Transportation Research Part E 45 (2009) 678–692

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Consolidation effects: Whether and how inventories should be pooled

Peter F. Wanke *, Eduardo Saliby

Center for Logistics Studies, The COPPEAD Graduate Business School, Federal University of Rio de Janeiro, Rua Paschoal Lemme, 355, Cidade Universitária, Rio de Janeiro, CEP. 21.949-900, Brazil

ARTICLE INFO

Article history:
Received 19 August 2008
Received in revised form 3 December 2008
Accepted 31 January 2009

Keywords:
Consolidation effect
Total cost
Service level
Regular transshipment
Inventory centralization
Independent system

ABSTRACT

This paper presents a framework for deciding whether and how inventories should be pooled, using the consolidation effect as a cornerstone tool to measure inventory costs, service levels, and total costs. Based on the random generation of different scenarios, it is indicated the adequacy of inventory centralization, regular transshipments, and independent systems to a given set of demand, lead time, and holding costs characteristics. Sensitivity analyses on mathematical expressions are performed to determine when one alternative is preferable in terms of total costs. Real settings are also presented in light of the framework developed.

1. Introduction

Research into virtual inventory management, variously referred to as inventory-pooling, inventory consolidation, portfolio effect, and consolidation effect, has been around in the literature in several forms for quite some time (Landers et al., 2000; Ballou and Burnetas, 2003). The basic idea is that inventory increases as the standard deviation of either demand or lead time increases, and, as a result, companies may attempt to reduce inherent variation by pooling it. The statistical economies that arise from uncertainty pooling can be achieved in numerous ways, including not only inventory centralization, but also transshipments and order splitting (Evers, 1999). Frequently, these economies are determined based upon an independent distribution system, in which each one of the markets is exclusively served by dedicated facilities.

Inventory centralization physically consolidates stock at a limited number of locations (often a single facility) from which all demand is satisfied. Since the mid 1970s, a substantial amount of research on consolidation effects and inventory portfolios has greatly enhanced the understanding of the effects of facility consolidation on inventory levels, total costs, and the major variables to account for. Most of this research originated from the seminal works of Maister (1976) and Zinn et al. (1989) and their basic idea of using the ‘square-root law’ to measure safety stock savings derived from consolidation. Generally speaking, inventory centralization results in demand pooling (Eppen, 1979); stocks should be consolidated at the facility with the smallest standard deviation of lead time, and distribution costs are higher when compared to decentralized systems (Wanke, 2009).

Order splitting occurs when a stock keeping location operates independently of all facilities in filling its demand, but divides its reorders (not necessarily evenly) among multiple suppliers (Evers, 1999). A great deal of research has focused on order splitting. For instance, Sculli and Wu (1981) examined two fundamental parameters in the case of two

* Corresponding author. Tel./fax: +55 21 2568 41 58.
E-mail address: peter@coppead.ufrj.br (P.F. Wanke).

1366-5545/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tre.2009.01.006

suppliers: the effective lead time and the time between the arrival of the first and the second shipment. Hayya et al. (1987) found that order splitting outperforms the use of a single vendor on the basis of backorder reduction. Other studies of the two-supplier case indicate that order splitting alone does not lead to demand pooling, but it is a way to pool lead times.

More recently, Thomas and Tyworth (2006) presented a comprehensive and critical review, over the past 25 years, on pooling lead time risk by order splitting. Although studies provide hypothetical evidence that pooling lead time risk by simultaneously splitting orders may be worthwhile, the authors affirm that it is rather difficult to find substantive empirical validation for this proposition. The reason for that is the lack of attention to transport economies of scale, as well as to the safety stock benefits from a total system cost perspective.

Transshipment occurs when a facility satisfies whatever demand coming from another territory. On a regular basis, transshipment implies that a given proportion of demand is supplied from facilities located in different markets, regardless of whether there is sufficient inventory in the original serving facility. If transshipment results because the original serving facility is out of stock, then it will be known as an emergency transshipment (Evers, 1997).

Much research on transshipments has focused on mathematical modelling of particular systems. Krishnan and Rao (1965) modelled the costs for emergency transshipments from n locations. Tagaras (1989) generalized Krishnan and Rao's cost structure. Lee (1987), Axsäter (1990), and Kukreja et al. (2001) used queuing theory methods to estimate the effect of transshipments in continuous review multi-location inventory systems with stochastic demand. Evers (1999) compared emergency transshipments to order splitting via simulation, analyzing their drivers on total costs and service levels. The study indicated that emergency transshipments tend to perform better than order splitting in terms of reductions in stock outs and order placements. Wong et al. (2007) described how game theoretic models are applied to analyze the cost-allocation problem in the context of spare parts inventory transshipment. In general, all these studies confirm that transshipments reduce stock outs by improving inventory availability when compared to centralized systems. This occurs, however, to the detriment of higher distribution costs associated to longer distances for a given market (in case of regular transshipments) and to expedited shipping (in case of emergency transshipments).

Evers' (1996 and 1997) and Ballou and Burnetas' (2003) studies specifically link transshipments to portfolio and consolidation effects. Evers (1996) showed how the portfolio effect model could be employed to evaluate the use of regular transshipments so as to reduce safety stocks. The author also assessed under what conditions it might be effective to do so. The analysis neither considered the effect of transshipments on cycle stocks nor the impact on distribution costs. The author found that the use of regular transshipments should result in significant inventory reductions in cases where markets exhibit high degrees of demand variability. In Evers (1997), emergency transshipments were included into the portfolio effect analysis indicating that they result in the pooling of demand and lead time uncertainties. Finally, Ballou and Burnetas (2003) incorporated cycle stock consolidating effects to the portfolio effects on safety stock, when studying the cross filling (transshipments) between locations. Several restrictive assumptions were made regarding no lead time uncertainty, uncorrelated demands, and equal inventory-related costs. The authors illustrated the countervailing inventory forces that affect the cross filling decision.

This paper aims to explore the impacts of inventory centralization and regular transshipments on total costs and service levels, employing the consolidation effect model in terms of their underlying demand allocation rules. More precisely, should a given centralized facility supply the same fraction of demand to each decentralized location, like in Tyagi and Das (1998), or should a fraction of the demand be supplied by a primary source and the remainder by secondary sources, like in Ballou and Burnetas (2003)? The impact of both allocation rules on the consolidation effect and their implications in terms of inventory levels, distribution costs, and service levels are addressed in order to show how the consolidation effect model can be used as a cornerstone tool to evaluate under what conditions it may be preferable to adopt one inventory-pooling model to the detriment of the other. These results are also compared to those of an independent distribution system, in which there is no inventory-pooling at all.

Specifically considering Tyagi and Das' (1998) and Ballou and Burnetas' (2003) results, this paper differs from them, not only for considering less restrictive assumptions in terms of lead time uncertainties, demand correlations, and inventory-related costs, but also for developing a framework for deciding whether and how inventories should be pooled in light of different product, lead time, and demand characteristics.

Rather than strictly adopting a given demand allocation rule for a maximal portfolio effect, like in Tyagi and Das (1998), or presenting a quantitative guideline to support the decision of whether or not to cross fill, like in Ballou and Burnetas (2003), this paper places these two demand allocation rules (and, consequently, their underlying inventory-pooling models) in a broader perspective based on the random generation of several different operational scenarios. This perspective allows not only to discriminate the adequacy of these demand allocation rules/pooling models in terms of different operational patterns, but also enables the measurement of their expected ranks in terms of distribution costs, inventory costs, and service levels. The adequacy of an independent system is also considered under this perspective.

This paper starts off by discussing the foundations of inventory consolidation planning; then, the two demand allocation rules are reviewed and linked to their respective inventory-pooling models in the ambit of the consolidation effect. Finally, several analytical expressions are derived and sensitivity analyses are performed, so as to address their adequacy in terms of product, demand, and operation characteristics for a minimal total cost.