# A Robust Optimization Framework for Analyzing Distribution Systems with Transshipment


Mabel Chou

*School of Business, National University of Singapore*

Melvyn Sim

*School of Business, National University of Singapore*

*Singapore-MIT-Alliance*

Kut C. So

*The Paul Merage School of Business, University of California, Irvine*


February 2006


**Abstract**

This paper studies a distribution system consisting of multiple retail locations with transshipment operations among the retailers. Due to the difficulty in computing the optimal solution imposed by the transshipment operations and in estimating shortage cost from a practical perspective, we propose a robust optimization framework for analyzing the impact of transshipment operations on such a distribution system. We demonstrate that our proposed robust optimization framework is analytically tractable and is computationally efficient for analyzing even large-scale distribution systems. From a numerical study using this robust optimization framework, we address a number of managerial issues regarding the impact of transshipment on reducing the costs of the distribution system under different system configurations and retailer characteristics. In particular, we consider two system configurations, line and circle, and study how inventory holding cost, transshipment cost, and demand size and variability affect the effectiveness of transshipment operations for the cases of both homogeneous and non-homogeneous retailers. The managerial insights obtained from our robust optimization framework can help to evaluate the potential benefits when investing in transshipment operations.

# 1 Introduction

Consider a distribution system consisting of multiple retail locations. Demands occur at each retail location, which replenishes its inventory from some central warehouse on a periodic basis. Demands at each retail location are first met from the available inventory at the location. When shortage occurs at one location, the shortage can be covered from available inventory at other retail locations through possible lateral transshipment. The objective is to determine the optimal order quantity of each retail location and the resulting optimal transshipment policy after demands are realized at each period so as to minimize the total expected replenishment costs, inventory holding costs, shortage costs and the transshipment costs among the multiple retail locations during some finite time horizon.

Transshipment, when possible, can be used as one effective way to reduce total inventory and increase service level in a distribution system. Essentially, transshipment allows the distribution system to take advantage of the risk pooling effect to deal with uncertain demands at different retail locations. Excess inventory at one retail location can be used to cover shortage at another location. Physically, one can interpret inventory stocking at each individual location as being "pooled" together to meet the demands at any other location within the distribution system. As such, the use of transshipment provides more flexibility in deploying the available inventory in the system to meet uncertain customer demand. Consequently, transshipment can help to reduce the total system inventory and stockout level at each individual location, at the expense of a higher transportation cost for transshipping the products among the different retail locations. It is interesting yet unclear as to what kind of system configurations and retailer characteristics would benefit most from using transshipment. One objective of this paper is to address a number of managerial issues regarding the impact of transshipment on reducing the costs of the distribution system under different system configurations and retailer characteristics.

Our distribution system with transshipment involves a convoluted decision problem consisting of two basic types of decisions that influence each other throughout the finite time horizon. The first type involves the decision for the optimal order quantity at the retail locations during each replenishment cycle. We refer to this decision as the optimal replenishment policy. The second type involves the decision for transshipping the products among the retail locations after demands are observed and shortages occur at different retail locations. We refer to this decision as the optimal transshipment policy.

The combined optimal replenishment problem with transshipment and stochastic demand is generally difficult to solve. The problem is complicated even for the single period model consisting of a two-stage decision problem, where the transshipment decisions are considered as a recourse action to cover shortage after the replenishment quantity has been selected and uncertain demands have been realized. For a finite time horizon, the optimal replenishment policy generally depends on the replenishment and transshipment decisions as well as realized demands in earlier periods. On the other hand, the optimal transshipment policy, which entails the decision of how much as well as from which location the transshipments should come from, also depends on the replenishment policy and realized demands in earlier periods. Such a convoluted dynamic decision comprising of the optimal transshipment and replenishment decisions throughout the finite time horizon makes this problem analytically intractable. One main contribution of this paper is to develop a tractable model to analyze this combined replenishment problem with transshipment.

In this paper we devise a robust optimization framework for analyzing the above distribution system with transshipment. We demonstrate that our proposed robust optimization framework is analytically tractable and is computationally efficient for analyzing the impact of deploying transshipment on large distribution systems. We further present a set of numerical results to address a number of important managerial issues regarding the use of transshipment in a distribution system. For instance, what is the impact of transshipment on reducing the total inventory in the distribution system? What is the balance between inventory savings and transshipment costs in the system? How does demand variability affect the effectiveness of transshipment? What type of system configurations is most beneficial when transshipment is allowed? Usually, substantial investments might be required to implement transshipment in an existing distribution system such as the need to upgrade the information system for demand and inventory information sharing among different retail locations or the need to build the physical distribution system for transshipping products among the locations. Our numerical results thus help to provide useful information for evaluating the tradeoff between the potential benefits and the required investments due to transshipment.

The paper is organized as follows. In Section 2 we provide a brief review of the literature on transshipment models and the relevant literature on the latest developments on robust optimization. In Section 3 we describe our transshipment problem. In Section 4 we provide a robust optimization model for solving our transshipment problem. In Section 5 we provide some further technical discussions regarding the applicability and limitations of our robust optimization model. In Section 6 we summarize a numerical study using our model and provide some useful managerial insights regarding the effectiveness of transshipment in a distribution system. We conclude the findings of our paper in Section 7. All proofs are given in the Appendix 1.

## 2 Literature Review

Decision models involving inventory replenishment and transshipment are difficult to solve. A number of researchers have studied the special cases where there are only two retail locations or the multiple retail locations are identical in terms of their cost structure in which the transshipment decisions can be much simplified.

Krishnan and Rao (1965) studied the transshipment problems with multiple retail locations with identical cost structure. They showed that the optimal stocking quantities satisfy the equal fractile property. Tagaras (1989) extended Krishnan and Rao's two-location model to allow for different cost structures, and analyzed the pooling effect due to transshipment. His model can also allow for a service constraint on the minimum acceptable fill rates. Taragas and Cohen (1993) later extended the two-location model to allow for positive replenishment leadtimes. With positive replenishment leadtimes, it might be beneficial to hold back stock for future demands, and so it is not necessarily optimal to always transship from the other location (complete pooling) when shortages occur. However, their numerical results showed that complete pooling generally dominates partial pooling. Herer and Rashit (1999) studied the two-location transshipment problem to include fixed and joint replenishment costs, and derived several properties regarding the structure of the corresponding optimal replenishment and transshipment policies. Herer and Tzur (2001) considered a dynamic two-location transshipment problem where demands are deterministic and the objective is to minimize the total replenishment, holding and transshipment costs over a finite horizon. They derived some structural results on the optimal policy and provided a polynomial time algorithm for finding the optimal policy. Rudi et al. (2001) studied a two location model with decentralized decision making. They analyzed the optimal transshipment prices to maximize the total profit. In a recent paper, Dong and Rudi (2004) analyzed how transshipment can benefit a manufacturer and multiple retailers in settings where the manufacturer can serve as a price setter or a price taking. In their model, the multiple retailers have the same cost structure and complete pooling among retailers is assumed. Zhang (2005) extended their results to general demand distributions.

When there are more than two locations in the system and the cost structures are non-identical, the optimal transshipment policy becomes more complex, as one needs to determine from which location, in addition to how many, to transship when a shortage occurs at any location. In general, it is analytically intractable to determine the joint optimal replenishment and transshipment policy. A number of papers studied different heuristic decision rules for lateral transshipment and then evaluated the optimal replenishment policy under these decision rules. This line of research includes the work of Alfredsson and Verrijdt (1999), Archibald, et al. (1997), Axsater (1990, 2003), Dada (1992), Grahovac and Chakravarty (2001), Lee (1987), Minner et. al (2003), and Robinson (1990). Most recently, Wee and Dada (2005) studied the optimal policies for transshipping inventory in a retail network. They focused on the integrated transshipment decisions instead of the interactions among the retailers and the impact of the network structure.

There is a closely related literature where transshipment is allowed in a distribution system periodically as a way to rebalance stock at different locations rather than to cover shortage. This includes the work of Cohen, et al. (1986), Das (1975), Diks and de Kok (1996), Hoadley and Heyman (1977), Jonsson and Silver (1987), and Karmarker and Patel (1977).

Robust optimization has recently gained substantial popularity as a useful methodology for addressing optimization models under uncertainties. The methodology immunizes uncertain mathematical optimization against infeasibility while preserving the tractability of the model; see Ben-Tal, et al. (2004), Ben-Tal and Nemirovski (1998, 1999, 2000), Bertsimas and Sim (2003, 2004, 2005), Bertsimas, et al. (2004), El-Ghaoui and Lebret (1997), and El-Ghaoui, et al. (1998). Robust optimization models have recently been applied to tackle decision problems in dynamic settings where future decisions (recourse variables) depend on the realization of present data including inventory management (Bertsimas and Thiele 2005, Ben-Tal, et al. 2004b), supply contracts (Ben-Tal, et al. 2004a) and project management (Chen, et al. 2005).

# 3 The Transshipment Problem

We consider a distribution system with $m$ retail locations. External demands arrive at the retail locations and each retail location replenishes its inventory from some central warehouse (or supplier). To simplify our analysis such that our model focuses only on the interactions among the retail locations through transshipment, we assume that the central warehouse always has ample inventory to meet the replenishment orders placed by all retail locations.

We consider a finite time period horizon, where the time line of events for each time period is as follows:

1. At the beginning of time period $t$, each retail location $i$ places an order of $x_i^t$ units to the central warehouse.

2. Before the end of time period $t$, the order $x_i^t$ is received at the retail location $i$.

3. Then, demand for period $t$ at each location is realized and each location meets its demand as much possible from its available inventory.

4. Excess inventory at any locations can be deployed, if desired, to meet demand shortage at other locations through lateral transshipment.

5. Any excess inventory and demand shortage at each location are charged for the period. All demand shortages are backlogged.

Our objective is to determine the optimal order quantity of each retail location at each period and the corresponding optimal (dynamic) transshipment policy so as to minimize the total expected replenishment costs, inventory holding costs, shortage costs and transshipment costs for the $m$ retail locations during this finite time horizon.

Throughout this paper, we denote a random variable with the tilde sign such as $\tilde{y}$ and vectors with bold face lower case letters such as $\boldsymbol{y}$. We use $\mathbf{y}'$ to denote the transpose of vector $\mathbf{y}$. Also, denote $y^+ = \max(y, 0)$, $y^- = \max(-y, 0)$, and $\|\boldsymbol{y}\|_2 = \sqrt{\sum y_i^2}$.

We introduce the following notation. Let

$T$ = number of time periods in the model;

$x_i^t$ = order quantity of retail location $i$ at period $t$;

$w_{ij}^t$ = transshipment quantity from retail location $i$ to location $j$ at period $t$, $i > j$;

$\tilde{d}_i^t$ = stochastic exogenous demand at retail location $i$ at period $t$;

$l_i^t$ = inventory position at retail location $i$ at the end of period $t$

$h_i^t$ = unit inventory holding cost at retail location $i$ at period $t$;

$a_i^t$ = unit replenishment cost for retail location $i$ at period $t$;

$b_i^t$ = unit shortage cost at retail location $i$ at period $t$;

$c_{ij}^t$ = unit transshipment cost from retail location $i$ to location $j$ at period $t$.

For each time period $t$, we assume that the transshipment costs are symmetric, i.e., $c_{ij}^t = c_{ji}^t$ for all $1 \leq j < i \leq m$, and allow $w_{ij}^t$ to be negative to represent a positive flow of inventory from location $j$ to location $i$ at the unit cost $c_{ij}$. We set $c_{ij}^t = \infty$ if transshipment between locations $i$ and $j$ are not allowed. For notational convenience, we define the set of retailers as

$$V = \{1, \ldots, m\},$$

and the set of arcs as

$$E = \{(i, j) : i, j \in V, i > j\}.$$

Given the order quantity $x_i^t$, stochastic exogenous demand $\tilde{d}_i^t$ and transshipment flows $w_{ij}^t$, the inventory position at retail location $i$ at the end of period $t$ is given by

$$l_i^t = l_i^{t-1} + x_i^t - \tilde{d}_i^t - \sum_{j:(i,j)\in E} w_{ij}^t + \sum_{j:(j,i)\in E} w_{ji}^t. \tag{1}$$

To ensure that only excess inventory after local demand has been met can be deployed for transshipment at each retail location $i$, we need to impose the constraint

$$(l_i^{t-1} + x_i^t - \tilde{d}_i^t)^+ \geq \sum_{j:(i,j)\in E} w_{ij}^t - \sum_{j:(j,i)\in E} w_{ji}^t. \tag{2}$$

Observe that positive value of $l_i^t$ represents the total amount of inventory at retail location $i$ at the end of period $t$ after meeting demand with transshipment while negative value of $l_i^t$ represents the total amount of demand shortage at retail location $i$ at the end of period $t$. Thus, the total replenishment costs, inventory holding costs, shortage costs and transshipment costs for period $t$ are equal to

$$\sum_{i\in V} \left( a_i^t x_i^t + h_i^t (l_i^t)^+ + b_i^t (l_i^t)^- + \sum_{j:(i,j)\in E} c_{ij}^t |w_{ij}^t| \right). \tag{3}$$

Essentially, transshipments $w_{ij}^t$ at period $t$, when deployed, help to reduce total costs by reducing excess inventory at some locations and demand shortage at others at the end of period $t$, at the expense of extra transshipment costs.

The transshipment problem at each time period $t$ can be considered as a two-stage decision problem where the first decision is to select the optimal replenishment quantity $x_i^t$ based on historical data including the demand information, the replenishment and the transshipment quantities at previous periods. The second decision is to determine the optimal transshipment flows $w_{ij}^t$ after the demands at period $t$ have been realized. Clearly, this decision also depends on historical data as in the first decision.

To capture the dependency of the decisions on the historical data in our model formulation, we let the vector $\tilde{\boldsymbol{d}}^t = (\tilde{d}_1^1, \ldots, \tilde{d}_m^1, \ldots, \tilde{d}_1^t, \ldots, \tilde{d}_m^t)$ denote the concatenation of realized demands at all retailers up to period $t$. In general, the optimal replenishment and transshipment decisions at each period depend on the demand realizations and decisions made in previous periods. Therefore, we use the notation $x_i^t(\tilde{\boldsymbol{d}}^{t-1})$ and $w_{ij}^t(\tilde{\boldsymbol{d}}^t)$ to represent the replenishment and transshipment quantities as a function of the historic demand realizations, with the understanding that the decisions at each period also depend on decisions made at previous periods so that we do not un-necessarily complicate the notations. We also use the notation $l_i^t(\tilde{\boldsymbol{d}}^t)$ to represent the inventory position at location $i$ at the end of period $t$ as a function of $\tilde{\boldsymbol{d}}^t$. To further simplify our exposition, however, we drop the dependency $\tilde{\boldsymbol{d}}^t$ in some of our subsequent discussions where there is no confusion. Thus, the multi-period transshipment problem can be formulated as the following multi-stage non-convex stochastic optimization model:

$$
\begin{aligned}
\min_{\boldsymbol{x}^t(\cdot),\boldsymbol{w}^t(\cdot),\boldsymbol{l}^t(\cdot)} \quad & \mathrm{E}\Big\{\sum_{t=1}^{T}\sum_{i\in V}\Big(a_i^t x_i^t(\tilde{\boldsymbol{d}}^{t-1}) + h_i^t(l_i^t(\tilde{\boldsymbol{d}}^t))^+ + b_i^t(l_i^t(\tilde{\boldsymbol{d}}^t))^- + \sum_{j:(i,j)\in E} c_{ij}^t|w_{ij}^t(\tilde{\boldsymbol{d}}^t)|\Big)\Big\} \\
\text{s.t.} \quad & l_i^t(\tilde{\boldsymbol{d}}^t) = l_i^{t-1}(\tilde{\boldsymbol{d}}^{t-1}) + x_i^t(\tilde{\boldsymbol{d}}^{t-1}) - \tilde{d}_i^t - \sum_{j:(i,j)\in E} w_{ij}^t(\tilde{\boldsymbol{d}}^t) + \sum_{j:(j,i)\in E} w_{ji}^t(\tilde{\boldsymbol{d}}^t) \qquad \forall i \in V, t = 1, ..., T \\
& (l_i^{t-1}(\tilde{\boldsymbol{d}}^{t-1}) + x_i^t(\tilde{\boldsymbol{d}}^{t-1}) - \tilde{d}_i^t)^+ \geq \sum_{j:(i,j)\in E} w_{ij}^t(\tilde{\boldsymbol{d}}^t) - \sum_{j:(j,i)\in E} w_{ji}^t(\tilde{\boldsymbol{d}}^t) \qquad \forall i \in V, t = 1, ..., T \\
& x_i^t(\tilde{\boldsymbol{d}}^{t-1}) \geq 0 \qquad \forall t = 1, \ldots, T.
\end{aligned} \tag{4}
$$

We assume that starting inventory position at each location $i$ at the beginning of the first period, $l_i^0(\tilde{\boldsymbol{d}}^0) = \bar{l}_i$ is given.

In our basic model (4), we assume that the replenishment leadtime is zero, i.e., the order quantity at time $t$, $x_i^t$, is received some time during the period, before the demand of period $t$ is realized. However, we can easily extend our model to allow for positive replenishment leadtime. With a positive replenishment leadtime $L$, we can assume that order decisions are made in periods 1 through $T$ and transshipment decisions are made in periods 1 through $(T + L)$. We also assume that the order quantities for the first $L$ periods are given as initial conditions to the transshipment model. The model then extends in a straightforward manner, with the order quantity $x_i^t$ now received during period $(t + L)$. To simplify our exposition, we shall focus our discussions for the zero leadtime case only.

We can also consider the infinite time horizon case with the objective of minimizing the total expected costs per period. Under the assumption that all model parameters are stationary, it is then reasonable to deploy a base-stock replenishment policy at each retail location such that the order quantity at each location at each time period is equal to the difference between the base-stock level and the inventory position. In this case, our proposed model for the single-period case can be used to find the optimal base-stock level and corresponding optimal transshipment policy. However, it is unclear as to whether the base-stock policy is optimal under our modeling assumptions.

In practice, the shortage costs $b_i^t$ are often hard to estimate, and it is common practice to manage the distribution system to achieve some minimum service requirement. Therefore, we replace the shortage cost in the objective (4) by a chance constraint, which specifies some minimum stockout probability that demand will not be fulfilled at any of the retail locations. Specifically, the transshipment problem can now be formally stated as the following stochastic programming model:

$$
\begin{array}{rll}
Z^* = \min\limits_{\boldsymbol{x}^t(\cdot),\boldsymbol{w}^t(\cdot),\boldsymbol{l}^t(\cdot)} & \mathrm{E}\left\{\sum\limits_{t=1}^{T}\sum\limits_{i\in V}\left(a_i^t x_i^t(\tilde{\boldsymbol{d}}^{t-1}) + h_i^t(l_i^t(\tilde{\boldsymbol{d}}^t))^+ + \sum\limits_{j:(i,j)\in E} c_{ij}\left|w_{ij}(\tilde{\boldsymbol{d}}^t)\right|\right)\right\} & \\
\text{s.t.} & l_i^t(\tilde{\boldsymbol{d}}^t) = l_i^{t-1}(\tilde{\boldsymbol{d}}^{t-1}) + x_i^t(\tilde{\boldsymbol{d}}^{t-1}) - \tilde{d}_i^t - \sum\limits_{j:(i,j)\in E} w_{ij}^t(\tilde{\boldsymbol{d}}^t) + \sum\limits_{j:(j,i)\in E} w_{ji}^t(\tilde{\boldsymbol{d}}^t) & \forall i \in V, t = 1, ..., T \\
& \Pr\left(l_i^t(\tilde{\boldsymbol{d}}^t) \geq 0,\ \forall i \in V\right) \geq 1 - \epsilon_t & \forall t = 1, ..., T \\
& x_i^t(\tilde{\boldsymbol{d}}^{t-1}) \geq 0 & \forall t = 1, \ldots, T.
\end{array}
\tag{5}
$$

Observe that

$$
\begin{aligned}
& \Pr\left(\bigcap_{i\in V}\left\{(l_i^{t-1}(\tilde{\boldsymbol{d}}^{t-1}) + x_i^t(\tilde{\boldsymbol{d}}^{t-1}) - \tilde{d}_i^t)^+ \geq \sum_{j:(i,j)\in E} w_{ij}^t(\tilde{\boldsymbol{d}}^t) - \sum_{j:(j,i)\in E} w_{ji}^t(\tilde{\boldsymbol{d}}^t)\right\}\right) \\
\geq\ & \Pr\left(\bigcap_{i\in V}\left\{(l_i^{t-1}(\tilde{\boldsymbol{d}}^t) + x_i^t(\tilde{\boldsymbol{d}}^{t-1}) - \tilde{d}_i^t) \geq \sum_{j:(i,j)\in E} w_{ij}^t(\tilde{\boldsymbol{d}}^t) - \sum_{j:(j,i)\in E} w_{ji}^t(\tilde{\boldsymbol{d}}^t)\right\}\right) \\
=\ & \Pr\left(\bigcap_{i\in V}\left\{l_i^t(\tilde{\boldsymbol{d}}^t) \geq 0\right\}\right) \\
\geq\ & 1 - \epsilon_t
\end{aligned}
$$

Hence, when $\epsilon_t$ is small, the joint chance constraint in the model (5) implies that violation of (2) will be rare at period $t$ and we may ignore constraint (2) in the above stochastic programming formulation for computational purposes, as constraint (2) is not convex. We can interpret the parameter $(1 - \epsilon_t)$ as the probability that the transshipment system can meet all customers' demands. For simplicity, we shall assume further that the service requirement $\epsilon_t$ is the same for all time period $t$, denoted simply by $\epsilon$, in our subsequent discussions.

The stochastic programming model given in (5) is difficult to solve. For the past fifty years, sampling approximation seems to be the sole method for tackling multistage stochastic optimization problems. (Very recently, Shapiro and Nemirovski (2004) and Dyer and Stougie (2005) established the computational intractability of multistage stochastic optimization problems.) Furthermore, the stochastic programming model (5) does not conform to a standard multi-stage stochastic optimization model as it encompasses recourse variables, i.e., the "wait-and-see" transshipment decisions with joint chance constraints, which further makes the problem analytically intractable and computationally challenging. To the best of our knowledge, there does not exist any stochastic optimization methodology that is capable of solving the above stochastic programming model. In view of these difficulties, we next develop a robust optimization framework to tackle the stochastic optimization model (5).

## 4 A Robust Optimization Framework

The robust optimization framework for the general multiple-period model involves rather complex notation. To simplify the exposition for the development of our robust optimization framework, we shall focus on the single-period model (i.e., $T = 1$) for the remainder of this paper. In this case, we can drop the superscript $t$ in our notation. The robust optimization framework for the general multiple-period model can be extended in a straightforward manner, as presented in Appendix 2 for any interested reader.

To apply the robust optimization methodology, we assume that the uncertain demand at each retail location $\tilde{\boldsymbol{d}}$ is linearly dependent on a set of independent random variable $\tilde{\boldsymbol{z}}$, $k = 1, 2, ..., N$, as follows:

$$\boldsymbol{d}(\tilde{\boldsymbol{z}}) = \boldsymbol{d}^0 + \sum_{k=1}^{N} \boldsymbol{\Delta d}^k \tilde{z}_k, \tag{6}$$

where $\boldsymbol{d}^0$ is the mean demand, $\boldsymbol{\Delta d}^k$ is a direction of demand perturbation, and each random variable $\tilde{z}_k$ has mean zero with support in $[-\underline{z}_k, \bar{z}_k]$ and $\underline{z}_k, \bar{z}_k > 0$. Without loss of generality, we assume that the standard deviations of $\tilde{z}_k$ are normalized to one. This representation is useful for relating multivariate random variables. For instance, if the uncertain demands at the retail locations are independent, we have $\boldsymbol{\Delta d}^k = \boldsymbol{e}_k$, where $\boldsymbol{e}_k$ is a unit vector at the $k$th element. On the other hand, this assumption can also allow for the uncertain demands at various retail locations to be collectively dependent of certain common factors such as market conditions, in which case, the demands will be correlated.

We refer to $\{\tilde{z}_k\}$ as the primitive uncertainties. For each primitive uncertainty $z_k$, we define $\alpha_k$ as its forward deviation given by

$$\alpha_k = \sup_{\theta > 0} \left\{ \sqrt{2 \ln(\mathrm{E}(\exp(\theta \tilde{z}_j)))/\theta^2} \right\} \tag{7}$$

and $\beta_k$ as its backward deviation given by

$$\beta_k = \sup_{\theta > 0} \left\{ \sqrt{2 \ln(\mathrm{E}(\exp(-\theta \tilde{z}_j)))/\theta^2} \right\}. \tag{8}$$

We assume that the transshipment flows $w_{ij}$ are also linearly dependent on the primitive uncertainties $\tilde{\boldsymbol{z}}$, i.e.,

$$w_{ij}(\tilde{\boldsymbol{z}}) = w_{ij}^0 + \sum_{k=1}^{N} w_{ij}^k \tilde{z}_k \quad \forall (i,j) \in E. \tag{9}$$

With this assumption, the constraint

$$l_i(\tilde{\boldsymbol{z}}) = \bar{l}_i + x_i - d_i(\tilde{\boldsymbol{z}}) - \sum_{j:(i,j)\in E} w_{ij}(\tilde{\boldsymbol{z}}) + \sum_{j:(j,i)\in E} w_{ji}(\tilde{\boldsymbol{z}}) \;\; \forall i \in V, t = 1, ..., T$$

implies that the inventory position is also one of linear decision rule as follows:

$$l_i(\tilde{\boldsymbol{z}}) = l_i^0 + \sum_{k=1}^{N} l_i^k \tilde{z}_k \qquad \forall i \in V. \tag{10}$$

For notational convenience, we denote the vectors $\boldsymbol{w}_{ij} = (w_{ij}^1, \ldots, w_{ij}^N)$ and $\boldsymbol{l}_i = (l_i^1, \ldots, l_i^N)$ so that

$$\begin{aligned} w_{ij}(\tilde{\boldsymbol{z}}) &= w_{ij}^0 + \boldsymbol{w}_{ij}'\tilde{\boldsymbol{z}} \\ l_i(\tilde{\boldsymbol{z}}) &= l_i^0 + \boldsymbol{l}_i'\tilde{\boldsymbol{z}}. \end{aligned}$$

This assumption is also known as the linear decision rule, which can be viewed as the first order approximation of the optimal solution. As will be shown later, the linear decision rule enables us to design a tractable robust optimization approach for our two-stage stochastic optimization model.

We first need two preliminary results. The first result is used to show that our robust solution satisfies the joint chance constraint given in the stochastic programming model (5).

**Proposition 1** *Let $(y^0, y^1, ..., y^N)$ be a feasible solution satisfying the following set of constraints*

$$\left\{ \begin{array}{ll} y^0 \geq \Omega p^0 + \boldsymbol{q}'\underline{\boldsymbol{z}} + \boldsymbol{r}'\bar{\boldsymbol{z}} & \\ p^k \geq \beta_k(y^k - q^k + r^k) & \forall k \in \{1, \ldots, N\} \\ p^k \geq -\alpha_k(y^k - q^k + r^k) & \forall k \in \{1, \ldots, N\} \\ \|\boldsymbol{p}\|_2 \leq p^0 & \\ p^0 \geq 0, \boldsymbol{p} \in \Re^N, \boldsymbol{q}, \boldsymbol{r} \in \Re^N_+. & \end{array} \right\}. \tag{11}$$

*Then,*

$$\mathrm{P}\left( y^0 + \sum_{k=1}^{N} y^k z_k < 0 \right) \leq \exp(-\Omega^2/2).$$

The second result is used to provide an upper bound for the value $\mathrm{E}\left(|w_{ij}(\tilde{\boldsymbol{z}})|\right)$ under the linear decision rule for $w_{ij}(\tilde{\boldsymbol{z}})$ as defined in (9).

**Theorem 1** *Let $\tilde{z}_1, \ldots, \tilde{z}_j$ be uncorrelated random variables with zero mean, unit standard deviation, and support in $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$. Then,*

$$\mathrm{E}\left( |y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}| \right) \leq G(y^0, \boldsymbol{y}),$$

*where*

$$G(y^0, \boldsymbol{y}) = \min_{\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} \geq \boldsymbol{0}} \left\{ (\boldsymbol{s} + \boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t} + \boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(y^0 + (\boldsymbol{s} - \boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t} - \boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{y} - \boldsymbol{s} + \boldsymbol{t} + \boldsymbol{u} - \boldsymbol{v}\|_2^2} \right\}.$$

We next propose a robust optimization model for solving our transshipment problem and establish the relationship between this robust optimization model and the stochastic programming model (5). In particular, we show that the optimal solution of our robust optimization model is a feasible solution to the stochastic optimization model. Consider the following optimization problem:

$$
\begin{array}{lll}
Z_r^* = & \min & \sum_{i\in V}\left(a_i x_i + h_i f_i + \sum_{j:(i,j)\in E} c_{ij} g_{ij}\right) \\
& \text{s.t.} & l_i^0 = \bar{l}_i + x_i - d_i^0 - \sum_{j:(i,j)\in E} w_{ij}^0 + \sum_{j:(j,i)\in E} w_{ji}^0 \qquad \forall i \in V \\
& & l_i^k = -\Delta d_i^k - \sum_{j:(i,j)\in E} w_{ij}^k + \sum_{j:(j,i)\in E} w_{ji}^k \qquad \forall k \in \{1,\dots,N\}, i \in V \\
& & l_i^0 \geq \sqrt{-2\ln(\epsilon/m)}\, p_i^0 + \boldsymbol{q}_i{}'\underline{\boldsymbol{z}} + \boldsymbol{r}_i{}'\bar{\boldsymbol{z}} \qquad \forall i \in V \\
& & p_i^k \geq \beta_k (l_i^k - q_i^k + r_i^k) \qquad \forall k \in \{1,\dots,N\}, i \in V \\
& & p_i^k \geq -\alpha_k (l_i^k - q_i^k + r_i^k) \qquad \forall k \in \{1,\dots,N\}, i \in V \\
& & \|\boldsymbol{p}_i\|_2 \leq p_i^0 \qquad \forall i \in V \\
& & \boldsymbol{x} \geq 0 \\
& & l_i^0 \in \Re, \quad x_i, p_i^0 \in \Re, \quad \boldsymbol{l}_i, \boldsymbol{p}_i \in \Re^N, \quad \boldsymbol{q}_i, \boldsymbol{r}_i \in \Re_+^N \qquad \forall i \in V \\
& & g_{ij} \geq (\boldsymbol{s}_{ij} + \boldsymbol{u}_{ij})'\bar{\boldsymbol{z}} + (\boldsymbol{t}_{ij} + \boldsymbol{v}_{ij})'\underline{\boldsymbol{z}} + \\
& & \quad \sqrt{(w_{ij}^0 + (\boldsymbol{s}_{ij} - \boldsymbol{u}_{ij})'\bar{\boldsymbol{z}} + (\boldsymbol{t}_{ij} - \boldsymbol{v}_{ij})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{w}_{ij} - \boldsymbol{s}_{ij} + \boldsymbol{t}_{ij} + \boldsymbol{u}_{ij} - \boldsymbol{v}_{ij}\|_2^2)} \qquad \forall (i,j) \in E \\
& & g_{ij} \in \Re, \quad \boldsymbol{s}_{ij}, \boldsymbol{t}_{ij}, \boldsymbol{u}_{ij}, \boldsymbol{v}_{ij} \in \Re_+^N, \quad \boldsymbol{w}_{ij} \in \Re^N \qquad \forall (i,j) \in E \\
& & f_i \geq \frac{1}{2}\Big\{ l_i^0 + (\hat{\boldsymbol{s}}_i + \hat{\boldsymbol{u}}_i)'\bar{\boldsymbol{z}} + (\hat{\boldsymbol{t}}_i + \hat{\boldsymbol{v}}_i)'\underline{\boldsymbol{z}} + \\
& & \quad \sqrt{(l_i^0 + (\hat{\boldsymbol{s}}_i - \hat{\boldsymbol{u}}_i)'\bar{\boldsymbol{z}} + (\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{v}}_i)'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{l}_i - \hat{\boldsymbol{s}}_i + \hat{\boldsymbol{t}}_i + \hat{\boldsymbol{u}}_i - \hat{\boldsymbol{v}}_i\|_2^2)} \Big\} \qquad \forall i \in V \\
& & f_i \in \Re, \quad \hat{\boldsymbol{s}}_i, \hat{\boldsymbol{t}}_i, \hat{\boldsymbol{u}}_i, \hat{\boldsymbol{v}}_i \in \Re_+^{\hat{N}} \qquad \forall i \in V
\end{array}
\tag{12}
$$

**Theorem 2** *For $T = 1$, under Assumptions (6) and (9), the optimal solution $\boldsymbol{x}$ of the robust optimization model (12) is a feasible solution to the stochastic optimization model (5). Moreover, the optimal objective value of (12), $Z_r^*$, is an upper bound for the optimal objective value of (5), $Z^*$.*

The attractiveness of our robust optimization model (12) are two folded. First, the demand uncertainties in our model are based on mild distributional assumption such as the support and measurement of deviations. We do not make any specific assumptions on the form of the demand distributions.

Second, our robust optimization models is in the form of linear optimization and second order cone optimization problems (SOCP) which are computationally attractive. This allows us to leverage on the state-of-the-art LP and SOCP solvers (for instance SiDuMe, SDPT3, MOSEK and CPLEX 9.1), which are increasingly becoming more powerful, efficient and robust.

# 5 Applicability and Limitations

We develop a stochastic programming model (5) for analyzing the transshipment problem that involves multiple stages of decision making. However, even for $T = 1$, the recourse variables make the stochastic programming model intractable and there does not exist any effective optimization methodology that is capable of solving the stochastic programming model. Due to these difficulties, we propose a robust optimization model (12) that is analytically tractable and computationally efficient to solve. We show in Theorem 2 that the optimal solution of our robust optimization model provides a feasible solution to the stochastic programming model (5). We also derive an upper bound for the optimal value of (5). More importantly, our formulation allows for the applicability of several powerful optimization softwares for solving the proposed robust optimization model.

However, there remains several technical issues regarding how well the solution of the robust optimization model can serve as an approximation to the optimal solution of the underlying stochastic programming model. In this section, we discuss some of these technical issues regarding the applicability as well as the limitations of our proposed robust optimization model.

## 5.1 Joint Chance Constraints

Robust optimization technique cannot directly address joint chance constraints problems such as

$$\mathrm{P}\left( y_i^0 + \sum_{k=1}^{N} y_i^k \tilde{z}_k \geq 0, \forall i = 1, \ldots, m \right) \geq 1 - \epsilon. \tag{13}$$

However, we apply Bonferroni's inequality or union bound to derive a sufficient condition that ensures feasibility in (13) as follows:

$$\left\{ \begin{array}{ll} y_i^0 \geq \Omega p_i^0 + \boldsymbol{q}_i' \underline{\boldsymbol{z}} + \boldsymbol{r}_i' \bar{\boldsymbol{z}} & \\ p_i^k \geq \beta_k (y_i^k - q_i^k + r_i^k) & \forall k \in \{1, \ldots, N\} \\ p_i^k \geq -\alpha_k (y_i^k - q_i^k + r_i^k) & \forall k \in \{1, \ldots, N\} \\ \|\boldsymbol{p}_i\|_2 \leq p_i^0 & \\ p_i^0 \geq 0, \boldsymbol{p}_i \in \Re^N, \boldsymbol{q}_i, \boldsymbol{r}_i \in \Re_+^N & \end{array} \right\},$$

| $m$ | $\rho$ |
|---|---|
| 1 | 1 |
| 10 | 1.155 |
| 100 | 1.219 |
| 1000 | 1.414 |
| 10000 | 1.528 |
| 100000 | 1.633 |

Table 1: Conservative measure $\rho$ with $\epsilon = 0.001$.

where, $\Omega_1, \ldots, \Omega_m$ satisfy

$$\sum_{i=1}^{m} \exp(-\Omega_i^2/2) \leq \epsilon. \tag{14}$$

In the context of robust optimization, the parameter $\Omega_i$ is also known as the budget of uncertainty associated with the $i$th constraint. Unfortunately, it is difficult to determine the optimal choice of the budget of uncertainty that satisfies (14). It is thus reasonable to choose a budget allocation so that in the worst case, we are not too far off from the optimal budget allocation. Thus, one approach is to minimize the maximum budget, i.e., $\min\{\max_i\{\Omega_i\}\}$, subject to the constraint of (14). Hence, we choose a budget allocation of

$$\Omega_i = \sqrt{-2\ln(\epsilon/m)}, \quad \forall i = 1, \ldots, m. \tag{15}$$

Observe that $\Omega_i \geq \sqrt{-2\ln(\epsilon)}$ in order to satisfy (14). Therefore, we can use the relative sizes of uncertainty budget proposed by Ben-Tal and Nemirovski (1998),

$$\rho = \frac{\sqrt{-2\ln(\epsilon/m)}}{\sqrt{-2\ln(\epsilon)}},$$

as a conservation measure of how much optimality we might lose in using (15) instead of the optimal allocation of uncertainty budgets. Table 1 shows that $\rho$ changes marginally with $m$, suggesting that even for very large-scale distribution systems, we are not losing too much from the solutions with optimally allocated budgets, $\Omega_i$ .

## 5.2 Linear Decision Rule

To apply the robust optimization framework, we need to restrict a linear decision rule on the transshipment policy, so that the size of the model does not explode exponentially with the number of periods.

Such decision rule can be viewed as first order approximation of the expected future costs, so that we can determine the first stage or 'here-and-now' decision. In practice, we do not use the decision rule as the respond actions in subsequent stages. Instead, we resolve subsequent stages upon realizations of uncertainties at earlier stages. This strategy is similar to "rolling horizon strategy" in approximate dynamic programming.

### 5.3 Bounding $\mathrm{E}(|w_{ij}(\tilde{\boldsymbol{z}})|)$ and $\mathrm{E}((l_i(\tilde{\boldsymbol{z}}))^+)$

Under the linear decision rule (9), and since the primitive uncertainties are uncorrelated with normalized standard deviation, the second-stage transshipment decision, $w_{ij}(\tilde{\boldsymbol{z}})$, is a random variable with mean $w_{ij}^0$ and standard deviation $\|\boldsymbol{w}_{ij}\|_2$. Hence, using the well known result of (25), it is trivial to establish the following bound:

$$\mathrm{E}(|w_{ij}(\tilde{\boldsymbol{z}})|) \leq \sqrt{(w_{ij}^0)^2 + \|\boldsymbol{w}_{ij}\|_2^2}. \tag{16}$$

However, the above simple bound can perform very poorly. For instance, suppose that

$$w_{ij}(\boldsymbol{z}) \geq 0,\ \forall \boldsymbol{z} \in \mathcal{W} \quad \text{or} \quad w_{ij}(\boldsymbol{z}) \leq 0,\ \forall \boldsymbol{z} \in \mathcal{W}. \tag{17}$$

Then, it is obvious that

$$\mathrm{E}(|w_{ij}(\tilde{\boldsymbol{z}})|) = |w_{ij}^0|,$$

and the above simple bound would be rather weak. Instead, we incorporate support information of the underlying primitive uncertainties $\tilde{\boldsymbol{z}}$ to derive a better bound as given in Theorem 1. Indeed, the following result shows that the bound in Theorem 1 performs better than the simple bound of (16) and actually achieves the exact value under the condition of (17).

**Theorem 3** *Let $\tilde{z}_1, \ldots, \tilde{z}_j$ be uncorrelated random variables with zero mean, unit standard deviation, and support in $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$.*

**(a)**

$$G(y_0, \boldsymbol{y}) \leq \sqrt{y_0^2 + \|\boldsymbol{y}\|_2^2},$$

*where*

$$G(y_0, \boldsymbol{y}) = \min_{\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v} \geq \boldsymbol{0}} \left\{ (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(y_0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{y} - \boldsymbol{s} + \boldsymbol{t} + \boldsymbol{u} - \boldsymbol{v}\|_2^2} \right\}.$$

**(b)** *Suppose that*

$$y(\tilde{\boldsymbol{z}}) \geq 0,\ \forall \boldsymbol{z} \in \mathcal{W} \quad \text{or} \quad y(\tilde{\boldsymbol{z}}) \leq 0,\ \forall \boldsymbol{z} \in \mathcal{W}.$$

*Then,*

$$\mathrm{E}\left(|y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}|\right) = G(y_0, \boldsymbol{y}) = |y_0|.$$

As for the bound of $\mathrm{E}((l_i(\tilde{\boldsymbol{z}}))^+)$, we observe that under the linear decision rule of (10)

$$\mathrm{E}((l_i(\tilde{\boldsymbol{z}}))^+) = \frac{1}{2}\mathrm{E}(l_i(\tilde{\boldsymbol{z}}) + |l_i(\tilde{\boldsymbol{z}})|) = \frac{1}{2}\left(l_i^0 + \mathrm{E}(|l_i(\tilde{\boldsymbol{z}})|)\right). \tag{18}$$

Hence, we can adopt the same strategy for bounding $\mathrm{E}((l_i(\tilde{\boldsymbol{z}}))^+)$ by applying the bound of $\mathrm{E}(|l_i(\tilde{\boldsymbol{z}})|)$ in equality (18).

# 6 Computational Studies

In this section we present a set of results from our numerical experiments to demonstrate the applicability of our robust optimization framework for studying the transshipment problem discussed in this paper. Furthermore, our numerical results provide interesting insights for understanding the extent of how transshipment can help to reduce the average inventory and total cost in a network of retail locations.

Essentially, transshipment can be viewed as an alternate way to inventory for tackling the underlying demand uncertainties in the retail network. It is thus interesting to compare the tradeoff between inventory and transshipment costs at each retailer location so as to understand the potential benefits of transshipment. To isolate this effect, we shall ignore the replenishment cost from the warehouse, i.e., $a_i = 0$, in our numerical study, as the percentage savings due to transshipment would depend on the relative magnitude of these replenishment costs. As such, our results provide a clear impact of the transshipment operations on inventory savings. Without loss of generality, we also assume that the initial inventory level at each retailer location is zero. We consider both homogeneous and non-homogeneous retailers and study how inventory holding cost, transshipment cost, and demand size and variability affect the effectiveness of transshipment operations.

## 6.1 Homogeneous Retailers

In this section, we examine the cases when all retailers in the network are homogeneous in terms of their holding cost and demand parameters. We use the following numerical example as the base case for our discussions here. For this base case, we consider a single period problem with a network of 10 retailers ($m = 10$); all 10 retailers have the same model parameters with unit holding cost $h_i = 1$, and external demand $D_i$ following $Uniform(50, 200)$ distribution. We consider two simple retailer network configurations – the line and circle configurations – with unit transshipment cost between adjacent locations $c_{i,i+1} = 0.1$; see Figure 1.

**Insert Figure 1 about here**

Also, we set the system service level $\epsilon = 0.1$. For the base case, the total expected inventory and transshipment cost of the robust solution for the line configuration is equal to 479 and the total expected inventory and transshipment cost of the robust solution for the circle configuration is equal to 467. For comparison purposes, we also compute the total expected inventory cost for the system without transshipment. To do that, we assume that each retailer location will achieve the same individual service level $\epsilon_i$ and the overall system service level $\epsilon$ will be achieved. That is, we set $(1 - \epsilon_i)^n = \epsilon$. With this simplifying assumption, we can easily calculate the expected ending inventory at the period, $l_i$, to achieve the service level $\epsilon_i$ without transshipment for each retailer. For the base case, the total inventory cost for the system without transshipment is equal to 734. Thus, the line and circle configurations with transshipment represent a total cost reduction of 34.7% and 36.4% over that of the system without transshipment, respectively. Recall that our robust optimization model provides an upper bound of the actual optimal total cost of the distribution system with transshipment. As such, the total cost reduction from the line and circle configurations is at least 34.7% and 36.4% over that of the system without transshipment, respectively.

**Insert Figures 2 and 3 about here**

Figure 2 and Figure 3 show the stocking quantity $x_i$ and the expected ending inventory $E[l_i]$ respectively for each retailer under the two configurations with transshipment and the system without transshipment for the base case. We see that the stocking quantity shows the same pattern as the expected ending inventory, as the demand distribution is assumed to be identical among all retail locations in this case. First of all, the results demonstrate that transshipment reduces both the stocking quantity and the expected ending inventory at each retailer. Second, under the circle configuration, both the stocking quantities and the expected ending inventory at all retailers are the same. This is intuitive since transshipment costs are symmetric among all locations. In contrast, under the line configuration, both the stocking quantities and the expected ending inventory are higher at the end locations than those at the center locations. This can be explained by the fact that there are more nearby retailers for the center locations than the end locations so that, on average, it is likely to be more expensive to transship to the end locations from the other retailers than to the center locations. As a result, under the line configuration, the inventory position at the end locations in the beginning of the period is brought up to a higher level and therefore the expected ending inventory of the period is also higher. We observe this same pattern in our other numerical experiments.

Our next set of numerical results illustrates how the different model parameters can affect the robust solutions for both configurations with transshipment. To do so, we vary the value of one model parameter at a time while keeping all other model parameters at their base values. To measure the impact of transshipment, we compute the total cost and total ending inventory of the system for each configuration with transshipment as a percent savings over the corresponding total cost and total ending inventory of the system without transshipment.

**Insert Figure 4 about here**

Figure 4 illustrates the impact of the number of retailers on the system performance, with the value of $m$ ranging from 5 to 40. The results show that the percent savings in total cost and total ending inventory for both the line and circle configurations increase as the number of retailers increases, demonstrating a larger pooling effect due to transshipment among more retailers. However, all these curves flatten out, which show diminishing benefits as the number of retailers increases. Another interesting observation is that the circle configuration always gives large percent savings than those for the line configuration. A closer examination of the robust solutions for both configurations suggests that both the total transshipment cost and the total ending inventory are higher under the line configuration than those under the circle configuration. This is apparently due to the fact that transshipments towards the end locations under the line configuration tend to be more expensive and consequently, the total transshipment cost is higher under the line configuration, together with a higher total ending inventory (mostly due to the higher ending inventory level at or near the end locations). We observe the same pattern in our other numerical results.

**Insert Figure 5 about here**

Figure 5 illustrates the impact of unit transshipment cost on the system performance, with the value of $c_{i,i+1}$ ranging from 0 to 1. Observe that the percent savings in total cost under both configurations decrease as $c_{i,i+1}$ increases. Intuitively, the total transshipment cost becomes higher as $c_{i,i+1}$ increases. Also, as transshipments become more expensive, the robust solution accordingly sets a higher inventory level to reduce the amount of transshipments in the system. As a result, the total ending inventory under both configurations increases as $c_{i,i+1}$ increases, resulting in a decrease in percent savings in inventory. As the unit transshipment costs become too high, transshipment is no longer beneficial and the system performance coincides with that of the system without transshipment. This occurs when $c_{i,i+1} > 0.72$ under the line configuration, and when $c_{i,i+1} > 0.78$ under the circle configuration. Also, the case with $c_{i,i+1} = 0$ corresponds to the situation where all stock are essentially centralized to meet demands at any retailer, under which the corresponding savings represent the maximum possible benefits due to the pooling effect of centralized stock. We note that since the relative performance among the systems depends only on the ratio of $h_i/c_{i,i+1}$, the same observations also apply, in the opposite manner, to the unit holding cost parameter $h_i$.

**Insert Figure 6 about here**

Figure 6 illustrates the impact of demand variability on the system performance. Specifically, we use the range of the support of the uniform distribution, denoted by $R$, as our measure of demand variability, and vary this range $R$ from 30 to 250 while keeping the average demand at (50+200)/2=125. Figure 6 plots the percent savings for different values of $R$. The results show that the percent savings in total cost and total inventory under both configurations remain constant when $R$ changes. This can be explained by the fact that with zero replenishment cost ($a_i = 0$), the expected ending inventory level, the transshipment cost, and the cost when there is no transshipment all scale linearly with $R$. Hence, the percent cost savings remain the same.

**Insert Figure 7 about here**

Figure 7 illustrates the impact of required service level on the system performance, with the value of $(1-\epsilon)$ ranging from 0.7 to 0.995. The results show that the percent savings decreases as the required service level increases. Furthermore, the decrease in percent savings becomes increasingly rapid when the service level exceeds .9. Apparently, a much higher inventory is needed to achieve a very high service level with transshipment, which makes the benefit due to transshipment relative to the total inventory cost becomes less significant. Consequently, the savings due to transshipment, as a percent of the total cost, decrease rapidly when the service level is high.

To conclude our discussions on the homogeneous retailers case, we make two other general observations for our numerical results. First, the percent savings in both total cost and total ending inventory under the circle configuration are always higher than those under the line configuration. This is apparently due to the fact that it is generally more expensive to transship to the end locations under the line configuration. As a result, the total cost is higher under the line configuration. Furthermore, less transshipments are used under the line configuration, resulting in a higher total inventory to meet the required service level. Second, the savings in total cost and total inventory are substantial for low unit transshipment cost (or equivalently, relatively high unit holding cost) and large number of retailers in the system. This implies that transshipment can provide significant savings in a large network of retail locations, especially when the cost to transship a unit is significantly lower than that of holding the unit in the inventory.

## 6.2 Non-homogeneous Retailers

So far, we only look into the cases when all retailers in the network are homogeneous in terms of their holding cost and demand parameters. The next set of numerical results further investigates the impact of transshipment on a more general network with non-homogeneous retailers. To isolate some single factors to better understand the model behavior, we use the same base case network with 10 retailers, where only one of the 10 retailers is different, while keeping the other 9 retailers the same as in the base case. Specifically, we study how the change in the holding cost or demand parameters of one different retailer could affect the performance of the systems with transshipment. Again, we consider the two basic network configurations – line and circle. For the circle configuration, it is irrelevant as to where this different retailer is located since the cost structure is symmetric. However, the location of this different retailer will affect the system performance for the line configuration. Therefore, we consider three different locations of this particular retailer, at locations 1, 3, and 5, respectively. In each scenario, we again calculate the percent savings in total cost for the system with transshipment over that of the corresponding system without transshipment.

### 6.2.1 Difference in Holding Cost

Figure 8 illustrates the impact of the holding cost of this different retailer, as vary its value from 0.3 to 3.0 while keeping the unit holding cost of all other retailers at one.

**Insert Figure 8 about here**

There are several interesting observations. First, the results show that the percent savings is the lowest with transshipment when this retailer has the same holding cost as the other retailers, i.e., all retailers are homogeneous. Second, in all cases, the percent savings is always higher under the circle configuration than all three locations under the line configuration. Third, more importantly, the percent savings increases sharply with transshipment when the holding cost of this different retailer becomes substantially lower than that of the other retailers. In this case, the system will take advantage of the lower holding cost of this retailer to stock inventory for **all** other retailers and use transshipments to meet demands for these retailers. Furthermore, under the line configuration, transshipment yields the maximum benefit when this "centralized" stock is located in the middle of the retailer network (location 5), as the required transshipment cost is minimized. We provide a graphical illustration of the corresponding robust solutions for the circle and line (different retailer at location 1) configurations when the unit holding cost $h_1 = 0.9$ in Figures 9 and 10. In both figures, the size of each node represents the size of the ending inventory level of the individual retailer and the thickness of each arc represents the transshipment amount between the corresponding adjacent retailers.

**Insert Figures 9 and 10 about here**

On the other hand, the percent savings increases only moderately with transshipment when the holding cost of this different retailer is higher than that of the other retailers. In this case, due to the higher holding cost, the system will now stock inventory at other nearby retailers to meet the demand of this one particular retailer. Actually, when the holding cost of this different retailer exceeds the sum of the holding cost and transshipment cost from its immediate neighboring retailers ($h \geq 1.1$ in this example), this retailer will no longer hold any inventory and will always use transshipments from other retailers to meet its demand. As a result, any further increase in its unit holding cost will have no impact on the total cost of the system. The increase in percent savings beyond this point is purely due to the corresponding increase in the total cost of the system without transshipment. Observe that the specific location of the different retailer under the line configuration does not has significant impact, as the holding cost at any of these corresponding immediate neighboring retailers are the same.

The above results suggest that transshipment provides the maximum benefit when the holding costs of the retailers are different, with few retailers with lower holding cost in more centralized locations such that the transshipment cost from these retailers to the others is low. In this case, the majority of other high cost retailers can take advantage of transshipment by stocking their inventory at these low cost retailers.

### 6.2.2 Difference in Demand

We next illustrate the impact of the demand size and demand variability of the different retailer, respectively, on the system performance. To isolate the effect of the demand size of this particular retailer on the benefits of transshipment, we first set the average demand of this retailer as a proportion of the average demand of the other (homogeneous) retailers, with the proportion ranging from 0.1 to 3, while keeping the coefficient of variation the same as other retailers.

**Insert Figure 11 about here**

Figures 11 shows that the savings due to transshipment is the highest when this retailer is the same as the others in demand size (proportion=1). Apparently, this is due to the fact that the pooling effect due to transshipment is most beneficial when all retailers face the same distribution of demand. Furthermore, if the average demand of this different retailer is lower than the others, the percent savings is higher when this retailer is located towards the end (location 1) under the line configuration. On the other hand, if the average demand of this different retailer is higher than the others, the percent savings is higher when this retailer is located towards the center (location 5) under the line configuration. In all cases, the percent savings is always higher under the circle configuration than that of any of the three locations under the line configuration.

**Insert Figure 12 about here**

We next analyze the impact of demand variability of the different retailer on the system. In this case, we set the average demand of this different retailer to be the same as the other retailers, but vary the demand variability. Specifically, we use the range of the support of the uniform distribution, denoted by $R$, as our measure of demand variability, and vary this range $R$ from 30 to 250 while keeping the average demand at (50+200)/2=125. The results are summarized in Figure 12. Consistent with that observed in Figure 11, we observe that the savings due to transshipment is the highest when this retailer is the same as the others in demand variability ($R = 150$). Furthermore, our results suggest that for the line configuration, when the demand variability of this different retailer is smaller than the others, the percent savings is higher when this different retailer is located towards the end. On the other hand, when the demand variability of this different retailer is larger than the others, the percent savings is higher when this different retailer is located towards the center for the line configuration. The above observations contract what we observe in Figure 6 for the case of homogeneous retailers. Finally, consistent again with that observed in Figure 11, in all cases, the percent savings is always higher under the circle configuration than that of any of the three locations under the line configuration.

Our results suggest that transshipment provides the maximum benefit when the demand characteristics at all retailers are identical. Furthermore, for non-homogeneous retailers with different demand characteristics, transshipment provides more benefits for systems when retailers with lower average demand or lower variability are located towards the end locations and retailers with higher average demand or higher variability are located at the center locations. One plausible explanation of this observation is that a higher average demand or variability will result in a higher inventory level at these locations, and it is generally more beneficial to have the higher inventory retailers to locate toward the center of the network to reduce the potential transshipment costs to use these inventories to meet demands at other locations.

For all our computations, we use CPLEX 9.1 commercial solver to solve the second order cone optimization models on a 2.8GHz desktop with 512Mb memory. Computational time depends on the size of the network configuration, that is, the number of retailers and connecting arcs. A circle configuration with 10 retailers takes less than a second to solve, while with 40 retailers, it takes about four seconds. In Figure 13, we show the solution of a randomly generated fully connected network with Euclidian distances. It takes about three seconds to solve the 20-retailer transshipment problem with 190 arcs. (It is interesting to observe that only a fraction of the transshipment routes are active in this example.) Our computational results demonstrate that our robust optimization framework can be used to analyze large distribution networks efficiently.

**Insert Figure 13 about here**

# 7 Conclusions

In this paper, we study a distribution system consisting of multiple retail locations with transshipment operations among the retailers. We first develop a stochastic programming model for analyzing the optimal replenishment and transshipment policy. However, the stochastic programming model is complex to solve. Therefore, we propose a robust optimization framework for analyzing the impact of transshipment operations on such a distribution system. We show that the system under our robust optimization framework is analytically tractable and is computationally efficient for solve large distribution networks.

Therefore, our proposed robust optimization framework provides a powerful tool for analyzing possible transshipment designs and for evaluating whether to invest in transshipment operations when facing different types of network configurations and retailer/demand characteristics.

Using this robust optimization framework, we provide some numerical results to address a number of managerial issues regarding the impact of transshipment on reducing the costs of the distribution system under different system configurations and different retailer/demand characteristics. Our results show that the savings due to transshipment can be very substantial in a larger network of many retail locations, especially when the cost to transship a unit is significantly lower than that of holding the unit in the inventory. Our results also suggest that, for both line and circle configurations, transshipment provides the maximum benefit when the demand characteristics at all retailers are identical. On the other hand, transshipment provide the maximum benefit when the holding costs of the retailers are different, with few retailers with lower holding cost in more centralized locations. Finally, under the line configuration and non-homogeneous retailers with different demand characteristics, we show that transshipment provides more benefits for systems when retailers with lower average demand or lower variability are located towards the end locations and retailers with higher average demand or higher variability are located at the center locations. Finally, our results show that the circle configuration always yields higher savings from transshipment than the line configuration in both total cost and total ending inventory.

While our numerical study only deals with the line and circle configurations, it remains an interesting issue on how to combine different models to design a configuration that is most efficient in capturing the benefit of transshipment operations. We leave that for future research.

# References


[1] Alfredsson, P. and J. Verrijdt. 1999. Modeling emergency supply flexibility in a two-echelon inventory system. *Management Science* **45** 1416-1431.

[2] Archibald, A. W., S. A. Sassen and L. C. Thomas. 1997. An optimal policy for a two depot inventory system with stock transfer. *Management Science* **43** 173-183.

[3] Axsäter, S. 1990. Modelling emergency lateral transshipments in inventory systems. *Management Science* **36** 1329-1338.

[4] Axsäter, S. 2003. A new decision rule for lateral transshipments in inventory systems. *Management Science* **49** 1168-1179.

[5] Ben-Tal, A. and A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23** 769-805.

[6] Ben-Tal, A. and A. Nemirovski. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters* **25** 1-13.

[7] Ben-Tal, A. and A. Nemirovski. 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* **88** 411-424.

[8] Ben-Tal, A., A. Goryashko, E. Guslitzer and A. Nemirovski. 2004a. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99** 351-376.

[9] Ben-Tal, A., B. Golany, A. Nemirovski and J. Vial. 2004b. Supplier-retailer flexible commitments contracts: A robust optimization approach. *Manufacturing and Service Operations Management* **7** 248-273.

[10] Bertsimas, D., D. Pachamanova and M. Sim. 2004. Robust linear optimization under general norms. *Operations Research Letters* **32** 510-516.

[11] Bertsimas, D. and M. Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* **98** 49-71.

[12] Bertsimas, D. and M. Sim. 2004. Price of robustness. *Operations Research* **52** 35-53.

[13] Bertsimas, D. and M. Sim. 2005. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, forthcoming.

[14] Bertsimas, D. and A. Thiele. 2005. A robust optimization approach to supply chain management. *Operations Research*, forthcoming.

[15] Birge, J. R. and F. Louveaux. 1997. *Introduction to stochastic programming*. Springer, New York.

[16] Chen, X., M. Sim and P. Sun. 2005. A robust optimization perspective of stochastic programming. *Working Paper*, National University of Singapore.

[17] Cohen, M.A., P. R. Kleindorfer and H. L. Lee. 1986. Optimal stocking policies for low usage items in multi-echelon inventory systems. *Naval Research Logistics Quarterly* **33** 17-38.

[18] Dada, M. 1992. A two-echelon inventory system with priority shipments. *Management Science* **38** 1140-1153.

[19] Das, C. 1975. Supply and redistribution rules for two-location inventory systems: One-period analysis. *Management Science* **21** 765-776.

[20] Diks, E. B. and A. G. de Kok. 1996. Controlling a divergent two-echelon network with transshipments using the consistent approximate share rationing policy. *International Journal of Production Economics* **45** 369-379.

[21] Dong, L. and N. Rudi. 2004. Who benefits from transshipment? Exogenous vs endogenous wholesale prices. *Management Science* **50** 645-657.

[22] Dyer, M. and L. Stougie. 2005. Computational complexity of stochastic programming problems, *Mathematical Programming*, to appear.

[23] El-Ghaoui, L, and H. Lebret. 1997. Robust solutions to least-square problems to uncertain data matrices. *SIAM Journal of Matrix Analysis and Applications* **18** 1035-1064.

[24] El-Ghaoui, L., F. Oustry and H. Lebret. 1998. Robust solutions to uncertain semidefinite programs. *SIAM Journal of Optimization* **9** 33-52.

[25] Grahovac, J. and A. Chakravarty. 2001. Sharing and lateral transshipment of inventory in a supply chain with expensive, low-demand items. *Management Science* **47** 579-594.

[26] Herer, Y. T. and A. Rashit. 1999. Lateral stock transshipments in a two-location inventory system with fixed and joint replenishment costs. *Naval Research Logistics* **46** 525-548.

[27] Herer, Y. T. and M. Tzur. 2001. The dynamic transshipment problem. *Naval Research Logistics* **48** 386-408.

[28] Hoadley, B. and D. P. Heyman. 1977. A two-echelon inventory model with purchases, dispositions, shipments, returns, and transshipments. *Naval Research Logistics Quarterly* **24** 1-19.

[29] Jonsson, H. and E. A. Silver. 1987. Analysis of a two-echelon inventory control system with complete redistribution. *Management Science* **33** 215-227.

[30] Karmarker, U. S. and N. R. Patel. 1977. The one-period N-location distribution problem. *Naval Research Logistics Quarterly* **24** 559-575.

[31] Krishnan, K. S. and V. R. K. Rao. 1965. Inventory control in N warehouses. *Journal of Industrial Engineering* **16** 212-215.

[32] Lee, H. L. 1987. A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science* **33** 1302-1316.

[33] Minner, S., E. A. Silver and D. J. Robb. 2003. An improved heuristic for deciding on emergency transshipments. *European Journal of Operational Research* **148** 384-400.

[34] Robinson, L. W. 1990. Optimal and Approximate Policies in Multiperiod, Multilocation Inventory Models with Transshipments. *Operations Research* **38** 278-295.

[35] Rudi, N., S. Kapur and D. F. Pyke. 2001. A two-location inventory model with transshipment and local decision making. *Management Science* **47** 1668-1680.

[36] Scarf, H. 1958. A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*, Stanford University Press, California.

[37] Shapiro, A. and A. Nemirovski. 2004. On Complexity of stochastic programming problems. *Optimization Online*.

[38] Soyster, A. L. 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* **21** 1154-1157.

[39] Tagaras, G. 1989. Effects of pooling on the optimization and service levels of two-location inventory systems. *IIE Transactions* **21** 250-257.

[40] Tagaras, G. and M. A. Cohen. 1992. Pooling in two-location inventory systems with non-negligible replenishment lead times. *Management Science* **38** 1067-1083.

[41] Wee, K. E. and M. Dada. 2005. Optimal policies for transshiping inventory in a retail network. *Management Science* **51**, No.10, 1519-1533.

[42] Zhang, J. 2005. Transshipment and its impact on supply chain members' performance. *Management Science* **51**, No. 10, 1534-1539.

# A Appendix 1

**Proof of Proposition 1:** See Chen, et al. (2005).

**Proof of Theorem 1:** To establish the bound that $\mathrm{E}\left(|y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}|\right) \leq G(y^0, \boldsymbol{y})$, we show equivalently that for all $\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} \geq \boldsymbol{0}$, we have

$$\begin{aligned}
& (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(y^0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v}\|_2^2} \\
\geq\ & (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \mathrm{E}\Big(|y^0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}} + (\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v})'\tilde{\boldsymbol{z}}|\Big) && (19) \\
=\ & \mathrm{E}\Big((\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + |y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + (\boldsymbol{s}-\boldsymbol{u})'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + (\boldsymbol{t}-\boldsymbol{v})'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}})|\Big) \\
=\ & \mathrm{E}\Big((\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + (y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + (\boldsymbol{s}-\boldsymbol{u})'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + (\boldsymbol{t}-\boldsymbol{v})'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+ + \\
& \quad (-y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}} - (\boldsymbol{s}-\boldsymbol{u})'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) - (\boldsymbol{t}-\boldsymbol{v})'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+\Big) && (20) \\
\geq\ & \mathrm{E}\Big((\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + (y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} - \boldsymbol{u}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) - \boldsymbol{v}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+ + \\
& \quad (-y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}} - \boldsymbol{s}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) - \boldsymbol{t}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+\Big) && (21) \\
=\ & \mathrm{E}\Big((\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \\
& \quad y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} - \boldsymbol{u}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) - \boldsymbol{v}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}) + (-y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}} + \boldsymbol{u}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + \boldsymbol{v}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+ + \\
& \quad -y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}} - \boldsymbol{s}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) - \boldsymbol{t}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}) + (y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \boldsymbol{s}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + \boldsymbol{t}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+\Big) && (22) \\
=\ & \mathrm{E}\Big((-y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}} + \boldsymbol{u}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + \boldsymbol{v}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+ + (y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} + \boldsymbol{s}'(\bar{\boldsymbol{z}}-\tilde{\boldsymbol{z}}) + \boldsymbol{t}'(\underline{\boldsymbol{z}}+\tilde{\boldsymbol{z}}))^+\Big) \\
\geq\ & \mathrm{E}\Big((-y^0 - \boldsymbol{y}'\tilde{\boldsymbol{z}})^+ + (y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\Big) && (23) \\
=\ & \mathrm{E}\Big(|y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}|\Big). && (24)
\end{aligned}$$

Inequality (19) follows from the well-known result that

$$\mathrm{E}(|\tilde{v}|) \leq \sqrt{\mu^2 + \sigma^2} \tag{25}$$

for any given random variable $\tilde{v}$, with mean $\mu$ and standard deviation $\sigma$; for instance, see Scarf (1958). Equalities (20) and (24) follow from

$$|x| = (x)^+ + (-x)^+,$$

and inequality (22) follows from

$$x^+ = x + (-x)^+.$$

Inequalities (21) and (23) are due to the fact that $-\underline{\boldsymbol{z}} \leq \tilde{\boldsymbol{z}} \leq \bar{\boldsymbol{z}}$, and hence,

$$(\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{s} \geq 0$$
$$(\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{t} \geq 0$$
$$(\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{u} \geq 0$$
$$(\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{v} \geq 0$$

for all $\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} \geq \boldsymbol{0}$. ■

**Proof of Theorem 2:** The model of (12) restricts the recourse variables $w_{ij}$ to a linear decision rule as given in (9). Hence, the inventory position at retailer location $i$ also conforms to the linear decision rule

$$l_i(\tilde{\boldsymbol{z}}) = l_i^0 + \sum_{k=1}^{N} l_i^k \tilde{z}_k$$

in view of the linear constraints

$$\begin{aligned} l_i^0 &= x_i - d_i^0 - \sum_{j:(i,j)\in E} w_{ij}^0 + \sum_{j:(j,i)\in E} w_{ij}^0 \\ l_i^k &= -\Delta d_i^k - \sum_{j:(i,j)\in E} w_{ij}^k + \sum_{j:(j,i)\in E} w_{ij}^k. \end{aligned} \tag{26}$$

Then, it follows from Proposition 1 that the constraints

$$\begin{aligned} &l_i^0 \geq \sqrt{-2\ln(\epsilon/m)}p_i^0 + \boldsymbol{q}_i{}'\bar{\boldsymbol{z}} + \boldsymbol{r}_i{}'\underline{\boldsymbol{z}} \\ &p_i^k \geq \beta_k(l_i^k - q_i^k + r_i^k) \\ &p_i^k \geq -\alpha_k(l_i^k - q_i^k + r_i^k) \\ &\|\boldsymbol{p}_i\|_2 \leq p_i^0 \end{aligned}$$

imply that for all $i \in V$,

$$\mathrm{P}(l_i(\tilde{\boldsymbol{z}}) < 0) \leq 1 - \frac{\epsilon}{m}.$$

Therefore,

$$\mathrm{P}(l_i(\tilde{\boldsymbol{z}}) \geq 0) > 1 - \frac{\epsilon}{m}.$$

Using Bonferroni's inequality, we have

$$\mathrm{P}\left(l_i(\tilde{\boldsymbol{z}}) \geq 0,\ \forall i \in V\right) > 1 - \epsilon.$$

Thus, the solution $\boldsymbol{x}$ and linear transshipment policy that is feasible in the stochastic programming model (5).

It remains to show that $Z_r^*$ is an upper bound to the stochastic programming model (5). It follows directly from Theorem 1 and $c_{ij} \geq 0$ that

$$\mathrm{E}(c_{ij}|w_{ij}(\tilde{\boldsymbol{z}})|) \leq c_{ij} G(w_{ij}^0, \boldsymbol{w}_{ij}),$$

where

$$G(w_{ij}^0, \boldsymbol{w}_{ij}) = \min_{\boldsymbol{s}_{ij}, \boldsymbol{t}_{ij}, \boldsymbol{u}_{ij}, \boldsymbol{v}_{ij} \geq \boldsymbol{0}} \left\{ (\boldsymbol{s}_{ij} + \boldsymbol{u}_{ij})'\bar{\boldsymbol{z}} + (\boldsymbol{t}_{ij} + \boldsymbol{v}_{ij})'\underline{\boldsymbol{z}} + \right.$$
$$\left. \sqrt{(w_{ij}^0 + (\boldsymbol{s}_{ij} - \boldsymbol{u}_{ij})'\bar{\boldsymbol{z}} + (\boldsymbol{t}_{ij} - \boldsymbol{v}_{ij})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{w}_{ij} - \boldsymbol{s}_{ij} + \boldsymbol{t}_{ij} + \boldsymbol{u}_{ij} - \boldsymbol{v}_{ij}\|_2^2} \right\}.$$

On the other hand, the last constraint in (12) requires that $g_{ij} \geq G(w_{ij}^0, \boldsymbol{w}_{ij})$ for any feasible $g_{ij}$. Since $c_{ij} \geq 0$, we must have $c_{ij}G(w_{ij}^0, \boldsymbol{w}_{ij}) = c_{ij}g_{ij}^*$, where $g_{ij}^*$ is the optimal solution of $g_{ij}$ to (12). Finally we note that the inventory position follows decision rule, hence, by using the same argument, we have

$$\mathrm{E}(h_i l_i(\tilde{\boldsymbol{z}})^+) = \frac{h_i}{2}(l_i^0 + \mathrm{E}(|l_i(\tilde{\boldsymbol{z}})|) \leq \frac{h_i}{2}(l_i^0 + G(l_i^0, \boldsymbol{l}_i)) = h_i f_i^*,$$

where $f_i^*$ is the optimal solution of $f_i$ to (12). Comparing the objective functions in (5) and (12), the result follows easily. ■

**Proof of Theorem 3** (a) Observe that when $\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} = \boldsymbol{0}$, we have

$$(\boldsymbol{s} + \boldsymbol{u})'\bar{z} + (\boldsymbol{t} + \boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(y_0 + (\boldsymbol{s} - \boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t} - \boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\boldsymbol{y} - \boldsymbol{s} + \boldsymbol{t} + \boldsymbol{u} - \boldsymbol{v}\|_2^2} = \sqrt{y_0^2 + \|\boldsymbol{y}\|_2^2}.$$

Therefore, the result follows easily.

**(b)** Suppose that

$$y_0 + \boldsymbol{y}'\boldsymbol{z} \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}.$$

Let $\boldsymbol{s} = \boldsymbol{t} = \boldsymbol{0}$, $u_k = (-y_k)^+$, $v_k = (y_k)^+$ for $k = 1, \ldots, N$ and

$$z_k^* = \begin{cases} \bar{z}_k & \text{if } y_k < 0 \\ -\underline{z}_k & \text{otherwise} \end{cases}.$$

Since $\boldsymbol{z}^* \in \mathcal{W}$, we have $y_0 + \boldsymbol{y}'\boldsymbol{z}^* \geq 0$. Furthermore, it is easy to verify that

$$\boldsymbol{y} = \boldsymbol{v} - \boldsymbol{u}$$

and

$$y_0 - \boldsymbol{u}'\bar{\boldsymbol{z}} - \boldsymbol{v}'\underline{\boldsymbol{z}} = y_0 + \boldsymbol{y}'\boldsymbol{z}^* \geq 0.$$

We have

$$\underbrace{(\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}}+(\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}}}_{=\boldsymbol{u}'\bar{\boldsymbol{z}}+\boldsymbol{v}'\underline{\boldsymbol{z}}}+\sqrt{(\underbrace{y_0+(\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}}+(\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}}}_{=y_0-\boldsymbol{u}'\bar{\boldsymbol{z}}-\boldsymbol{v}'\underline{\boldsymbol{z}}\geq 0})^2+\underbrace{\|\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v}\|_2^2}_{=0}}=y_0.$$

Hence,

$$y_0=\mathrm{E}\left(|y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}}|\right)\leq G(y_0,\boldsymbol{y})\leq y_0.$$

Similarly, if

$$y_0+\boldsymbol{y}'\boldsymbol{z}\leq 0 \qquad \forall \boldsymbol{z}\in\mathcal{W}$$

then let $\boldsymbol{v}=\boldsymbol{u}=\boldsymbol{0}$, $s_k=(y_k)^+$, $t_k=(-y_k)^+$ for $k=1,\ldots,N$ and

$$z_k^*=\begin{cases}\bar{z}_k & \text{if } y_k>0\\ -\underline{z}_k & \text{otherwise}\end{cases}.$$

Since $\boldsymbol{z}^*\in\mathcal{W}$, we have $y_0+\boldsymbol{y}'\boldsymbol{z}^*\leq 0$. Furthermore, it is easy to verify that

$$\boldsymbol{y}=\boldsymbol{s}-\boldsymbol{t}$$

and

$$y_0+\boldsymbol{s}'\bar{\boldsymbol{z}}+\boldsymbol{t}'\underline{\boldsymbol{z}}=y_0+\boldsymbol{y}'\boldsymbol{z}^*\leq 0.$$

Hence, we have

$$\underbrace{(\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}}+(\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}}}_{=\boldsymbol{s}'\bar{\boldsymbol{z}}+\boldsymbol{t}'\underline{\boldsymbol{z}}}+\underbrace{\sqrt{(\underbrace{y_0+(\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}}+(\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}}}_{=y_0+\boldsymbol{s}'\bar{\boldsymbol{z}}+\boldsymbol{t}'\underline{\boldsymbol{z}}\leq 0})^2+\underbrace{\|\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v}\|_2^2}_{=0}}}_{-y_0-\boldsymbol{s}'\bar{\boldsymbol{z}}-\boldsymbol{t}'\underline{\boldsymbol{z}}}=-y_0=|y_0|.$$

Therefore,

$$|y_0|=\mathrm{E}\left(|y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}}|\right)\leq G(y_0,\boldsymbol{y})\leq |y_0|.$$

■

# B Appendix 2: Multiple-Period Model

We assume that the uncertain demand at each retail location $\tilde{d}_i^t$ is linearly dependent on a set of independent random variables

$$\tilde{\boldsymbol{z}}^t:=(\tilde{z}_1^1,\ldots,\tilde{z}_{N_1}^1,\ldots,\tilde{z}_1^t,\ldots,\tilde{z}_{N_t}^t)$$

as follows:

$$d_i^t(\tilde{\boldsymbol{z}}^t) = d_i^{t0} + \sum_{\tau=1}^{t}\sum_{k=1}^{N_\tau} \Delta d_i^{t\tau k}\tilde{z}_k^\tau. \tag{27}$$

Each random variable $\tilde{z}_k^\tau$ has mean zero, unit standard deviations, and support in $[-\underline{z}_k^\tau, \bar{z}_k^\tau]$ with $\underline{z}_k^\tau, \bar{z}_k^\tau > 0$. Define the forward deviation $\alpha_k^\tau$ and backward deviation $\beta_k^\tau$ as given in (7) and (8). At time period $t$ progresses, a new set of $N_t$ primitive uncertainties $(\tilde{z}_1^t, \ldots, \tilde{z}_{N_t}^t)$ unfolds. Define

$$\begin{aligned}
\bar{\boldsymbol{z}}^t &:= (\bar{z}_1^1, \ldots, \bar{z}_{N_1}^1, \ldots, \bar{z}_1^t, \ldots, \bar{z}_{N_t}^t) \\
\underline{\boldsymbol{z}}^t &:= (\underline{z}_1^1, \ldots, \underline{z}_{N_1}^1, \ldots, \underline{z}_1^t, \ldots, \underline{z}_{N_t}^t).
\end{aligned}$$

We shall assume the linear decision rules on the recourse variables, order quantity and the transshipment flows, i.e.,

$$\begin{aligned}
x_i^t(\tilde{\boldsymbol{z}}^{t-1}) &= \begin{cases} x_i^{t0} + \displaystyle\sum_{\tau=1}^{t-1}\sum_{k=1}^{N_\tau} x_i^{t\tau k}\tilde{z}_k^\tau & \text{if } t = 2, \ldots, T \\ x_i^{10} & \text{if } t = 1 \end{cases} && \forall i \in V, t = 1, \ldots, T \\
w_{ij}^t(\tilde{\boldsymbol{z}}^t) &= w_{ij}^{t0} + \sum_{\tau=1}^{t}\sum_{k=1}^{N_\tau} w_{ij}^{t\tau k}\tilde{z}_k^\tau && \forall (i,j) \in E, t = 1, \ldots, T
\end{aligned} \tag{28}$$

For notational convenience, define

$$\hat{N}_t = \sum_{\tau=1}^{t} N_\tau$$

and denote the vectors

$$\begin{aligned}
\boldsymbol{x}_i^t &= (x_i^{t11}, \ldots, x_i^{t1N_1}, \ldots, x_i^{t(t-1)1}, \ldots, l_i^{t(t-1)N_{t-1}}) \\
\boldsymbol{w}_i^t &= (w_{ij}^{t11}, \ldots, w_{ij}^{t1N_1}, \ldots, w_{ij}^{tt1}, \ldots, w_{ij}^{ttN_t})
\end{aligned}$$

such that

$$\begin{aligned}
x_i^t(\tilde{\boldsymbol{z}}) &= x_i^{t0} + \boldsymbol{x}_i^{t\prime}\tilde{\boldsymbol{z}}^{t-1} \quad \forall t = 2, \ldots, T \\
w_{ij}^t(\tilde{\boldsymbol{z}}) &= w_{ij}^{t0} + {\boldsymbol{w}_{ij}^t}'\tilde{\boldsymbol{z}}^t \quad \forall t = 1, \ldots, T.
\end{aligned}$$

The multiple-period robust optimization model is given as follows:

$$
\begin{aligned}
Z_r^* = \quad \min \quad & \sum_{t=1}^{T} \sum_{i \in V} \left( a_i^t x_i^{t0} + h_i f_i^t + \sum_{j:(i,j) \in E} c_{ij}^t g_{ij}^t \right) \\
\text{s.t.} \quad & l_i^{t0} = l_i^{(t-1)0} + x_i^{t\tau 0} - d_i^0 - \sum_{j:(i,j)\in E} w_{ij}^{t0} + \sum_{j:(j,i)\in E} w_{ji}^{t0} && \forall i \in V \\
& && t = 1, \ldots, T \\
& l_i^{t\tau k} = \begin{cases} l_i^{(t-1)\tau k} - \Delta d_i^{t\tau k} - \sum_{j:(i,j)\in E} w_{ij}^{t\tau k} + \sum_{j:(j,i)\in E} w_{ji}^{t\tau k} & \text{if } \tau < t \\ -\Delta d_i^{ttk} - \sum_{j:(i,j)\in E} w_{ij}^{ttk} + \sum_{j:(j,i)\in E} w_{ji}^{ttk} & \text{if } \tau = t \end{cases} && \forall t = 1, \ldots, T, \\
& && \tau \le t \\
& && k \in \{1, \ldots, N_\tau\} \\
& && i \in V \\
& l_i^{t0} \ge \sqrt{-2\ln(\epsilon_t/m)}\, p_i^{t0} + \boldsymbol{q}_i^{t\prime} \underline{\boldsymbol{z}}^t + \boldsymbol{r}_i{}' \bar{\boldsymbol{z}}^t && \forall i \in V \\
& && t = 1, \ldots, T \\
& p_i^{t\tau k} \ge \beta_k^\tau (l_i^{t\tau k} - q_i^{t\tau k} + r_i^{t\tau k}) && \forall t = 1, \ldots, T, \\
& && \tau \le t, \\
& && k \in \{1, \ldots, N_\tau\}, \\
& && i \in V \\
& p_i^{t\tau k} \ge -\alpha_k^\tau (l_i^{t\tau k} - q_i^{t\tau k} + r_i^{t\tau k}) && \forall t = 1, \ldots, T, \\
& && \tau \le t, \\
& && k \in \{1, \ldots, N_\tau\}, \\
& && i \in V \\
& \|\boldsymbol{p}_i^t\|_2 \le p_i^{t0} && \forall i \in V, t = 1, \ldots, T \\
& l_i^{t0} \in \Re, \boldsymbol{l}_i^t \in \Re^{\hat{N}_t} && \forall i \in V, t = 1, \ldots, T \\
& p_i^{t0} \in \Re_+, \boldsymbol{p}_i^t \in \Re^{\hat{N}_t} && \forall i \in V, t = 1, \ldots, T \\
& \boldsymbol{q}_i^t, \boldsymbol{r}_i^t \in \Re_+^{\hat{N}_t} && \forall i \in V, t = 1, \ldots, T \\
& f_i^t \ge \frac{1}{2} \Big\{ l_i^{t0} + (\hat{\boldsymbol{s}}_i^t + \hat{\boldsymbol{u}}_i^t)' \bar{\boldsymbol{z}}^t + (\hat{\boldsymbol{t}}_i^t + \hat{\boldsymbol{v}}_i^t)' \underline{\boldsymbol{z}}^t + \\
& \quad \sqrt{(l_i^{t0} + (\hat{\boldsymbol{s}}_i^t - \hat{\boldsymbol{u}}_i^t)' \bar{\boldsymbol{z}}^t + (\hat{\boldsymbol{t}}_i^t - \hat{\boldsymbol{v}}_i^t)' \underline{\boldsymbol{z}}^t)^2 + \|\boldsymbol{l}_i^t - \hat{\boldsymbol{s}}_i^t + \hat{\boldsymbol{t}}_i^t + \hat{\boldsymbol{u}}_i^t - \hat{\boldsymbol{v}}_i^t\|_2^2} \Big\} && \forall i \in V \\
& && t = 1, \ldots, T \\
& f_i^t \in \Re, \quad \hat{\boldsymbol{s}}_i^t, \hat{\boldsymbol{t}}_i^t, \hat{\boldsymbol{u}}_i^t, \hat{\boldsymbol{v}}_i^t \in \Re_+^{\hat{N}_t}, && \forall i \in V \\
& && t = 1, \ldots, T \\
& \ldots
\end{aligned} \tag{29}
$$

$$\ldots$$

$$g^t_{ij} \geq (\boldsymbol{s}^t_{ij} + \boldsymbol{u}^t_{ij})'\bar{\boldsymbol{z}}^t + (\boldsymbol{t}^t_{ij} + \boldsymbol{v}^t_{ij})'\underline{\boldsymbol{z}}^t + \sqrt{(w^{t0}_{ij} + (\boldsymbol{s}^t_{ij} - \boldsymbol{u}^t_{ij})'\bar{\boldsymbol{z}}^t + (\boldsymbol{t}^t_{ij} - \boldsymbol{v}^t_{ij})'\underline{\boldsymbol{z}}^t)^2 + \|\boldsymbol{w}^t_{ij} - \boldsymbol{s}^t_{ij} + \boldsymbol{t}^t_{ij} + \boldsymbol{u}^t_{ij} - \boldsymbol{v}^t_{ij}\|^2_2}\Big) \qquad \forall (i,j) \in E,\ t = 1, \ldots, T$$

$$g^t_{ij} \in \Re, \quad \boldsymbol{s}^t_{ij}, \boldsymbol{t}^t_{ij}, \boldsymbol{u}^t_{ij}, \boldsymbol{v}^t_{ij} \in \Re^{\hat{N}_t}_+, \quad \boldsymbol{w}^t_{ij} \in \Re^{\hat{N}_t} \qquad \forall (i,j) \in E,\ t = 1, \ldots, T$$

$$x^{10}_i \geq 0 \qquad \forall i \in V$$

$$x^{t0}_i \geq \check{\boldsymbol{q}}^{t-1'}_i \underline{\boldsymbol{z}}^{t-1} + \check{\boldsymbol{r}}^{t-1'}_i \bar{\boldsymbol{z}}^{t-1} \qquad \forall i \in V,\ t = 2, \ldots, T$$

$$\boldsymbol{x}^t_i = \check{\boldsymbol{q}}^{t-1}_i - \check{\boldsymbol{r}}^{t-1}_i \qquad \forall i \in V,\ t = 2, \ldots, T$$

$$\check{\boldsymbol{q}}^t_i, \check{\boldsymbol{r}}^t_i \in \Re^{\hat{N}_t}_+ \qquad \forall i \in V,\ t = 1, \ldots, T-1$$

We can extend the results of Theorem 2 to the multiple-period model (29).

**Theorem 4** *Under Assumptions (27) and (28), the optimal solution $\boldsymbol{x}$ of the robust optimization model (29) is a feasible solution to the stochastic optimization model (5). Moreover, the optimal objective value of (29), $Z^*_r$, is an upper bound for the optimal objective value of (5), $Z^*$.*

**Proof :** The model of (29) requires the recourse varaibles $w^t_{ij}$ and $x^t_i$ to follow the linear decision rules as given in (28). In view of the linear constraints in (29), the inventory position at retailer location $i$ at time $t$ also conforms to the linear decision rule, i.e.,

$$l^{t0}_i = l^{(t-1)0}_i + x^{t\tau 0}_i - d^0_i - \sum_{j:(i,j)\in E} w^{t0}_{ij} + \sum_{j:(j,i)\in E} w^{t0}_{ji} \qquad \forall i \in V,\ t = 1, \ldots, T$$

$$l^{t\tau k}_i = \begin{cases} l^{(t-1)\tau k}_i - \Delta d^{t\tau k}_i - \sum_{j:(i,j)\in E} w^{t\tau k}_{ij} + \sum_{j:(j,i)\in E} w^{t\tau k}_{ji} & \text{if } \tau < t \\ -\Delta d^{ttk}_i - \sum_{j:(i,j)\in E} w^{ttk}_{ij} + \sum_{j:(j,i)\in E} w^{ttk}_{ji} & \text{if } \tau = t \end{cases} \qquad \forall t = 1, \ldots, T,\ \tau \leq t,\ k \in \{1, \ldots, N_\tau\},\ i \in V. \tag{30}$$

Observe that the linear decision rule on the order quantity $x^t_i(\tilde{\boldsymbol{z}}^{t-1})$ in (28) implies that $\mathrm{E}(a^t_i x^t_i(\tilde{\boldsymbol{z}}^{t-1})) = a^t_i x^{t0}_i$. The rest of the proof follows in the same manner as in the proof of Theorem 2, and we omit the details here. ■

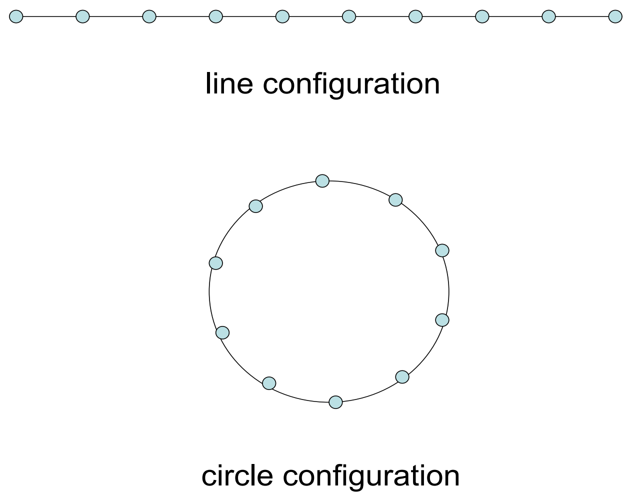


Figure 1: Line Configuration vs. Circle Configuration

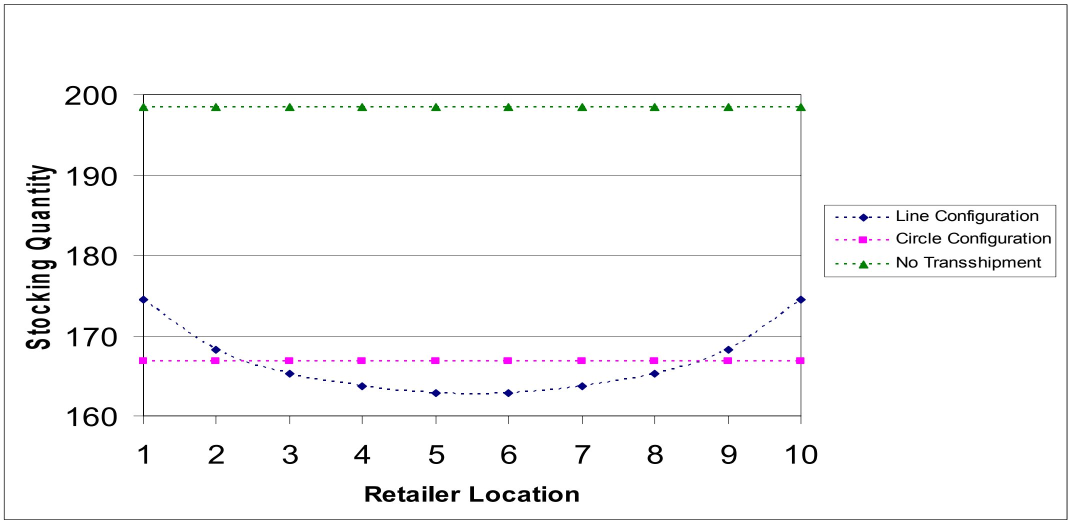


Figure 2: Stocking Quantity for Each Retailer

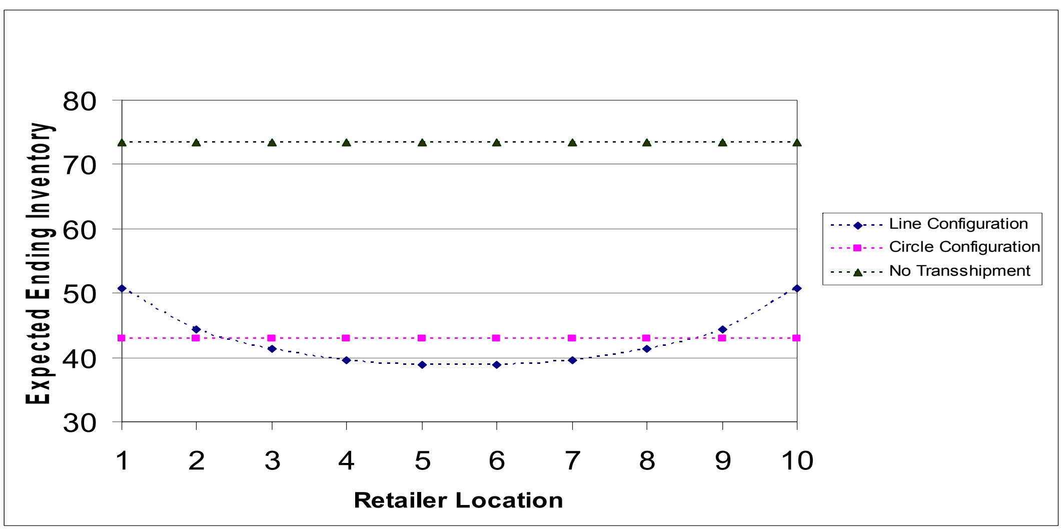


Figure 3: Expected Ending Inventory for Each Retailer

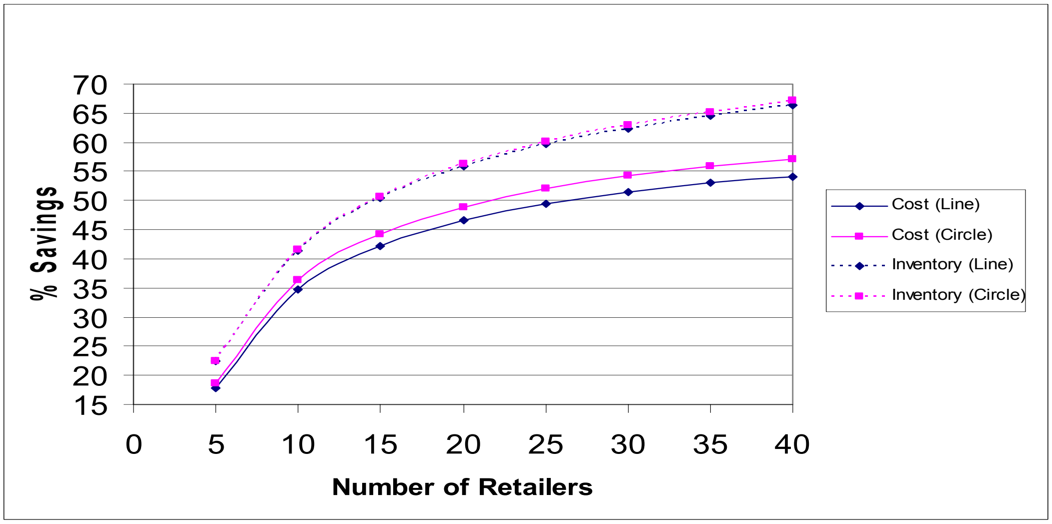


Figure 4: Impact of Number of Retailers

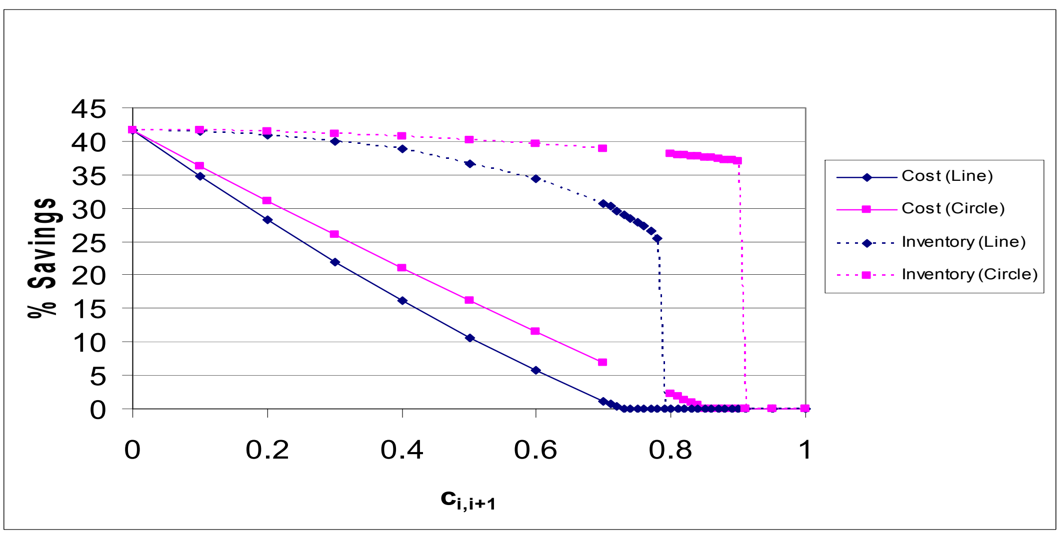


Figure 5: Impact of Unit Transshipment Cost

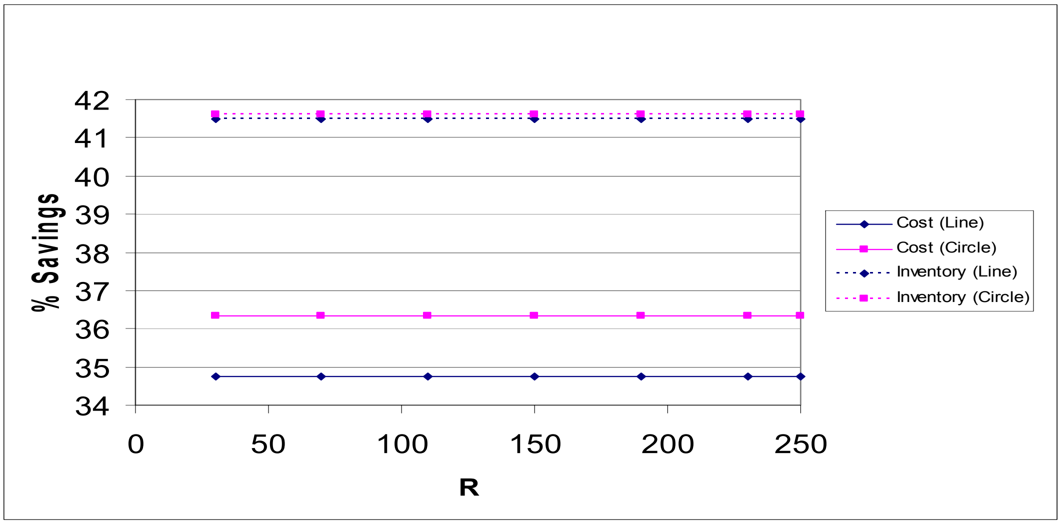


Figure 6: Impact of Demand Variability

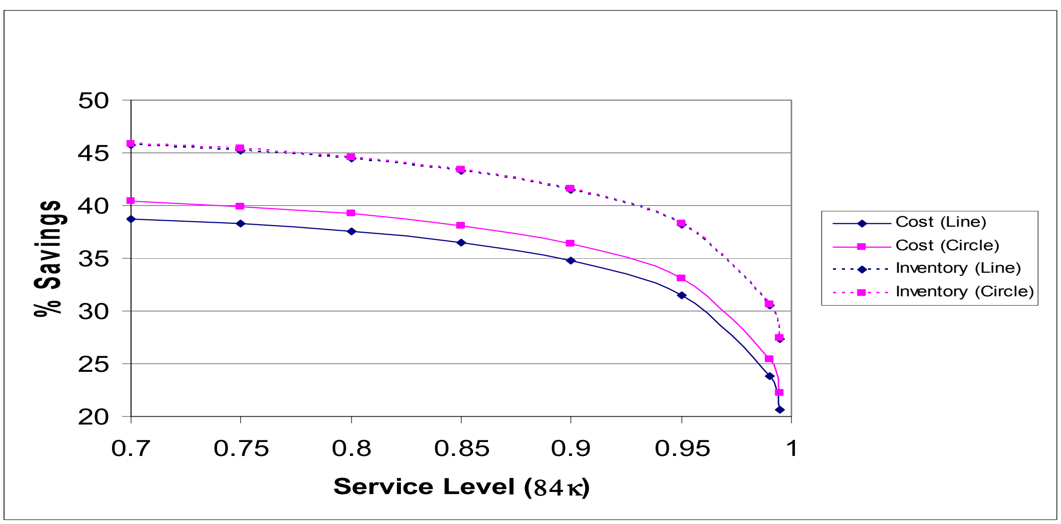


Figure 7: Impact of Service Level

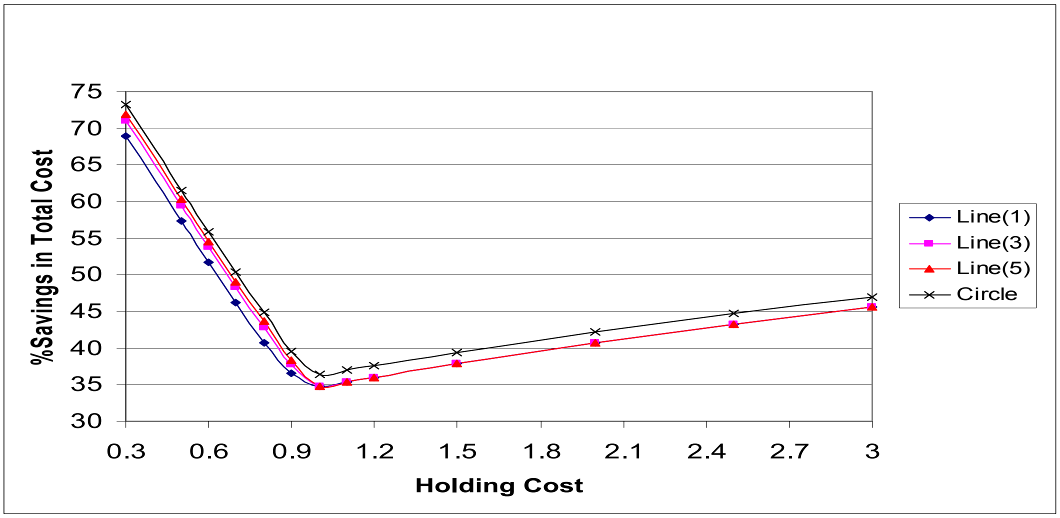


Figure 8: Non-homogeneous Retailer - Holding Cost

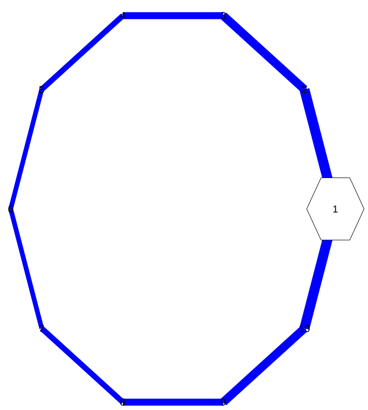


Figure 9: Circle network with $h_1 = 0.9$

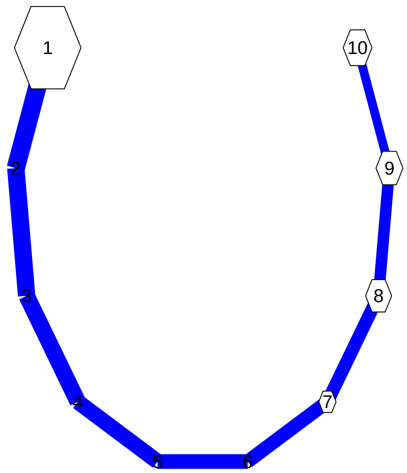


Figure 10: Line network with $h_1 = 0.9$

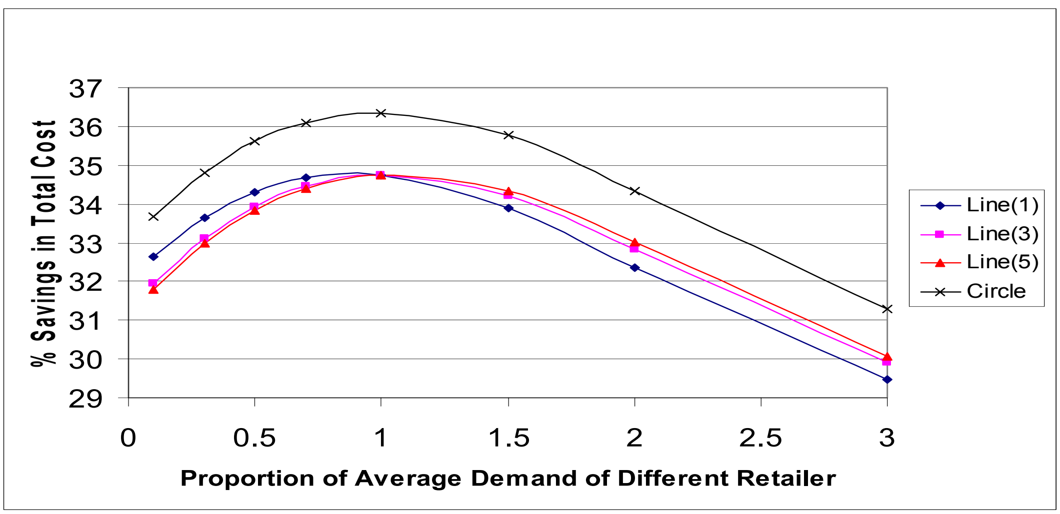


Figure 11: Non-homogeneous Retailers - Average Demand

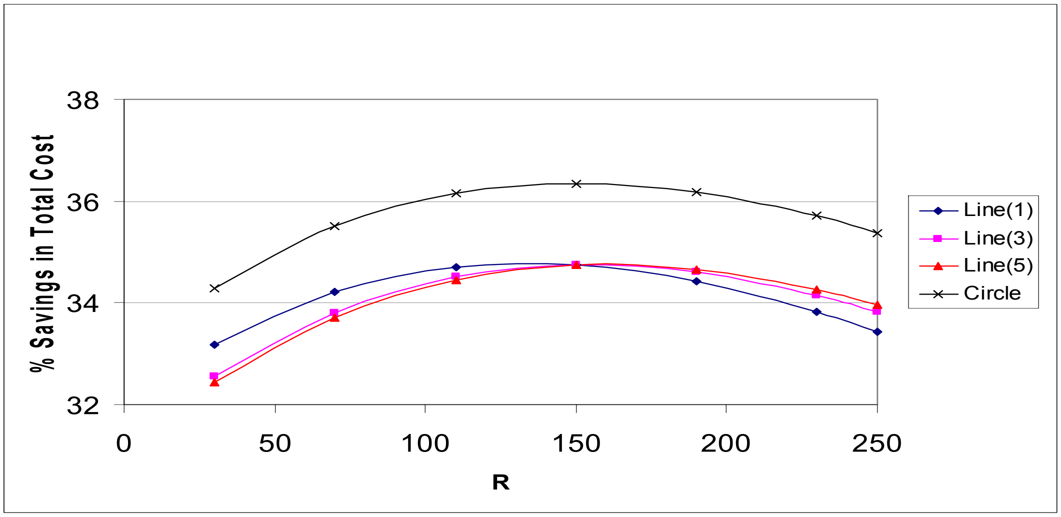


Figure 12: Non-homogeneous Retailer - Demand Variability

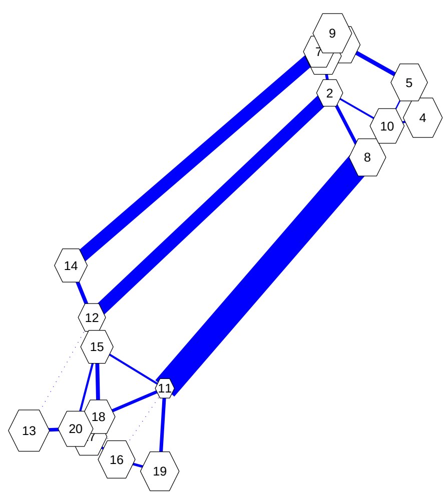


Figure 13: Fully connected random network