Int. J. Production Economics 143 (2013) 518–526

Contents lists available at SciVerse ScienceDirect

Int. J. Production Economics

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

The selective use of emergency shipments for service-contract differentiation

E.M. Alvarez*, M.C. van der Heijden, W.H.M. Zijm

School of Management and Governance, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

A R T I C L E I N F O

Available online 15 March 2012

Keywords:
Inventory
Customer differentiation
Emergency shipments
Service contracts
Critical levels

A B S T R A C T

Suppliers of capital goods increasingly offer performance-based service contracts with customer-specific service levels. To handle such differentiated service levels, we use selective emergency shipments of spare parts. We apply emergency shipments in out-of-stock situations for combinations of parts and customer classes that yield service levels close to the class-specific targets. We develop two heuristics to solve this problem. An extensive numerical experiment reveals average cost savings of 4.4% compared to the one-size-fits-all approach that is often used in practice. Furthermore, it is particularly beneficial to combine our policy with critical levels, which yields an average cost saving of 13.9%.

1. Introduction

To service advanced capital goods (e.g., defense or medical systems), suppliers increasingly offer service contracts to their customers. This particularly applies when system downtime can have serious consequences (e.g., loss of production output, failure of military missions). Service contracts typically contain quantified targets for key performance measures such as a maximum response time in case of failures or a minimum system availability. As users typically value downtime differently, service level agreements may differ among customer groups. For example, the minimum system availability may be 90% or 99%.

In practice, suppliers often service customers that have varying service levels using a uniform logistics fulfillment process (a so-called one-size-fits-all approach, cf. Cohen et al. 2006). This approach can be very costly if a supplier designs the fulfillment process based on the premium service level. Also, standard customers have no incentive to switch to premium contracts. The fulfillment process should thus be such that the actual service levels reflect the contractual agreements. In this paper, we focus on differentiation in *spare parts supply*.

In literature, *critical level policies* are common differentiation approaches. Such policies reserve parts for premium customers once the inventory level drops below a certain threshold. Then, demand from non-premium customers is either backordered or satisfied from a source that is usually assumed to have infinite supply (e.g., a production facility). Although shown to be effective and efficient, there are barriers for implementation in practice. For instance, service engineers responsible for system repair are often unwilling to wait for a part that is in fact in stock when they are primarily accountable for the speed of repair.

These drawbacks prompt us to investigate the *selective use of emergency shipments* as an alternative. Now, we use on-hand stock to meet demand first-come-first-served. If we are out of stock, we can request an emergency shipment from a secondary source. Emergency shipments are both faster and more expensive than regular replenishments. We should thus investigate for which combinations of customer segments and item types it is a viable approach. As main advantage, this approach is easier to implement in practice than critical level policies, while still being a tool for differentiation. We will show that the approach leads to clear savings over using simple one-size-fits all strategies. Also, we will show that it is very effective to *combine* selective emergency shipments and critical level policies.

The remainder of the paper is structured as follows. In Section 2, we discuss relevant literature and state our contribution. Then, we state our optimization problem and solution approach in Sections 3 and 4, respectively. It will become clear that we must analyze various single-item models as building blocks. In Section 5, we analyze these models for the special case with two customer classes. We give the results of an extensive numerical experiment in Section 6. In Section 7, we give conclusions and discuss options for model extension.

2. Literature overview

Our research is related to literature on service differentiation and the use of emergency shipments for parts supply. The service differentiation stream focuses on critical level policies, introduced by Veinott (1965). The optimality of this policy has been shown under periodic review for backordering and lost sales (Topkis, 1968). Under continuous review, optimality has been shown for

* Corresponding author.
E-mail address: e.m.alvarez@utwente.nl (E.M. Alvarez).

0925-5273/$ - see front matter
http://dx.doi.org/10.1016/j.ijpe.2012.02.019

Poisson demand and exponential or Erlang lead times, both for lost sales (Ha, 1997a; 2000) and backorders (Ha, 1997b; De Véricourt et al., 2002; Gayon et al., 2009).

Several approaches have been developed to find (near-) optimal stock levels and critical levels. For *fast movers*, the focus is on continuous demand distributions with unmet demand being backordered. Ha (1997b) shows that it is optimal to only use arriving replenishment orders to clear non-premium backorders if the inventory level is at least the critical level for premium demand. As the mathematical analysis of such models is intractable – we must keep track of the non-premium backorders – heuristics are often used, see e.g., Möllering and Thonemann (2010) and Arslan et al. (2007). Models for *slow movers*, common in service logistics, focus on Poisson demand and one-for-one replenishment (Dekker et al., 2002). Our work is most similar to Kranenburg and Van Houtum (2007), who try to minimize holding and shipment costs in a multi-item multi-class model with class-dependent waiting time restrictions. Unmet demand is satisfied through emergency shipments. The authors use a solution approach based on decomposition and column generation, combined with greedy heuristics.

The second relevant literature stream focuses on emergency shipments, possibly combined with lateral transshipments among locations at the same echelon level. Most papers consider a single- or two-echelon model with a central location that has infinite supply (see e.g., Muckstadt and Thomas, 1980; Hausman and Erkip, 1994). Alfredsson and Verrijdt (1999) combine lateral shipments with emergency shipments for a two-echelon single-item network. Other recent contributions include Van Utterbeeck et al. (2009) and Wong et al. (2007). We have not yet found literature that uses emergency shipments as a differentiation tool.

The contribution of our paper is fourfold: First, we give a *new differentiation approach* in spare parts supply using selective emergency shipments. Second, we develop *two efficient and effective heuristics* to find near-optimal stock levels and shipment strategies. Third, we show the *added value of selective emergency shipments* compared to one-size-fits-all policies and critical level policies. Finally, we show the *added value of combining selective emergency shipments and critical level policies* for service differentiation.

3. Model

We first give an outline of our model. Next, we discuss the validity of our selection of shipment policies (Section 3.2). In Section 3.3, we present our model assumptions and notation. We give the formal optimization problem in Section 3.4.

3.1. Model outline

Consider a local warehouse that supplies various types of parts to multiple customer classes, and a central depot with infinite supply that replenishes the local warehouse. All customers have the same system, with each item in the system being critical (i.e., an item failure causes a system failure). Each customer class has a distinct amount of time it is willing to wait for parts on average. The warehouse fills demand from all classes first-come-first-served. If it is out of stock, the warehouse may backorder the demand or request an emergency shipment from the central depot. We achieve service differentiation by only using emergency shipments for customer classes with tight waiting time restrictions. We expect this to be particularly beneficial for expensive slow movers that often have low fill rates (making the difference between regular and emergency shipment times crucial). Still, it will sometimes be better to avoid stocks altogether and use emergency shipments for all classes. Conversely, for cheap fast movers it is probably better to keep large stocks (avoiding expensive emergency shipments) and use full backordering. The shipment mode should thus depend on both the item characteristics and waiting time constraints per customer class.

In addition to the above model, we also consider a model where critical levels and selective emergency shipments are jointly used for differentiation. This combined model only satisfies demand from on-hand stock if it exceeds the critical level for the customer's class. Unmet demand is met using either backordering or emergency shipments.

The objective in both models is to minimize system holding and shipment costs, under restrictions on the mean aggregate waiting time per class. Firms like Philips Healthcare and Océ Technologies usually have service level requirements with their clients in terms of e.g., average failure resolution times, with delays often being caused by waiting time for spares. Penalties may apply if the supplier violates the agreements, but we have not seen explicit backorder costs in service contracts. Therefore, we do not include penalty costs per unit waiting time in our objective function. Our decision variables are the item stock levels, and the shipment mode (regular, emergency) and critical level for each item and customer class.

3.2. Selection of shipment policies

In our model, we always use emergency shipments in out-of-stock settings if that shipment mode is chosen for a customer class. However, if the pipeline contains many items, the emergency shipment time might exceed the backorder waiting time, making backordering the faster *and* cheaper option. Ideally, we should thus consider the system state and the customer's class when deciding what shipment mode is most effective. Still, we do not consider such policies to keep the notation transparent and reduce computational effort. In the end, we are mainly interested in the suitability of selective emergency shipments for differentiation compared to critical level policies and the "one-size-fits-all" approach.

3.3. Assumptions and notation

3.3.1. Main assumptions

1. *Demand for each item occurs according to a Poisson process.*
2. *An $(S-1,S)$ base stock policy is applied for all items.* In practice, spares often tend to be expensive slow movers. Therefore, holding costs usually dominate ordering costs and hence the optimal ordering quantity is usually 1.
3. *Regular shipment times from depot to warehouse are exponentially distributed.* This assumption facilitates Markov chain analysis. Also, we show in Appendix A that our model is quite insensitive to lead time distribution used.
4. *The shipment time from the local warehouse to the customer is negligible.*
5. *An emergency shipment is shipped directly from central depot to customer* (i.e., *the shipment does not pass through the local warehouse*).
6. *We consider an infinite horizon.* As a result, the mean waiting time for any customer in class j will equal the average waiting time of class j as a whole.

3.3.2. Notation

For each item $i=1,2,\ldots,I$, we denote the mean replenishment lead time by T_i^{reg}, the emergency shipment time by T_i^{em}, the holding costs per time unit by h_i and the additional costs for an emergency shipment over a normal replenishment by EC_i^{em}. The