Impact of Bimodal and Lognormal Distributions in Ocean Transportation Transit Time on Logistics Costs

Lita Das, Basak Kalkanci, and Chris Caplice

This paper shows that ignoring bimodality and lognormality in transit time distributions can cause large increases in the logistic costs of maritime transportation. Bimodal and lognormal transit time distributions are observed to be of moderate frequency (approximately 17%) but high impact in the volume of shipment carried (approximately 85%) in these lanes. Ignoring and assuming incorrect distribution of transit time can have dramatic implications on the safety stock levels and reorder points and hence inventory cost incurred by the shipper. To display the incorrectness of such assumptions, the paper compares the typical approach of using a deterministic value (Case 1) for transit time and Hadley–Whitin (1963) with normal approximation (Case 2) to the authors' simulation and empirical analysis on bimodal and lognormal transit time distributions (Case 3). This paper further explores how the shipper should optimally manage inventory in the parameters of transit time distribution and critical ratio (or the service level of the shipper). Specifically, different regions are defined by the transit time distribution parameters and critical ratios that determine the magnitude of relative cost differences when the three cases are compared.

Maritime transportation traffic has seen tremendous growth in recent years because of the surge in globalization and international trade. Container shipments between North America, Asia, and Europe have nearly tripled in the past 16 years, rising from just over 15 million 20-ft equivalent units in 1995 to over 45 million in 2011 (*1*).

Higher trade volumes coupled and increased shipping locations have made conducting and coordinating shipping operations a complex task. Transporting a shipment from the point of origin (typically the manufacturer) to the final destination (usually a distribution center) involves a series of individual and often independently managed activities. The transportation transit time is the cumulative effect of these individual movements. Companies shipping goods across the globe are concerned with the average total end-to-end transit time as well as the unreliability in the shipments, variability, and shape of the distribution itself.

L. Das, 77 Massachusetts Avenue, and C. Caplice, 75 Massachusetts Avenue, Building E40-275, Engineering Systems Division, Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, MA 02139. B. Kalkanci, Scheller College of Business, Georgia Institute of Technology, 800 West Peachtree Street, NW, Atlanta, GA 30308. Corresponding author: L. Das, litadas@mit.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2409, Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 63–73.
DOI: 10.3141/2409-09

In recent years, a number of issues have affected global ocean transit times, including port congestion, bad weather, slow steaming, and sailing longer distances to avoid pirates. Schedule unreliability was identified as a major problem for manufacturers in surveys carried out in 2011 by the U.S. Federal Maritime Commission and logistics firm BDP International. As discussed in a recent news article, only 63.7% of containers were on time in the first 20 weeks of 2012 versus 65.9% a year earlier, according to INTTRA, a U.S. e-commerce platform that handles 525,000 shipments a week (*2*). Higher fuel costs, increased competition, and lower revenues are affecting the service quality of container shipping companies and leading some to even shed service on certain trade lanes.

Variability and unreliability in ocean transport ultimately affect the shippers' operational performance. In theory, the uncertainty in transit time can be modeled as a probability distribution and considered explicitly in the calculation of an optimal inventory policy for a shipper. However, a review of common practices in industry and conversations with companies and academic experts suggest that the companies do not necessarily consider transport time variability or unreliability in their inventory planning. In fact, most planning systems such as SAP and Oracle are configured to consider only deterministic transit times. Other systems, such as SAP Advanced Planning and Optimization, use a simplified approach to address variability: standard deviation of the transit time is calculated and used in the classical Hadley–Whitin formula (*3*). However, calculations done using this formula usually assume that the transit times are normally distributed.

The authors analyzed transit time information of more than 125,000 container movements from four major U.S. shippers; the movements occurred in 2011 and 2012. All the container movement information analyzed and discussed in this paper uses line (and not charter) shipping. The analysis has shown that ocean transport times are rarely normal and are often either lognormal or bimodal. Moreover, interactions with participants (including around 40 ocean shippers, carriers, and third-party logistics providers) at a global ocean transportation roundtable (*4*) in November 2012 further strengthened the belief of the existence of lognormal and bimodal distributions in transit time. However, neither of these transit time distributions has been well examined. Lognormal distributions occur because of few and long-delayed shipments. Bimodal transit times can occur for several reasons. For example, ocean carriers tend to follow weekly schedules in which shipments from a particular origin are picked up only once a week. If a container arrives late or is bumped, it will have to wait a week for the next vessel. A bimodal distribution will also arise when a shipper has multiple carriers, each with a different transit time, serving the same trade lane. From the shipper's perspective, the resultant transit time on the trade lane being served by more than carrier is the combination of faster and slower carriers. A bimodal distribution can also be caused when carriers have the flexibility to adjust speeds (e.g., slow steaming). Finally, because shippers can handle both high- and low-value products, they choose to use different service levels of shipping that have different transit times on the same trade lane. For instance, the company could decide to prioritize the high-value products and bump the low-value ones.

Following the observation that transit time distributions are often lognormal or bimodal, the aim is to characterize the impact of the variability from such distributions on shippers under (*a*) common inventory planning practices and (*b*) ocean transportation transit time distributions observed in practice. Three case studies were compared in the analysis. The deterministic case (Case 1) corresponds to the scenario of completely ignoring variability and using only the average transit time when setting inventory levels. This method is the most commonly used in practice. The normal case (Case 2) corresponds to the scenario in which the inventory policy decisions are calculated with the Hadley–Whitin formula (*3*) (with the inherent assumption that transit times are normally distributed). The bimodal (or lognormal) case (Case 3) denotes optimizing inventory decisions using the actual transit time distribution. Figure 1 shows the three cases for a given mean and standard deviation of the transit time distribution.

Using transactional data from several shippers (importers and exporters) on their ocean shipments, the authors analyzed and characterized the transit time distributions. Unlike data available from well-recognized sources such as Drewry Shipping Consultants Ltd. and Lloyds Maritime Intelligence, which focus only on the port-to-port component of a global shipment, the authors analyzed the end-to-end transit times of ocean transportation in global supply chains as observed by the shippers. The authors then examined the performance of common inventory practices to account for variability under the two dominant transit time distributions observed in the data. This analysis will help provide insights on the conditions (in the characteristics of the lane and service level) for when shippers should (and should not) consider transit time variability in setting their inventory levels. Finally, the authors aim to aid in deciding an effective inventory policy for shippers who observe bimodally or lognormally distributed transit times and a specific service level. In this paper, the authors show the detailed analysis of bimodal distribution of transit time and discuss the results for the lognormal case only. It was observed that the results for the two distributions were consistent.

The rest of the paper is structured as follows. A literature review outlines the relevant research on the effect of uncertainty in transit time and how to design inventory policies under such circumstances. Next, the analysis is discussed in two parts. The first part refers to the statistical test performed to identify nonunimodality in total transit time, and the second part derives the cost structure used to compare the three cases. The subsequent section discusses the results obtained from the simulation on the three cases for both theoretical and empirical distributions. Finally, conclusions are presented.

LITERATURE REVIEW

Numerous papers have addressed uncertainty in transit time and its impact on inventory management. The effect of shape of transit time distribution on inventory management decisions has been studied in terms of coefficient of variation (CV), skewness, and kurtosis (*5–7*).

Literature suggests variable levels of effects of assumption of normality in transit time demand distribution. For instance, whereas Eppen and Martin (*8*) discuss that using a normal approximation can lead to erroneous inventory management decisions (reorder points and safety stock levels), Tyworth and O'Neill (*9*) state that normal approximation is actually robust (for CV of transit time demand of less than 0.45). However, Lau and Lau (*10*) show that normal approximations are not robust, even for distributions with low CVs.

Bagchi et al. (*11*) also show the impact of ignoring variability in transit time on safety stock levels and hence logistics cost. The paper asserts that normal approximation for demand over transit time when it is otherwise can lead to significant errors. The authors state that one of the reasons for the use of normal distribution is familiarity and extensive tabulation of it.

In specific relation to this paper, Tyworth and O'Neill (*9*) and Bookbinder and Lordahl (*12*) discuss bimodally distributed demand over transit time. The former use a symmetrical bimodal distribution (50% of products on time and the rest delayed). However, it is different from the observations of transit time data of the authors of the current paper; these observations suggest different bimodal distribution characteristics. The assumption is that such a bimodal distribution does not encompass all possible levels of bimodality as seen in the data

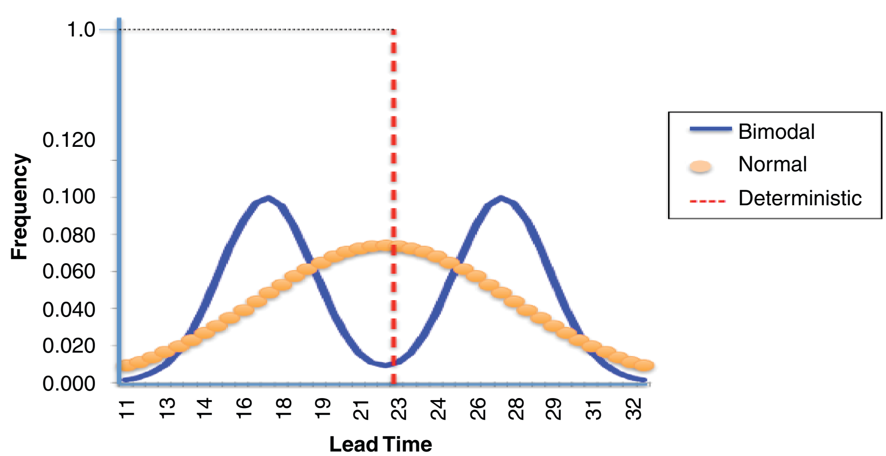


FIGURE 1 Shape approximations of transit time distribution in three cases.

available to the authors of the current paper. However, Bookbinder and Lordahl (*12*) discuss a different approach, a distribution-free approach, to tackling the problem of transit time demand distributions that are not normal. They suggest that a bootstrap method is more accurate than normal approximation for modeling demand over transit time that is not normal.

Summarily, it was ascertained that the distribution of demand over transit time can have large effects on inventory management decisions on reorder quantities and safety stock levels desired, and hence the total logistics cost. It is also seen that distribution of demand over transit time is preferably approximated to a normal distribution because of its familiarity compared with other distributions. However, the review suggests that the effect of incorrectly assuming normality in demand over transit time distribution is not consistent. Although on one hand it was shown that normal approximations are robust for low CVs, on the other hand it was disproven even at much lower CV values.

In addition, bimodality and lognormality have been observed in transit time distributions in the maritime transportation data available to the authors this paper. However, such distributions have not been concentrated on. This paper aims to study the combined effect of bimodality in transit time distributions and a range of service levels set by the shipper. Moreover, the authors shed light on another aspect of variability in transit time pertaining to industry—the issue of using a deterministic value for transit time as a common practice for making inventory management decisions. The authors observe different regions of comparison on the basis of service levels (critical ratios) and the extent to which one is worse off by choosing a deterministic value for transit time or a normal approximation or the actual bimodal and lognormal distribution.

ANALYSIS

For a significant portion of the shipments, the null hypothesis that the transit times are unimodally distributed is rejected using Hartigan's dip test of unimodality (*13*). This test calculates a dip statistic, which is the maximum difference between the empirical distribution and the unimodal distribution that minimizes the maximum difference.

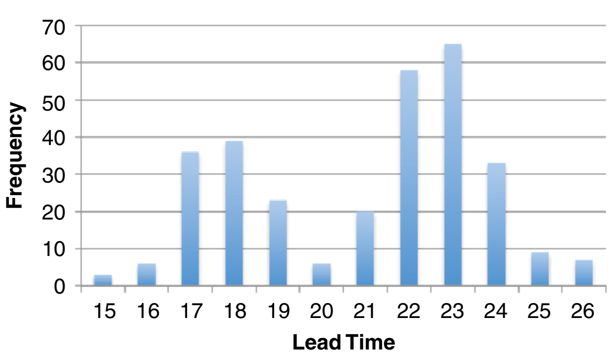
Inventory calculations involve simulating the demand over transit time that is observed in reality and then comparing the inventory levels calculated separately for each case to obtain the costs that the firm incurs in stocking excess or not enough of the product. The analysis of inventory cost is based on a given critical ratio (defined as the ratio of shortage to surplus cost). The critical ratio discussed is considered to be equivalent to the service level targeted by the firm.

The authors evaluate the impact of variability of transit time on logistics cost. The analysis for calculation of safety stock and order-up-to levels is performed using a simulation model. It is assumed that transit time and demand are independent. For the purposes of the simulation, it is assumed that a normally distributed demand is observed. The resulting value of safety stock is the average of 10,000 runs of Monte Carlo simulation. The authors analyze the performance of different cases by varying (*a*) critical ratio and (*b*) level of bimodality for bimodal transit time distribution (captured as the normalized difference between the two means of transit time distribution under a fixed mixture rate and standard deviation); the bimodal transit time distribution could also translate to CV of the distribution. Similarly, CV is varied for the case of lognormal distribution.

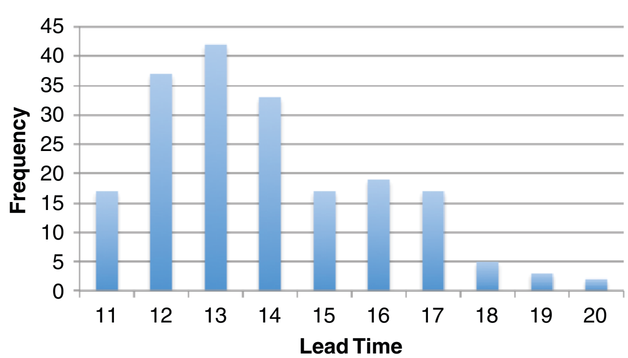
A stochastic inventory model and simulations for the calculation of the cost and safety stock level capture the impact of variability in transit times. For nonunimodal lanes, the authors model the transit time as a bimodal distribution for tractability and for consistency with the authors' data. The bimodal transit time distribution is created by mixing two normal distributions. The lanes are represented with long right tails using lognormal distribution.

Evidence of Nonunimodality in Transit Time Distribution

Hartigan's dip test is done on the transit time values of lanes on container-level data available. The lanes are identified as unique origin–destination pairs for all data. It was observed that nonunimodality is prevalent across different kinds of shippers. However, it occurs at different levels for retailer, manufacturer, and freight forwarder data. For the retailer, nonunimodal distributions occurred in only 2% to 4% of origin–destination lanes but account for 12% of shipment volume. However, for the manufacturer, the corresponding number averaged for 22% of lanes accounting for 75% of shipment volume. For the freight forwarder, nonunimodality was observed for 24% of the lanes, equivalent to 85% in shipment volume. A common norm across many trade lanes also indicated that the transit times are heavily right-tailed. Examples of histograms of empirical distributions from available data that have multiple modes and long right tails are shown in Figure 2.



(a)



(b)

FIGURE 2 Distribution of transit time: (*a*) bimodal and (*b*) long right-tailed.

Simulation of Cases of Bimodality and Its Effects

Variability in Transit Time Distribution and Use of Normal Approximation in Hadley–Whitin Formula

For a variable demand and transit time distribution, let

E[DoLT] = mean of demand over transit time,
σ_{DoLT} = standard deviation of demand over transit time,
E[LT] = mean transit time,
σ^2_{LT} = variance of transit time,
$E[D]$ = mean demand during one time period,
σ^2_D = variance of demand, and
k = service level, which is the level at which the probability of demand is always less than the quantity ordered.

Under the assumption that observed demand and transit time are independent and that demand is uncorrelated between the transit time periods, the mean and variance of demand over transit time is given by

$$E[\mathrm{DoLT}] = E[\mathrm{LT}]E[\mathrm{D}]$$

$$\sigma^2_{\mathrm{DoLT}} = E[\mathrm{LT}]\sigma^2_D + (E[D])^2\,\sigma^2_{\mathrm{LT}}$$

Performing a random sum of random numbers derives the above equations.

The normality assumption incorporated in the Hadley–Whitin formula (*3*) gives the value for the reorder point (R).

$$R = E[\mathrm{LT}]E[D] + k\sqrt{\left(E[\mathrm{LT}]\sigma^2_D + (E[D])^2\,\sigma^2_{\mathrm{LT}}\right)}$$

The above equation for the reorder point is derived when demand over transit time is approximated to be a normal distribution.

Creation of Bimodal Distribution from Mixture of Two Normal Distributions

Bimodality can be simulated as a mixture of two normal distributions. Applications of mixing two normal distributions to create multimodal distributions have been seen in many fields, such as economics, finance, and astronomy.

A bimodal distribution is simulated at a certain mixture rate, π, such that any point in the resultant distribution lies in the first normal distribution with a probability of π and in the second distribution with a probability of $1 - \pi$. The probability density function of the resulting mixture distribution of the transit time is obtained as a linear combination of two normal distributions such that

Probability density function:

$$f(x) = \pi f_1(x) + (1-\pi) f_2(x)$$

Cumulative distribution function:

$$F(x) = \pi F_1(x) + (1-\pi) F_2(x)$$

where $f_i(x)$ has a mean μ_i and standard deviation σ_i and $0 \le \pi \le 1$.

Let the resulting bimodal distribution have a mean μ and standard deviation σ. By using formulas from probability theory, the authors obtain

$$\mu = \pi\mu_1 + (1-\pi)\mu_2$$

$$\sigma^2 = \pi\left(\mu_1^2 + \sigma_1^2\right) + (1-\pi)\left(\mu_2^2 + \sigma_2^2\right) - \left(\pi\mu_1 + (1-\pi)\mu_2\right)^2$$

Different levels of bimodal distributions are obtained by changing the mean of one of the normal distributions while keeping the mean of the other distribution constant.

Simulation Model Used for Calculation of Logistics Cost

The second part of the analysis includes evaluation of cost for the three cases. Inventory calculations involve simulating demand over transit time that is observed in reality and then comparing the inventory levels calculated separately for each of the three cases—deterministic (Case 1), normal (Case 2), and bimodal (Case 3)—to obtain the costs that the firm incurs in case of stocking excess or not enough of the product. Inventory cost is analyzed on the basis of a given critical ratio (service level).

Simulation Structure with Inventory Policy Used

The inventory policy is as follows:

1. Order-up-to level. If the ending inventory level goes below R by x units, x units are ordered.
2. Complete back ordering. Any demand that is not fulfilled in a time period is back-ordered and is satisfied in the next period.
3. Frequency of ordering. Ordering is done at the end of every unit time period.

The order-up-to level when the inventory management system uses the actual distribution is the optimum amount calculated according to the critical ratio that is being observed. The order-up-to level when the system uses normal approximation uses the Hadley–Whitin formula (*3*), shown in the section on simulating cases of bimodality and its effects. The case of deterministic transit time uses an order-up-to level, which is the expected value of demand over transit time. The amount ordered at the end of every time period to reach the respective order-up-to levels is the same for all the three cases, which is equal to the demand observed in that period. Therefore, the ordering and purchase cost is the same across all the three cases. Hence, the net difference in the cost is given by the understocking and overstocking costs.

Derivation of Cost Model

A stationary infinite horizon inventory model is considered, in which the optimal base stock is calculated from the critical ratio. The derivation of the cost model used for the evaluation is adapted from the cost structure derived by Zipkin (*14*) when transit time is a random variable. This cost structure is demonstrated below. The notation is as follows:

γ = discount cost rate $0 < \gamma \le 1$ on fixed ordering cost,
$L(t)$ = transit time that randomly changes over t,

$h(t)$ = inventory holding cost rate at time t,
$b(t)$ = back order penalty cost rate at time t,
$x(t)$ = inventory position at time t before ordering,
$y(t)$ = inventory position at time t after ordering,
$\hat{C}(t, x(t))$ = inventory holding or back order cost at time t,
$C(t, y)$ = expected inventory back order cost,
$[a]^+$ = maximum between $(a, 0)$,
$D(t)$ = demand at time t,
$D[t, t + L)$ = transit time demand starting at time t,
T = time horizon, which could be finite or infinite,
$z(t)$ = order size at time t, and
CSL = customer (cycle) service level.

A given is that the order placed at time t will arrive at some future time denoted by $+L(t)$. The decision-making process that helps in the formulation of the cost model is composed of two steps at time $t < \mathrm{T}$.

Step 1. Net inventory, $x(t)$, is observed.
Step 2. An order size, $z(t)$, is decided.

The assumptions follow:

1. There are no crossovers: orders arrive in the same sequence in which they were issued.
2. Transit time $L(t)$ is independent of demand.
3. Stationary transit time would imply that $L(t)$ has the same distribution over time and is denoted by random variable L.

From work by Zipkin (*14,* p. 409), the expected inventory back order cost $C(t, y)$ after the order is placed in Step 2 can be written as

$$C(t, y) = E\left[\gamma^L \hat{C}(t + L, y - D[t, t + L))\right]$$

where

$$\hat{C}(t, x) = h(t)[x - D]^+ + b(t)[D - x]^+$$

By definition, the inventory position observed just after ordering, is $x(t + 1)$:

$$x(t+1) = y - D[t, t + L)$$

Given stationary transit time and infinite horizon, the cost function becomes the following:

$$C(y) = E\left[\gamma^L \hat{C}(y - D)\right]$$

If an average ordering cost (i.e., no discounting of cost or $\gamma = 1$) is assumed, then

$$C(y) = E\left[\hat{C}(y - D)\right]$$

By definition, for any inventory level i,

$$\hat{C}(i) = h[i - D]^+ + b[D - i]^+$$

In a single or unit period of time, the quantity that maximizes profit or minimizes the total cost for a firm is given by solution to the newsboy or the newsvendor problem.

$$F(Q^*) = \frac{b}{b + h} = \mathrm{CSL}$$

But by definition $F(Q^*)$ = CSL or the level at which the probability of demand is always less than the quantity ordered (Q). This service level is equivalent to the critical ratio in this paper. Returning to infinite horizon problem, the general cost equation from above becomes

$$C(y) = E\left[h[y - D]^+ + b[D - y]^+\right]$$

Dividing and subtracting a term $(b + h)$ gives the following:

$$(b+h)\left\{\frac{h}{b+h}\left(E[y - D]^+\right) + \frac{b}{b+h}\left(E[D - y]^+\right)\right\}$$
$$= (b+h)\left\{(1 - \mathrm{CSL}) * \left(E[y - D]^+\right) + \mathrm{CSL} * \left(E[D - y]^+\right)\right\}$$

For given values of (b, h), the effective cost is given by

$$\text{effective cost} = \left\{(1 - \mathrm{CSL}) * \left(E[y - D]^+\right) + \mathrm{CSL} * \left(E[D - y]^+\right)\right\}$$

For a given value of $(b + h)$, various values are possible for CSL. However, because CSL is a probability value, it can range only between [0, 1]. This limit enables the authors to bind the calculations for the cost for values of CSL in the range of [0, 1]. Therefore, there are not infinite options for (h, b) values. Hence, the analysis is simpler. The term $[y - D]^+$ is positive when the shipper carries excess inventory and $[D - y]^+$ is positive when demand is back ordered.

DISCUSSION OF RESULTS

The plots that are used to understand the relative cost differences between the cases use the mean of the resultant bimodal distribution as a proxy for levels of bimodality because it evenly spaces the points on the plot.

Comparing Deterministic Case 1 and Normal Approximation Case 2

Generally, ignoring variability in transit time could have grave impacts on the inventory cost. However, digging deeper reveals that this statement is not true under all values of critical ratios and levels of bimodality. On the basis of the results, it is recommended that shippers

- May choose to ignore variability and avoid updating their inventory management system to account for variability for (*a*) low service levels (critical ratio < 0.6) for all levels of bimodality and (*b*) intermediate service levels (0.6 ≤ critical ratio < 0.8), but only for low levels of bimodality, and

• Should consider updating their inventory management system to account for variability to allow for approximating the distribution to normal for the remaining combinations of critical ratios and levels of bimodality (especially for high service levels (critical ratio ≥ 0.8) under all levels of bimodality).

Relative difference is obtained by dividing the difference of the costs between the cases by the cost obtained in Case 2. In terms of positive and negative values, if the relative difference is negative, it implies that the cost obtained by ignoring variability (Case 1) is more expensive than considering a normal approximation of the transit time distribution (Case 2) (see Figure 3).

A clear demarcation is observed across levels of bimodality and critical ratios as shown in Figure 4. The first range that corresponds to positive values of relative difference is seen in two regions:

• For low levels of critical ratios (<0.55) and over all levels of bimodality and
• For intermediate critical ratios but only for low levels of bimodality.

The magnitude of the values of this range is only 0% to 1%. So the difference is very low and could probably also be attributed to variations caused by simulation. However, inventory managers can choose to be indifferent between the two cases for low levels of bimodality and low critical ratios. The reason behind negligible yet positive differences between the two cases is the result of low critical ratios. The normal approximation simulation model is forced to stock less to avoid high surplus costs. The result is a higher logistics cost caused by more back orders than occur in the case of deterministic transit time.

The second range covers values for which maximum of the absolute of the relative difference is less than 10%. The relative difference is negative, implying that it is more expensive to ignore variability. This range encompasses values from medium to high values of mean of the resultant distribution and intermediate values of critical ratios (0.6 ≤ critical ratio < 0.85). The authors observe worse effects of approximating the transit time distribution to a deterministic value as compared with a normal approximation when there is a larger variability (CV) in transit time.

The third range corresponds to situations that are most affected by ignoring variability in transit time. This range covers all relative differences for which absolute value is greater than 10%. In fact, the maximum value of this difference can go as high as 395%. The extremely high differences correspond to situations of very high critical ratios (≥0.85) and very high levels of bimodality. With higher variability, the cost structure penalizes the deterministic case much more than low variability distributions, and hence it becomes more expensive than normal approximation.

Comparing Deterministic Case 1 and Bimodal Case 3

The relative cost difference plotted in Figure 5 is obtained by dividing the difference between the costs by the cost obtained in the bimodal case.

On the basis of the results, the recommendations for the shipper are similar to those in the comparison of Case 1 and Case 2. The difference, however, is that the magnitude of relative differences becomes higher in the comparison of Case 1 and Case 3, because using the actual distribution makes the effect of ignoring variability worse. The differences are all negative, implying that Case 1 is always more expensive (could be approximately six times) than Case 3, except when the difference is minimal or zero.

It is easy to see the demarcation in the difference between Case 1 and Case 3 with different levels of bimodality and critical ratios.

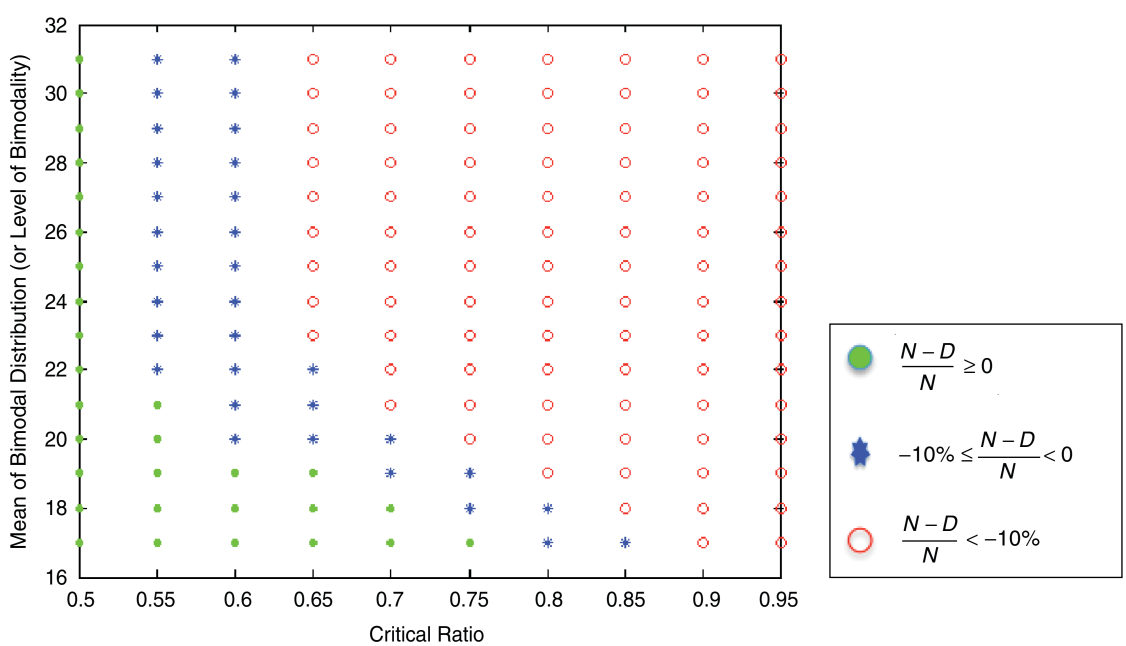


FIGURE 3 Relative difference between logistics cost of Case 1 [deterministic (*D*)] and Case 2 [normal approximation (*N*)].

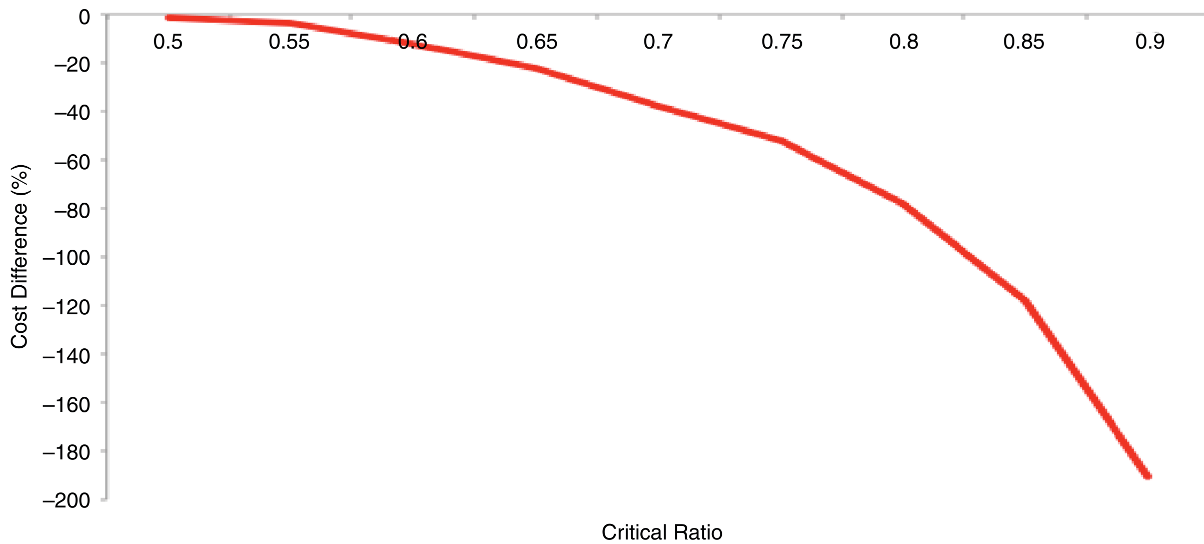


FIGURE 4 **Change of relative cost difference with increasing critical ratio for given level of bimodality.**

The first range corresponds to an absolute difference of less than 15%, which occurs

- For low levels of critical ratios (≤0.55) and across all levels of bimodality and
- For intermediate critical ratios (between 0.6 and 0.7), but only for low levels of bimodality.

The second and the third ranges, which correspond to moderate values between 15% and 50%, correspond to intermediate critical ratios (0.65 ≤ critical ratio ≤ 0.8). The plot indicates that these ranges progressively transition from lower critical ratios and high levels of bimodality to higher critical ratios and all levels of bimodality.

The fourth and fifth ranges, corresponding to high values between 50% and 100%, behave in a manner similar to the trend in the second and the third ranges of progressive increase across critical ratios and levels of bimodality. The ranges occur in critical ratios between 0.75 and 0.9. For ranges between 0.75 and 0.8, the high differences occur for high levels of bimodality. These differences also occur for lower levels of bimodality in cases of higher critical ratios (between 0.8 and 0.9).

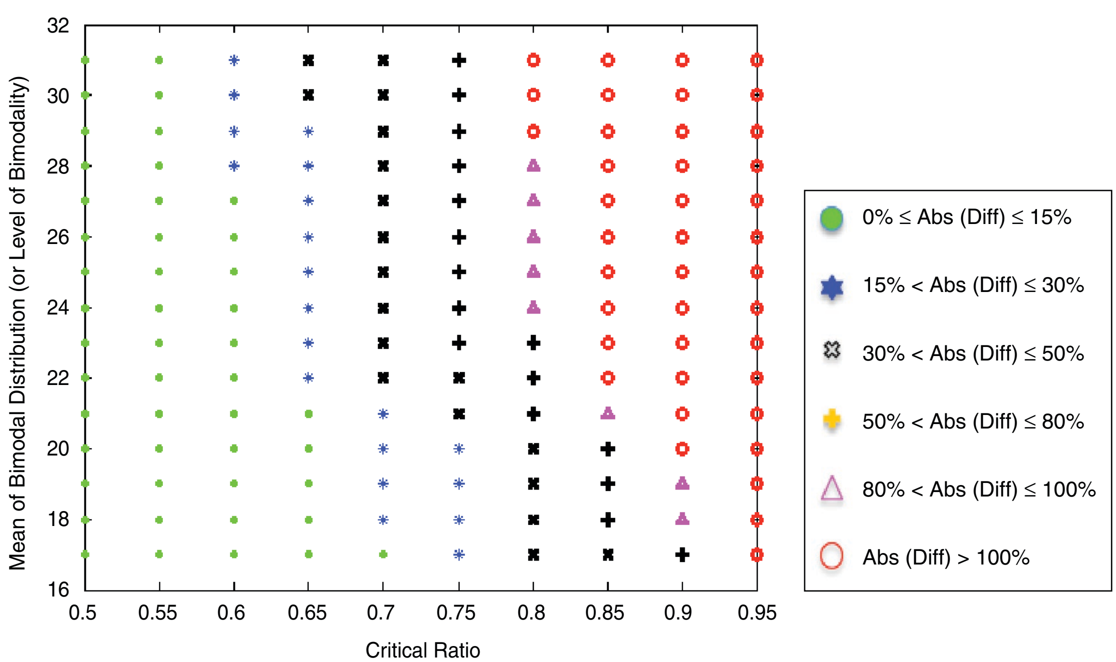


FIGURE 5 **Difference between logistics cost of Case 1 [deterministic (*D*)] and Case 3 [bimodal (*B*)] [Abs(Diff) = absolute difference = *D* – *B*/*B*].**

The final, very high, value range, corresponds to cost differences that are greater than 100%. As expected, they occur at regions that correspond to high critical ratios and high levels of bimodality. However, for a very high critical ratio of 0.95, they occur over all levels of bimodality.

Summarily, for a given level of variability in bimodal transit time, the effect of completely ignoring variability becomes worse as critical ratio increases. It becomes more and more expensive to ignore variability in transit time with high critical ratios and high levels of bimodality. The trend in the relative cost difference with increasing critical ratio for a given level of bimodality is plotted in Figure 4.

Comparing Normal Approximation Case 2 and Bimodal Case 3

Figure 6 shows the relative cost difference between Case 2 and Case 3. The relative difference is obtained by dividing the difference between the costs in the two cases by the cost obtained in the bimodal case. The authors observe three main regions on the basis of the values of critical ratios and levels of bimodality.

In the first, small difference, range, the maximum value of differences is 5%, corresponding to

- Low critical ratios and low levels of bimodality and
- Overall critical ratios for levels of bimodality that are equivalent to the mean of the resultant distribution of less than 21 units.

The second range incudes moderate levels of differences that are between 5% and 10%, corresponding to critical ratios and levels of bimodality that are greater than those in the range of small differences. These differences occur for intermediate critical ratios (around 0.8) for higher levels of bimodality than in the previous range.

The final range of high differences occurs in the middle regions of the plot that correspond to the intermediate values of critical ratios and high levels of bimodality. They are also observed for very high levels of bimodality and very high critical ratios (0.95).

It was observed that the differences peak in terms of magnitude at intermediate critical ratios and in the case of large critical ratio of 0.95 for high levels of bimodality. The differences also show an increasing pattern up to a peak followed by decreasing pattern and finally an increasing pattern at very high critical ratios and high levels of bimodality. This V-shaped trend (corresponding to critical ratios between 0.5 and 0.9) is shown in Figure 7.

At low critical ratios (e.g., 0.5), which imply that surplus cost is equal to shortage cost, the cost from normal approximation and optimal should be close because optimal stocking quantity is close to average of demand over transit time. As the critical ratio increases, the normal approximation case stocks more to account for the growing shortage cost and so the difference between the logistics cost increases.

However, the change in critical ratio also changes the cost structure, such that the relative difference between per unit shortage and the surplus cost changes. As critical ratio increases even further, the stocking quantity in the normal approximation simulation increases, but it also penalizes little for overstocking (but penalizes a lot for understocking). Hence, the difference in costs goes down.

Results for Lognormally Distributed Transit Time

Similar to the analysis for bimodal distribution, lognormal distributions of transit time were analyzed by comparing the three cases of deterministic, normal approximation, and lognormal. The results were consistent with that of bimodal distribution. The CV of the lognormal distribution is used for comparing the three cases.

In the comparison of the deterministic and normal approximation cases, it was observed that below a critical ratio of 0.7, it is better to opt for the deterministic case, as shown in Figure 8. The interactions with shippers suggest that they usually fall in the region that corresponds

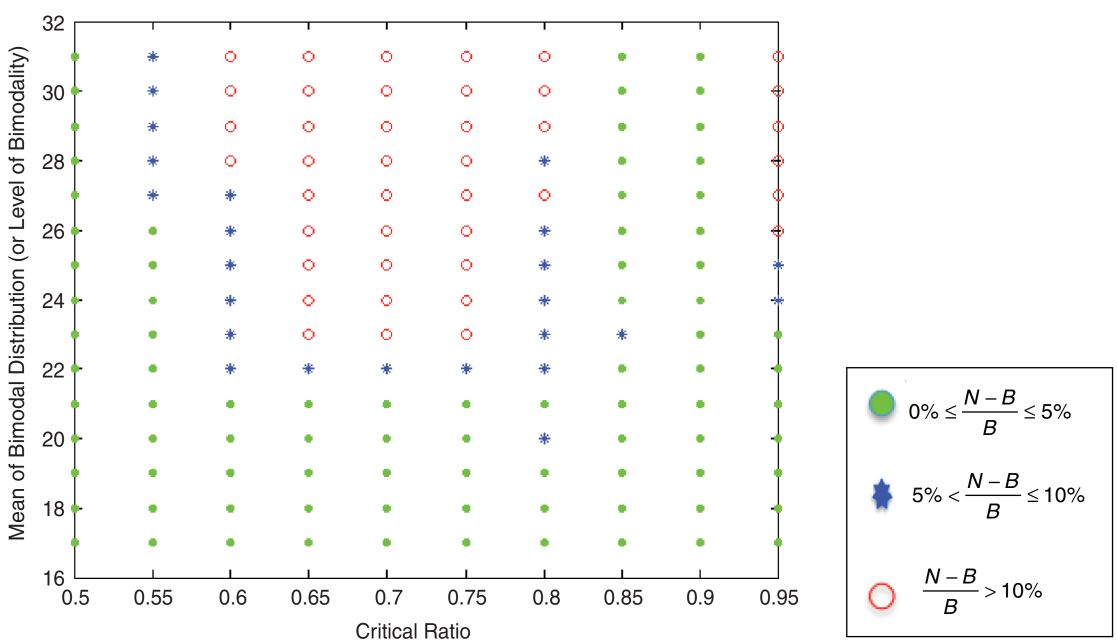


FIGURE 6 Difference between logistics cost of Case 2 [normal approximation (*N*)] and Case 3 [bimodal (*B*)].

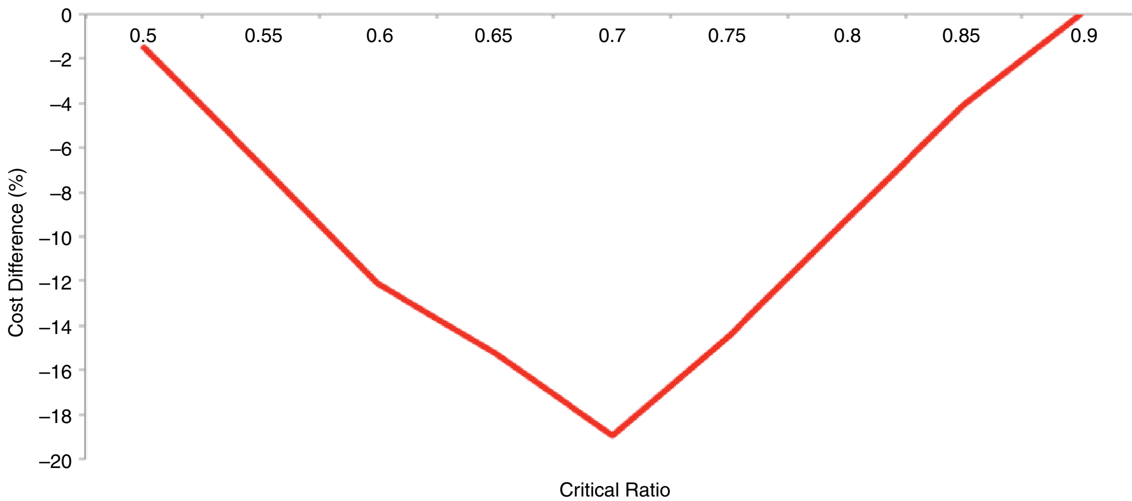


FIGURE 7 **Change of relative cost difference with increasing critical ratio for given level of bimodality.**

to high critical ratios (service level) and high variability and hence are worse off ignoring variability.

The difference between the deterministic case and the actual distribution (lognormal here) keeps increasing with critical ratio and variability of the distribution of transit time. This result is consistent with the results when the distribution of transit time is bimodal.

Finally, the comparison between the normal approximation and the actual distribution shows that the difference between the cases decreases with the critical ratio. The trend is hence not V-shaped as observed in the case of bimodal distribution. Further research is needed to understand this difference.

Results for Empirical Nonunimodal Distributions

The authors now show the use of empirical information available to understand the cost implications under the three cases for the transit time distributions of trade lanes that were identified as nonunimodal. The trends of the relative cost difference with respect to empirical distributions were found to be similar to those discussed in the previous section for theoretical bimodal distributions.

The authors used transit time information of a leading manufacturing firm, one of the four shippers from the data set used to characterize variability in the distribution. The authors evaluated the 17 (of 73) trade lanes that were identified as nonunimodal by Hartigan's dip test (*13*). The results of one such empirical distribution are discussed next. The histogram of the transit time for a trade lane, shown in Figure 9, suggests a bimodal distribution.

It was observed that with increasing critical ratios, the relative difference between the inventory costs under the normal approximation and the use of actual empirical distribution (Figure 10) follows a V-shaped trend similar to the one observed in the theoretical distributions (Figure 7). A negative relative cost difference suggests

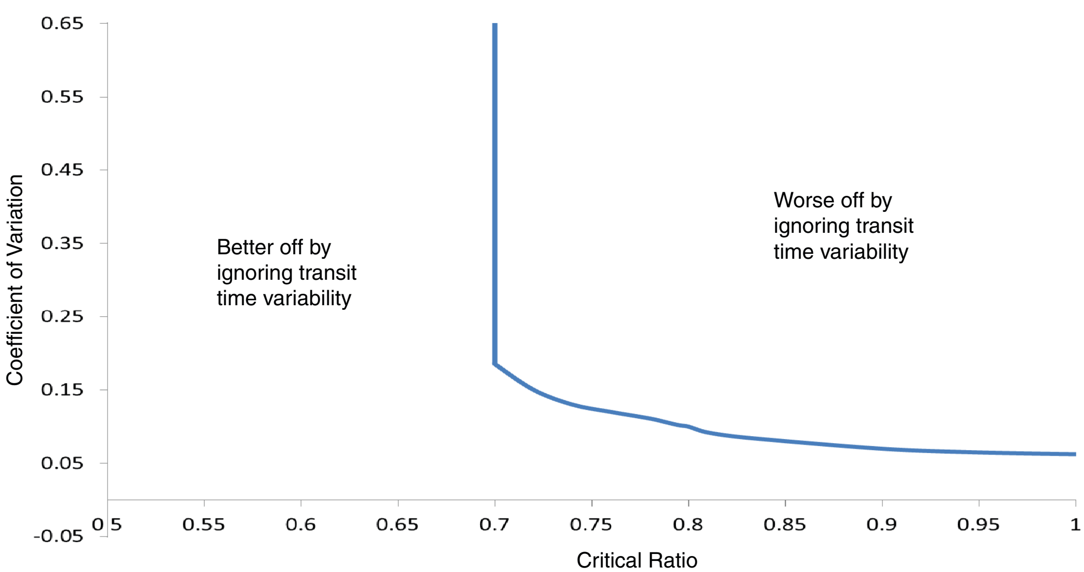


FIGURE 8 **Comparison of deterministic versus normal approximation when transit time is lognormal.**

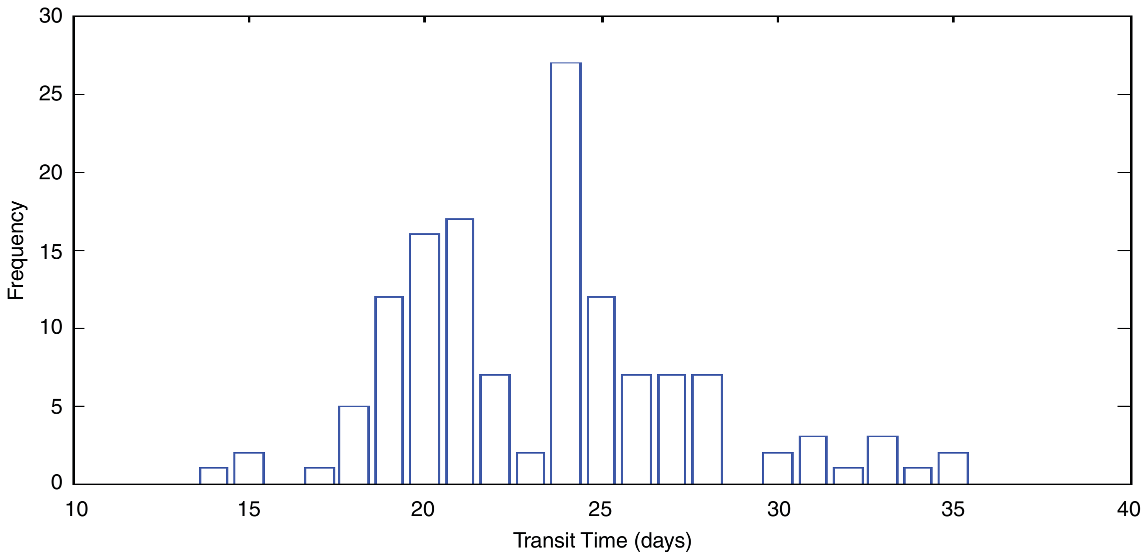


FIGURE 9 Histogram of empirical nonunimodal transit time distribution for shipper.

that the normal assumption case is more expensive than the actual distribution for inventory decisions and logistics.

Similarly, the relative cost difference between the deterministic and the actual distribution increased tremendously with increasing critical ratios, showing a trend exactly the same as in Figure 4.

Although the results provide an idea about the trend of relative logistics cost difference with increasing critical ratios, further research is required to evaluate the relationship between the magnitude of this difference and that observed in the theoretical distribution.

CONCLUSION

Multimodality and long right-tailed distributions were observed in ocean transit time distribution in the data set available. Cumulatively, the percentage of lanes that were not unimodal was 17% of the total lanes, but these lanes constitute large shipment volumes, totaling about 85%. The analysis to understand the effect of bimodality (and lognormality) in transit time was done by comparing three ways of approaching variability in transit time: deterministic, normal approximation, and bimodal (or lognormal) cases. It was observed that it makes sense to use inventory management systems, which do not have the added capabilities to account for variability, for low service levels (critical ratio < 0.55) under all levels of bimodality. This finding is also true for intermediate service levels (≤0.7), but only for low levels of bimodality. However, the shippers should consider updating their inventory management systems to account for variability in situations of high service levels desired and high variability of bimodal transit time distributions. As expected, the system that uses the actual distribution to make inventory management decisions gives the best cost results. The case when the distribution is approximated to normal and the one that uses actual

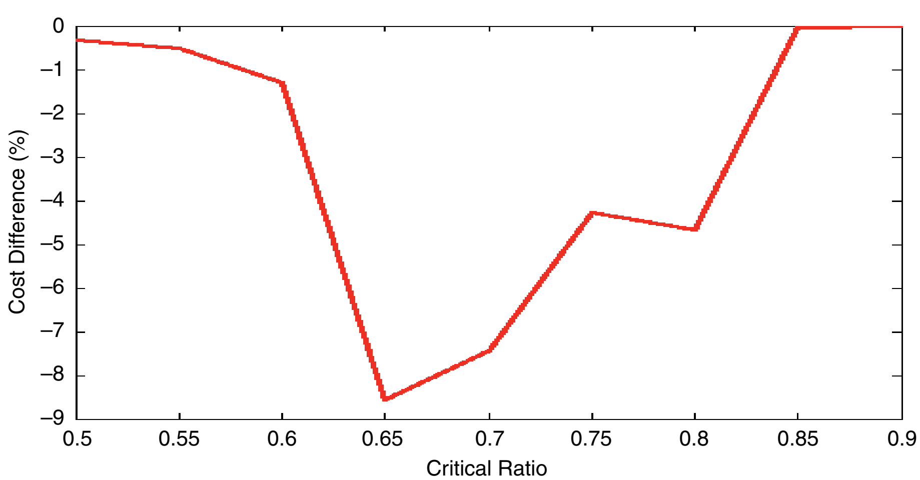


FIGURE 10 Change of relative cost difference between normal approximation and using actual distribution across increasing critical ratio.

bimodal distribution differ the most in terms of cost at a critical ratio of 0.65 or 0.7 for high levels of bimodality. The shape of the relative difference between the costs is a V-shaped trend. The results for the lognormal distribution were found to be consistent with that of bimodal distribution.

The authors see the following potential extensions of this research. First, it would be interesting to analyze the theoretical bounds of the relative difference in logistics costs for the three cases discussed as a function of parameters of the transit time distribution and the service levels targeted. Next, it would be worth investigating the occurrence of cases of higher levels of multimodality (besides bimodality) and their effects on logistics costs. This paper was able to identify non-unimodality only in transit time distributions. Bimodality was used because it is the simplest case of multimodality. In addition, frequency and impact of bimodally and lognormally distributed transit time lanes warrant a thorough investigation of the reasons that lead to such a phenomenon of large differences in logistics cost in the three cases. This investigation would enable shippers to predict bimodality (or lognormality) of the lane, given a set of trade lane characteristics. It would also be useful to further validate the results pertaining to the two distributions with expert interviews. Finally, an investigation could be conducted of the optimal inventory frequency of ordering, order-up-to levels, and stocking policy that should be used when the shippers face bimodality or lognormality of transit time.

REFERENCES

1. Rodrigue, J. *The Geography of Transport Systems,* 2013. http://people.hofstra.edu/geotrans/eng/ch3en/conc3en/worldcontainerflows.html. Accessed March 13, 2014.
2. Bloomberg News, 2012. http://www.bloomberg.com/news/2012-06-21/container-ships-miss-boat-as-maersk-pushes-punctuality-freight.html. Accessed March 13, 2014.
3. Hadley, G., and T. M. Whitin. *Analysis of Inventory Systems.* Prentice Hall, Upper Saddle River, N.J., 1963.
4. Center for Transportation and Logistics, MIT Global Ocean Transportation Roundtable, Nov. 2012. http://ctl.mit.edu/events/global_ocean_transportation_roundtable.
5. Lau, H. S., and A. Zaki. The Sensitivity of Inventory Decisions to the Shape of Lead Time–Demand Distribution. *IIE Transactions,* Vol. 14, No. 4, 1982, pp. 265–271.
6. Tadikamalla, P. R. A Comparison of Several Approximations to the Lead Time Demand Distribution. *Omega,* Vol. 12, No. 6, 1984, pp. 575–581.
7. Naddor, E. Note-Sensitivity to Distributions in Inventory Systems. *Management Science,* Vol. 24, No. 16, 1978, pp. 1769–1772.
8. Eppen, G. D., and R. K. Martin. Determining Safety Stock in the Presence of Stochastic Lead Time and Demand. *Management Science,* Vol. 34, No. 11, 1988, pp. 1380–1390.
9. Tyworth, J. E., and L. O'Neill. Robustness of the Normal Approximation of Lead-Time Demand in a Distribution Setting. *Naval Research Logistics (NRL),* Vol. 44, No. 2, 1997, pp. 165–186.
10. Lau, H. S., and A. H. L. Lau. Nonrobustness of the Normal Approximation of Lead-Time Demand in a (Q, R) System. *Naval Research Logistics (NRL),* Vol. 50, No. 2, 2003, pp. 149–166.
11. Bagchi, U., J. C. Hayya, and C. H. Chu. The Effect of Lead-Time Variability: The Case of Independent Demand. *Journal of Operations Management,* Vol. 6, No. 2, 1986, pp. 159–177.
12. Bookbinder, J. H., and A. E. Lordahl. Estimation of Inventory Re-Order Levels Using the Bootstrap Statistical Procedure. *IIE Transactions,* Vol. 21, No. 4, 1989, pp. 302–312.
13. Hartigan, J. A., and P. M. Hartigan. The Dip Test of Unimodality. *The Annals of Statistics,* 1985, pp. 70–84.
14. Zipkin, P. H. *Foundations of Inventory Management,* Vol. 2. McGraw-Hill, New York, 2000.

The International Trade and Transportation Committee peer-reviewed this paper.