inventory rationing problem

As mentioned earlier, only one store j can be selected at each period t ($\Sigma_j X_{j,t} = 1$ for each period t). A method based on the basis of maximum expected shortage is adopted here to determine the variable $X_{j,t}$. Details of such a method can be found in Monthatipkul (2006). It is briefly explained as follow.

Determination of $X_{j,t}$

Let $ES_{j,t}$ be the expected shortage cost for store j in period t. $X_{j,t}$ is therefore determined as follows:

$$ES_{j,t} = \int_{V_{j,t-1}}^{\infty} P(y)\{(y - V_{j,t-1})\, C_j\}dy \quad (3)$$

$$X_{j^*,t} = 1 \mid ES_{j^*,t} = \text{Max}_j \{ES_{j,t}\} \text{ else } X_{j,t} = 0 \quad (4)$$

where P(y) is the normal probability density function for the distribution $N(\mu_{Dj}, \sigma_{Dj})$.

Equation 3 represents the expected shortage cost at store *j*, which is determined by an integration function over the lower limit $V_{j,t-1}$ to infinity. Equation 4 shows that the selected store must have maximum $ES_{j,t}$. Once a store is selected then the replenishment quantity $Q_{j,t}$ must be derived. It is obvious if the inventory at the warehouse is sufficient, the inventory rationing is not a must. Nevertheless, if the warehouse has limited inventories, an inventory rationing rule must be imposed. It is because positioning inventories to a wrong location with wrong quantity returns poor system performance. This paper aims to propose a new inventory rationing rule, which will be presented in the next section, and to compare it against the existing inventory rationing rules.

3. INVENTORY RATIONING RULES

This research proposes a new inventory rationing rule. It's performance has also been compared to the existing inventory rationing rules. Even though there are a lot of inventory rationing rules reported in the literature (see section 1.2 for details), none of them is developed based on the concerned situation. Thus, this paper selects some interesting inventory rationing rules from the literature and applies only the main ideas behind those rules for the current circumstance.

3.1 Existing Inventory Rationing Rules

- **The Current Practice rule (CP)**

The Current Practice rule is used in the frozen-food supply chain in Bangkok, Thailand. Basically, when this rule is adopted the product will be transferred to the selected store as much as possible until its order-up-to level is reached. This rule represents the neglect of the inventory replenishment opportunity of the other stores in the subsequent periods.

- **The Fill-Rate Based Fair Share rationing rule (FRBFS)**

The FRBFS is selected because it is a modified version of the Fair Share rule, which is widely studied by many researchers (see Dik *et al*., 1996). The FRBFS contains two main steps as follows (for details see Monthatipkul, 2006);

a) Calculating the inventory sufficiency indicator

The inventory sufficiency indicator is calculated as a function of (i) S_j, (ii) PIA_t-the total projected available to promise inventory at the regional warehouse for the next M-1 periods, and (iii) PSD_t-the total projected store replenishment demand for the current period. It is noted that PSD_t is calculated based on the present time other than the next M-1 periods, because one cannot determine the demand at the future periods due to unavailability of the on-hand inventories at the future periods. The second and third factors are derived as follows:

$$PIA_t = I_{t-1} + \Sigma_{\tau=t,\, t+M-1} A_\tau \quad (5)$$

$$PSD_t = \Sigma_{j\in M} \{S_j - V_{j,t-1}\} \quad (6)$$

For the current instance, the supply-demand ratio given by PIA_t/PSD_t is indicative of how constrained the replenishment problem is. If the ratio is greater than or equal to one then the regional warehouse is considered as having sufficient inventories, and the shipment quantity $Q_{j,t}$ is selected to meet the S_j target. If the ratio is less than one then the regional warehouse is considered as having limited inventories, and the shipment quantity $Q_{j,t}$ must be re-calculated by step b).

b) Determining the replenishment quantity $Q_{j,t}$

Since the rationing rule has effects when the ratio $PIA_t/PSD_t < 1$. The FRBFS suggests the deduction of the original shipment quantity $Q_{j,t}$ by the ratio PIA_t/PSD_t. Mathematically, we have

$$Q_{j,t} = (PIA_t/PSD_t)(S_j - V_{j,t-1}), \text{ where } PIA_t/PSD_t < 1 \quad (7)$$

- **The Balanced Stock rationing rule (BS)**

The BS rule is selected because it outperforms many existing rules in the literature (see Van der Heijden, 1997 and De kok and Fransoo, 2003). This rule is applied only if $PIA_t/PSD_t < 1$. The general concept of the Balanced Stock rationing rule can be seen in Van der Heijden (2000). Briefly, when the shortage at the warehouse equals *x*-for this paper, $x = PSD_t - PIA_t$, when $PIA_t/PSD_t < 1$-the inventory level of its successors are raised to the levels $S_j - p_j x$, where p_j are rationing fractions such that $\Sigma_j p_j = 1$, calculated as seen in equation 8.

$$p_j = \{1/2M\} + \{(\sigma_{Dj})^2 / 2\, \Sigma_j(\sigma_{Dj})^2\} \quad (8)$$

3.2 A Proposed Inventory Rationing Rule

- **The Expected Cost Minimization rationing rule (ECM)**

The ECM is a myopic allocation rule, it aims to minimize the expected cost in the period that the allocation takes effect and ignore all costs in all subsequent periods. Similar to the BS, this rule is applied only if $PIA_t/PSD_t < 1$. The ECM is applied by solving the problem **P** for each period *t*.

The problem **P**:

$$\text{Minimize } \Sigma_j ES^*_{j,t} \quad (9)$$

s.t.

$$ES^*_{j,t} = \int_{V_{j,t-1}+Q_{j,t}}^{\infty} P(y)\{(y - V_{j,t-1} + Q_{j,t})C_j\}dy \;\; \forall j \quad (10)$$

$$\Sigma_j Q_{j,t} = I_{t-1} + \Sigma_{\tau=t,\, t+M-1} A_\tau \quad (11)$$

$$Q_{j,t} \geq 0 \text{ for all } j \quad (12)$$

The objective of function 9 is to minimize the overall expected shortage of the entire supply chain. Equation 10 shows a cumulative cost function to estimate the shortage cost. The sign of $Q_{j,t}$ is a plus sign because it represents a shipping quantity of products which are transferred to the stores in order to increase the inventory levels of the stores. Therefore, the equation 10 represents the expected cumulative cost after transferring the products to each store. Equation 11 shows that the total projected shipping quantity must be equal to the total projected available to promise inventory at the regional warehouse for the next M-1 periods. Constraint 12 is the non-negative condition. A procedure to solve the problem **P** is presented in the next section. It is noted that the $Q_{j,t}$ is finally equal to all products available in period t ($I_{t-1}+A_t$) if it exceeds the sum of current inventories I_{t-1} plus shipment quantity A_t.

4. THE SOLUTION PROCEDURE

Figure 2 presents an iterative approach to solve the problem **P**. The drawback behind the flow chart is the selection of a giver and a taker from all stores. The giver will dedicate an amount of its $Q_{j,t}$ (represented by SCALE α) to the taker such that the overall expected shortage cost of the entire supply chain is decreased. The giver will be selected if the increase in its expected shortage cost $ES^*_{j,t}$ is at least. In contrast, the taker will be selected if the decrease in its expected cost is at most. The giver-taker selection process is stopped when there is no improvement in the objective function 9. The step-by-step explanation is as follow.

Step 1 Determine the initial solution $Q_{j,t}$ for all j as a weight ratio $[\{S_j - V_{j,t-1}\} / \Sigma_{j\in M} \{S_j - V_{j,t-1}\}]$ multiplied by the total projected available to promise PIA_t

Step 2 Determine $ES^*_{j,t}$ for all j using equation 10 by a numerical integration

Step 3 Determine the SCALE α. Initially, α is set to 50% of $\text{Min}_j \{Q_{j,t}\}$. It is then changed to Max $\{0.5\alpha^{old}, UM\}$, where UM is the smallest unit of measure of the product, when new value of $\Sigma_j ES^*_{j,t}$ ($\Sigma_j ES^{new*}_{j,t}$) is greater then the old value of $\Sigma_j ES^*_{j,t}$ ($\Sigma_j ES^{old*}_{j,t}$)

Step 4 Change the old α to the new α according to step 3

Step 5 Stop if the stopping condition is satisfied. The procedure stops when α = UM and $\Sigma_j ES^{new*}_{j,t} > \Sigma_j ES^{old*}_{j,t}$

Step 6 Select a GIVER and a TAKER. The GIVER will dedicate its fraction of $Q_{j,t}$ (by the amount of α) to the TAKER such that $\Sigma_j ES^{new*}_{j,t}$ will be less than $\Sigma_j ES^{old*}_{j,t}$. The GIVER j is selected only if its $Q_{j,t} > 0$ and the associated term ΔG_j is at least among all stores. In contrast, the TAKER j is selected when the related term ΔT_j is at most among all candidates, where

$$\Delta G_j = \left(\int_{V_{j,t-1}+Q_{j,t}-\infty}^{\infty} P(y)\{(y-V_{j,t-1}+Q_{j,t}-\infty)C_j\}dy\right) - \left(\int_{V_{j,t-1}+Q_{j,t}}^{\infty} P(y)\{(y-V_{j,t-1}+Q_{j,t})C_j\}dy\right) \quad (13)$$

$$\Delta T_j = \left(\int_{V_{j,t-1}+Q_{j,t}}^{\infty} P(y)\{(y-V_{j,t-1}+Q_{j,t})C_j\}dy\right) - \left(\int_{V_{j,t-1}+Q_{j,t}+\infty}^{\infty} P(y)\{(y-V_{j,t-1}+Q_{j,t}+\infty)C_j\}dy\right) \quad (14)$$

In case of GIVER = TAKER, the GIVER will be maintained and a new TAKER which has the second maximum ΔT_j will be selected.

Step 7 Decrease the GIVER by α and increase the TAKER by α, go back to step 2

Step 8 Keep $Q_{j,t}$ for all j which can give minimum $\Sigma_j ES^*_{j,t}$ as the final solutions

5. NUMERICAL EXPERIMENTS AND RESULT

Performance of the proposed rationing rule was evaluated against the selective rationing rules. The numerical experiments were conducted based on a real situation of a selected supply chain in Bangkok, Thailand, namely the frozen-food supply chain.

5.1 Experimental Base Case

The concerned supply chain consists of one regional warehouse and multiple stores supplying over 50 product types to its customers. To simplify the analysis, the experiments were concerned with only three selective stores (M = 3) and one truck which was dedicated to one selective product type. The product was selected in which its customers demand at each store could be assumed to follow normal distributions (in the long term). From the real data set, the mean demands, standard deviations, and stock out costs at the three stores were as follows, μ_{Dj} = {428, 418, 423), σ_{Dj}/μ_{Dj} = 0.1, and C_j = {6.00, 8.50, and 7.00}. The factor λ_j was set to 1.5 for all store j. The initial inventories at the three stores were $V_{j,0}$ = {450, 500, 475). The inventory replenishment cycle at the regional warehouse was set to five (FI = 5). This was equivalent to a five-day delivery schedule. The order processing plus transportation lead-time from the external source to the warehouse was equal to two days (TL = 2). Thus, the Head Office must place an order to the factories two days in advance. While placing an order, the Head Office used the current inventory status of the regional supply chain. Our base problem used the safety factor λ_W = 1. The initial inven-

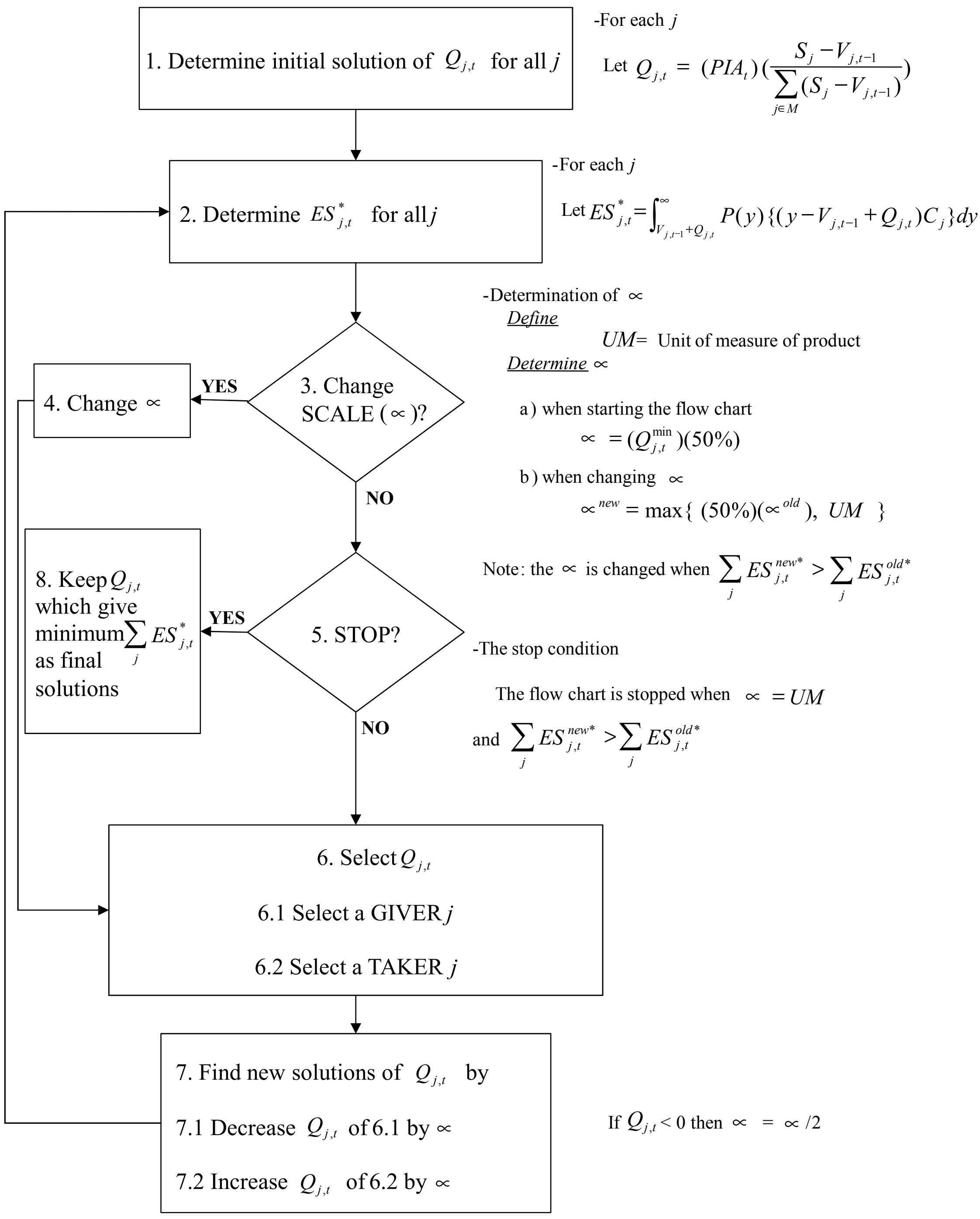

Figure 2. A flow chart to solve the problem **P**.

tory at the warehouse was 4500 ($I_0 = 4500$), and the experiment was run to cover 20 periods (T = 20).

5.2 Experiments

In this paper, there were two main experimental objectives which would be proved by three main tests. The details are described as follows.

1st Experimental Objective: Rationing Rule Performance Measurement

The first numerical test was to compare efficiencies among the rationing rules under the base case. After the 20-day run, the following numbers were recorded for each rule; the total lost sale cost, the total products remaining in the system (REMAIN), and the difference between the total customer demand and the total products entered into the system (DIFF). The experiment was repeated until the 95%-confidence intervals of the observed average numbers did not exceed their 5% errors.

For the general purpose, the second test was performed to evaluate the efficiencies of the rationing rules under various parameter settings. The test was conducted with 16 different sets of mean customer demand μ_{Dj}, standard deviation σ_{Dj}, and cost of a lost sale C_j. They were re-generated within the 50% gap of those used in the base case. For each parameter setting, the experiment was also repeated until the 95%-confidence intervals of the observed average numbers did not exceed their 5% errors.

2nd Experimental Objective: Effects of Interesting Parameters

The third numerical test was to study the effect of interesting parameters, namely, cost of a lost sale C_j and standard deviation σ_{Dj}. The base case was used and then the lost sale cost C_j was varied ±40% (step by 20%), while the other parameters were kept constant. The effect of the standard deviation σ_{Dj} was then studied. The ratio σ_{Dj} / μ_{Dj} was increased from 0.1 to 0.5 (step by 0.1).

5.3 Results

Table 1 concluded the average numbers of the total lost sale costs from the 1st and 2nd numerical tests. The base case was shown on the first row. The remaining 16 cases were the results of the 2nd tests with 16 different sets of parameter settings. One-way ANOVA was applied and it could be concluded that the rationing rule significantly affected the total lost sale costs at a significant level of 0.05. The multiple comparison test, known as the Duncan Method, was then preformed to classify the rationing rules into homogeneous subsets. Their ranks were shown by the numbers in the parentheses. The review of table 1 showed that the ECM outperformed other rules. It was always in the first rank. The FRBFS and BS were almost in the second rank. Interestingly, the CP rule, which was currently used in the frozen-food supply chain in Bangkok, Thailand, provided the worst results. In fact, the CP rule represented that the manager did not realize benefits of rationing some stocks to some other stores to take advantages of the transshipment in the subsequent periods. It was proved by this research that rationing some appropriate stocks for future transferring was beneficial and the current practice of the supply chain should be reviewed. The FRBFS and BS were poorer than the ECM because they sometimes allocated too less or much stock for the selected store. This was a result of the calculation which took into account of the demand distribution functions. The remaining stock was also kept at the regional warehouse due to lack of truck. This remark coincides with the paper of Van der Heijden (1997).

Table 2 presented the average numbers of DIFF and REMAIN. DIFF numbers were obtained by the total

Table 1. Total lost sale costs (average numbers) from the 1st and 2nd tests.

Total lost sale cost (average numbers)				
	ECM-Rule	FRBFS-Rule	BS-Rule	CP-Rule
Base Case	**7462(1)***	7681(2)	7680(2)	7914(3)
Case				
1	**8919(1)**	**9006(1)**	9147(2)	9405(3)
2	**7014(1)**	7222(2)	7217(2)	7416(3)
3	**8547(1)**	8704(2)	8661(2)	8930(3)
4	**6288(1)**	6504(2)	6552(2)	6698(3)
5	**5234(1)**	**5243(1)**	5316(2)	5488(3)
6	**5707(1)**	5878(2)	5879(2)	5999(3)
7	**4683(1)**	**4713(1)**	4887(2)	4995(3)
8	**6301(1)**	6447(2)	6433(2)	6566(3)
9	**3955(1)**	4095(2)	4120(2)	4124(2)
10	**6326(1)**	6368(2)	6372(2)	6507(3)
11	**6342(1)**	6543(2)	6535(2)	6746(3)
12	**4850(1)**	5004(2)	5014(2)	5033(2)
13	**5345(1)**	5520(2)	5545(2)	5699(3)
14	**4048(1)**	4121(2)	4126(2)	4248(3)
15	**7265(1)**	7405(2)	7382(2)	7698(3)
16	**4359(1)**	4492(2)	4468(2)	4599(3)

Note) * represents the rank number of the rule classified by the multiple comparison test-Duncan Method.

customer demand minus the total supplied products. The positive number represented the insufficiency of the total supply toward the total demand. Note that DIFFs of the FRBFS, BS, and CP rules (columns 4, 6, and 8) were presented in terms of percentage gaps which were determined using DIFF of the ECM as a base. For example, the -5% gap of the CP rule in the base case meant that its DIFF was 5% less than that of the ECM. REMAIN numbers showed the remaining products in the system after the 20-day running. The positive numbers indicated that some products were left in the system and not sold to the customer. Table 2 provided some observations as follows.

(a) DIFF values could indicate the insufficiency of the system because they were all positive. Thus, rationing rules were a must because the system was not overstocked.
(b) CP-DIFFs were all less than ECM-DIFFs. These results implied that more products were always supplied to the system when the CP rule was used instead of the ECM rule. It was because the total demand was kept constant for each replication across all the tests. A lower number of DIFF represented a smaller gap between the total demand and supply and consequently referred to a higher supply level in the system. This observation amplified the poor performance of the CP rule. It acquired more products, but returned higher total lost sale costs.
(c) DIFFs of FRBFS and BS rules could be more or less than those of the ECM. It meant that it was possible that the system acquired more or less products when the FRBFS or BS was in place.
(d) REMAINs were all positive for all rationing rules. It signified that there were always products remaining in the system after the 20-day running even though the system acquired insufficient products. The remaining products could be anywhere in the system and not sold to the customers. This condition implied there were always shortages in the system because products could be transferred to wrong places, or at the wrong times with wrong quantities. This phenomenon was a result of demand uncertainty and could not be eliminated by any rationing rules.

Figure 3 and Figure 4 exhibited the average-number graphs of the total lost sale costs from the 3rd numerical test when the parameter C_j and the ratio σ_{Dj}/μ_{Dj} was varied, respectively. In Figure 3, the horizontal axis represented factors to multiply the parameter C_j. For example, the number 1 referred to the base case. The number 0.8 represents the sensitivity analysis of the parameter C_j when 20% downside was concerned. In Figure 4, the horizontal axis represents the change of the ratio σ_{Dj}/μ_{Dj}. For example, the number 0.1 referred to the base case. The number 0.2 represents the sensitivity analysis of the ratio σ_{Dj}/μ_{Dj} when 100% upside was concerned. Figure 3 and Figure 4 confirmed that the ECM performance was better and more robust than the other rules throughout the wide ranges of the concerned pa-

Table 2. DIFF and REMAIN (average numbers) from the 1st and 2nd tests.

DIFF/REMAIN (Average numbers)								
	ECM-Rule		FRBFS-Rule		BS-Rule		CP-Rule	
	DIFF	REMAIN	DIFF	REMAIN	DIFF	REMAIN	DIFF	REMAIN
Base Case	**794.85**	318.28	**-3%***	318.18	**+2%**	318.18	**-5%**	318.86
Case								
1	**927.96**	286.14	**+1%**	289.77	**-2%**	290.56	**-4%**	300.33
2	**783.63**	304	**-1%**	312.64	**+3%**	312.64	**-3%**	311.91
3	**935.25**	301.28	**-1%**	304.04	**-3%**	302.68	**-3%**	305.98
4	**859.93**	309.77	**+4%**	314.19	**-4%**	316.09	**-8%**	318.91
5	**591.22**	357.22	**-3%**	362.51	**-2%**	364.23	**-8%**	362.46
6	**757.09**	332.59	**-1%**	334.48	**-3%**	333.7	**-7%**	331.21
7	**520.09**	378.38	**-5%**	382.86	**+1%**	383.29	**-5%**	387.21
8	**871.19**	309.07	**+3%**	313.6	**-2%**	311.91	**-4%**	312.28
9	**361.21**	403.81	**+1%**	404.07	**+3%**	404.15	**-4%**	405.21
10	**871.71**	292.2	**-5%**	296.77	**-2%**	298.57	**-8%**	295.8
11	**794.96**	321.49	**-1%**	326.25	**+3%**	327.8	**-6%**	326.21
12	**730.94**	299.33	**-2%**	301.03	**-4%**	300.33	**-6%**	298.08
13	**502.5**	340.54	**-6%**	344.57	**+2%**	344.96	**-7%**	348.48
14	**643.52**	278.16	**+5%**	282.24	**-3%**	280.71	**-4%**	281.05
15	**442.07**	363.42	**+3%**	363.66	**-1%**	363.73	**-4%**	364.68
16	**740.51**	262.98	**-1%**	267.09	**+1%**	268.71	**-2%**	266.22

Note) * represents the DIFF value in percentage term which is relative to that of the ECM rule.

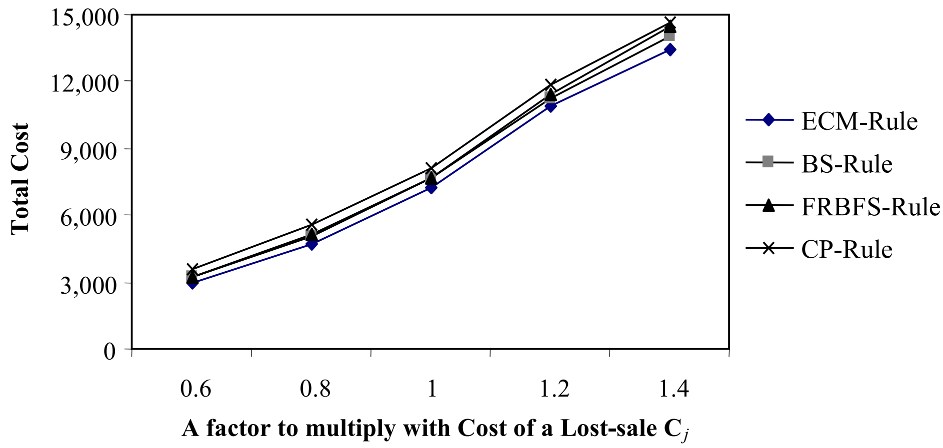

Figure 3. Average total cost graphs from the sensitivity analysis of C_j.

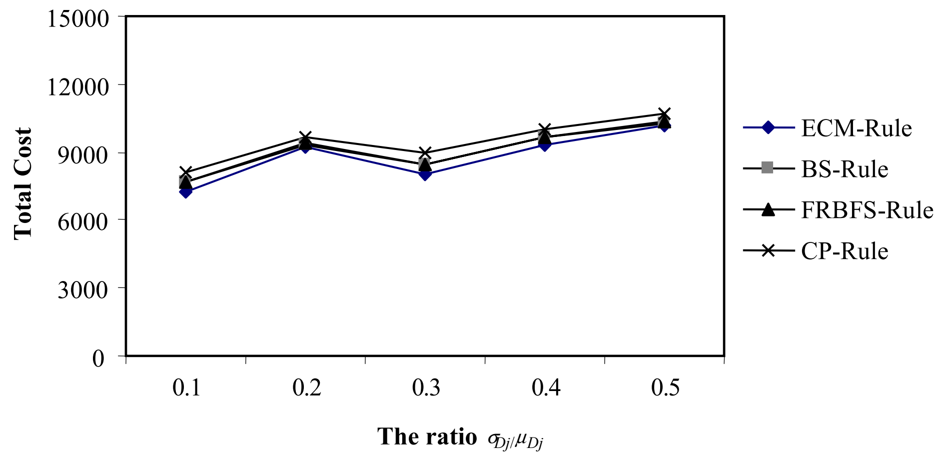

Figure 4. Average total cost graphs from the sensitivity analysis of the ratio σ_{Dj}/μ_{Dj}.

rameters. The CP rule, which was the current practice of the selective frozen-food supply chain in Bangkok, Thailand, returned poor performances. From Figure 3, it could be observed that the total cost had a positive relation with the value of C_j for all concerned rationing rules because it was increased when C_j was increased. However, it did not have a relationship with the ratio σ_{Dj}/μ_{Dj} (see Figure 4).

6. SUMMARY

We studied an inventory rationing problem embedded in the regional supply chain inventory replenishment problem (RSIRP). First, the concerned regional supply chain, which operated under the order-up-to level system, was introduced. It was modified from a small frozen-food supply chain in Bangkok, Thailand. It was supplied by a national unit through a nationwide planning system. To simplify the analysis, some conditions were made; the regional warehouse used single-truck round trips to transfer a single type of product to the stores in its region. When the warehouse had insufficient inventories, a rationing rule must be imposed. In this paper, a new inventory rationing rule developed based on the logical purpose was compared against the current practice and other well-known rationing rules in the existing literature. Based on the experiments, it could be concluded that the proposed rationing rule was better and more robust than the other rules.

The current research can be extended for future study in many directions. The first extension would be to consider the multiple product case, which is certainly a more common scenario in practice. In such a case the replenishment conflict will become two dimensional since need of stores and products may be in conflict. A second extension would be to increase the number of trucks, indicating that more than one store can be replenished in a period. Typically, as the number of stores increase then the number of trucks would also increase. Possibly, such a problem can be portioned into several one truck M-store problems. A final extension could be to study the demand correlation between stores. A key assumption in the design of many regional supply chains is that there is negative correlation between stores, and hence the net demand for the region is more steady state.

REFERENCES

Axsäter, S., Kleijn, M., and De Kok, T. G. (2004), Stock rationing in a continuous review two echelon inventory model, *Annals of Operations Research*, **126**, 177-194.

Ayanso, A., Diaby, M., and Nair, S. K. (2006), Inventory rationing via drop-shipping in internet retailing: a sensitivity analysis, *European Journal of Operational Research*, **171**, 135-152.

Diks, E. B. and De Kok, A. G. (1999), Computational results for the control of a divergent *N*-echelon inventory system, *International Journal of Production Economics*, **59**, 327-336.

Diks, E. B. and De Kok, A. G. (1998), Optimal control of a divergent multi-echelon inventory system, *European Journal of Operational Research*, **111**(1), 75-97.

Diks, E. B., De Kok, A. G., and Lagodimos, A. G. (1996) Multi-echelon systems: a service measure perspective, *European Journal of Operational Research*, **95**(2), 241-263.

Diks, E. B. and De Kok, A. G. (1996), Controlling a divergent two-echelon network with transshipments using the consistent appropriate share rationing policy, *International Journal of Production Economics*, **45**, 369-379.

De Kok, A. G. (1990), Hierarchical production planning for consumer goods, *European Journal of Operational Research*, **45**, 55-69.

De Kok, A. G. and Fransoo, J. C. (2003), *Planning Supply Chain Operations, in: Supply Chain Management: Design*, *Coordination and Operation*, In A. G. Dekok and S. C.Graves, eds., *Handbooks in Operations Research and Management Science*, **11**, Elsevier, Amsterda.

Deshpande, V., Cohen, M. A., and Donohue, K. (2003), A threshold inventory rationing policy for service-differentiated classes, *Management Science*, **49**, 683-703.

De Véricourt, F., Karaesman, F., and Dallery, Y. (2002), Optimal stock allocation for a capacitated supply chain, *Management Science*, **48**, 1486-1501.

Dotoli, M., Fanti, M. P., Meloni, C., and Zhou M. (2005), A multi-level approach for network design of integrated supply chains, *International Journal of Production Research*, **43**(20), 4267-4287.

Eppen, G., and Schrage, L. (1981), *Centralized Ordering Policies in a Multi-Warehouse System with Lead Times and Random Demand*, In L. B. Schwarz (Ed.), Multi-Level Production/Inventory Control Systems: Theory and Practice. North-Holland, Amsterdam, 51-67.

Federgruen, A. and Zipkin, P. (1984), Approximations of dynamic multilocation production and inventory problems, *Management Science*, **30**, 69-84.

GÜllÜ, R. and Erkip, N. (1996), Optimal allocation policies in a two-echelon inventory problem with fixed shipment costs, *International Journal of Production Economics*, **46/47**, 311-321.

Ha, A. (1997a), Inventory rationing in a make-to-stock production system with several demand classes and lost sales, *Management Science*, **43**, 1093-1103.

Ha, A. (1997b), Stock-rationing policy for a make-to-stock production system with two priority classes and backordering, *Naval Research Logistics*, **44**, 458-472.

Ha, A. (2000), Stock-rationing in an $M/E_k/1$ make-to-stock queue, *Management Science*, **46**, 77-87.

Huang, B. and Iravani, S. M. R. (2007), Optimal production and rationing decisions in supply chains with information sharing, *Operations Research Letters*, **35**(5), 669-676.

Jackson, P. L. (1988), Stock allocation in a two-echelon distribution system or what to do until your ship comes in, *Management Science*, **34**(7), 880-895.

Kaplan, A. (1969), Stock rationing, *Management Science*, **15**(5), 260-267.

Lagodimos, A. G. (1992), Multi-echelon service models for inventory systems under different rationing policies, *International Journal of Production Research*, **30**, 939-958.

Lee, J. E. and Hong, Y. (2003), A stock rationing policy in a (*s*, *S*)-controlled stochastic production system with 2-phase coxian processing times and lost sales, *International Journal of Production Economics*, **83**, 299-307.

McGavin, E. J., Schwarz, L. B., and Ward, J. E. (1993), Two-interval inventory-allocation policies in a one-warehouse *N*-identical-retailer distribution system, *Management Science*, **39**(9), 1092-1107.

Melchiors, P., Dekker, R., and Kleijn, M. J. (2000), Inventory rationing in an (*s*, *Q*) inventory model with two demand classes and lost sales, *Journal of the Operational Research Society*, **51**(1), 111-122.

Melchiors, P. (2003), Restricted time-remembering policies for the inventory rationing problem, *International Journal of Production Economics*, **81-82**, 461-468.

Monthatipkul, C. (2006), Production/distribution control systems and safety stock policies in supply chains, *A Dissertation Report*, Sirindhorn International Institute of Technology, Thummasat University, Thailand.

Nahmias, S. and Demmy, W. (1981), Operating characteristics of an inventory system with rationing, *Management Science*, **27**(11), 1236-1245.

Schneider, H. and Rinks, D. B. (1991), Empirical study of a new procedure for allocating safety stock in a wholesale inventory system, *International Journal of Production Economics*, **24**, 181-189.

Schneider, H., Rinks, D. B., and Kelle, P. (1995), Power approximations for a two-echelon inventory system using service levels, *Production and Operations Management*, **4**, 381-400.

Seidel, H. P. and De Kok, A. G. (1990), Analysis of stock allocation in a 2-echelon distribution system. *Technical Report 098*, CQM Eindhoven.

Topkis, D. M. (1968), Optimal ordering and rationing policies in a non-stationary dynamic inventory model with *n* demand classes, *Management Science*, **15**(3), 160-176.

Van der Heijden, M. C. (1997), Supply rationing in multi-echelon divergent systems, *European Journal of Operational Research*, **101**, 532-549.

Van der Heijden, M. C., Diks, E. B., and De Kok, A. G. (1997), Stock allocation in general multi-echelon distribution systems with (*R*, *S*) order-up-to-policies, *International Journal of Production Economics*, **49**, 157-174.

Van der Heijden, M. C. (1999), Multi-echelon inventory control in divergent systems with shipping frequencies, *European Journal of Operational Research*, **116**, 331-351.

Van der Heijden, M. C. (2000), Near cost-optimal inventory control policy for divergent networks under fill rate constraints, *International Journal of Production Economics*, **63**, 161-179.

Veinott Jr, A. F. (1965), Optimal policy in a dynamic, single product, non-stationary inventory model with several demand classes, *Operations Research*, **13** (5), 761-778.

Verrijdt, J. H. C. M. and De Kok, A. G. (1995), Distribution planning for a divergent *N*-echelon network without intermediate stocks under service restrictions, *International Journal of Production Economics*, **38**(2, 3), 225-244.

Verrijdt, J. H. C. M. and De Kok, A. G. (1996), Distribution planning for a divergent depotless two-echelon network under service constraints, *European Journal of Operational Research*, **89**(2), 341-354.

Zipkin, P. (1984), On the imbalance of inventories in multi-echelon systems, *Mathematics of Operations Research*, **9**, 402-423.

Zhang, J. (2003), Managing multi-customer service level requirements with a simple rationing policy, *Operations Research Letters*, **31**, 477-482.