European Journal of Operational Research 248 (2016) 963–970

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

On the benefits of delayed ordering

A. A. Syntetos[a,*], R. H. Teunter[b], M. Z. Babai[c], S. Transchel[d]

[a] Cardiff University, United Kingdom
[b] University of Groningen, The Netherlands
[c] Kedge Business School, France
[d] Kühne Logistics University, Germany

ARTICLE INFO

Article history:
Received 12 April 2014
Accepted 5 August 2015
Available online 11 August 2015

Keywords:
Inventory control
Delayed ordering
Intermittent demand
(s, S) policy
Marginal cost analysis

ABSTRACT

Practical experience and scientific research show that there is scope for improving the performance of inventory control systems by delaying a replenishment order that is otherwise triggered by generalised and all too often inappropriate assumptions. This paper presents the first analysis of the most commonly used continuous (s, S) policies with delayed ordering for inventory systems with compound demand. We analyse policies with a constant delay for all orders as well as more flexible policies where the delay depends on the order size. For both classes of policies and general demand processes, we derive optimality conditions for the corresponding delays. In a numerical study with Erlang distributed customer inter-arrival times, we compare the cost performance of the optimal policies with no delay, a constant delay and flexible delays. Sensitivity results provide insights into when the benefit of delaying orders is most pronounced, and when applying flexible delays is essential.

© 2015 Elsevier B.V. and Association of European Operational Research Societies (EURO) within the International Federation of Operational Research Societies (IFORS). All rights reserved.

1. Introduction

1.1. Motivation and research relevance

It is well known that for periodic review inventory systems, the order level, order-up-to level (s, S) policy is optimal under quite general conditions (Iglehart, 1963; Sahin, 1990; Scarf, 1959). In particular, the optimality under concern, in the case of backordering of unfilled demand, is associated with: (i) constant ordering cost; (ii) linear stockout and holding cost; and (iii) a fixed replenishment lead time.

The same is not true under continuous review, as is illustrated by the following simple example. Consider an item with a constant lead time L and a larger constant customer inter-arrival time I between unit-sized transactions. Then the optimal policy is obviously to have at most one unit on hand and always reorder $I - L$ time units after a transaction. In other words, compared to the $(s = 0, S = 1)$ policy, each replenishment order should be delayed by $I - L$ time units. An alternative interpretation is that the order is being placed L time units before it is needed to satisfy the next demand, thereby avoiding any time in inventory.

More generally, delaying orders seems suitable whenever the customer inter-arrival times do not exhibit the memory-less property of the exponential distribution. There are various settings where this situation may naturally occur. One is that of a multi-echelon system, where lot-sizing is applied at lower levels. Another occurs in spare parts management, where parts used for corrective maintenance may wear. Empirical results by Porras and Dekker (2008) under a continuous (s, S) system confirm that assuming demand is driven by a Poisson process results in overstocking spare parts having 0/1 demands. Numerous papers in the area of spare parts modelling assume a continuous review system; interested readers are referred to Kennedy, Patterson, and Fredendall (2002) for an overview in this area.

The demand for spare parts is known to arrive sporadically/intermittently and to be driven by increasing failure rate (IFR) distributions. This is true not only for engineering spares but for service parts kept at the wholesaling/retailing level as well. The stock-bases in the military context, process industries, aerospace, automotive and IT sectors are also dominated by such items. Two very comprehensive benchmarking reports by Aberdeen Group (2005) and Deloitte (2006) identify the increasing importance of after-sales service and parts business (please refer also to Inderfurth and Kleber (2013)). As stated in the latter report, the combined revenues of many of the world's largest manufacturing companies are more than 1.5 trillion US dollars. Further, on average, service revenues account for more than 25% of the total business, so delaying orders for (expensive) spare parts can have a considerable effect on the bottom line. For example, Dickinson (2013) states that "the rotable pool of high value assets for the EuroFighter is managed through delayed

* Corresponding author. Tel.: +44 2920 875672.
E-mail addresses: SyntetosA@cardiff.ac.uk (A.A. Syntetos), R.H.Teunter@rug.nl (R.H. Teunter), Mohamed-Zied.Babai@kedgebs.com (M.Z. Babai), sandra.transchel@the-klu.org (S. Transchel).

http://dx.doi.org/10.1016/j.ejor.2015.08.003
0377-2217/© 2015 Elsevier B.V. and Association of European Operational Research Societies (EURO) within the International Federation of Operational Research Societies (IFORS). All rights reserved.

ordering practices". Similarly, in many organisations, Maintenance, Repair and Operations (MRO) inventory accounts for as much as 40% of the annual procurement budget (Donnelly, 2013). Thus, small improvements regarding the management of the relevant inventories may be translated to substantial cost savings; whereas it is also true to say that any research in this area has a direct relevance to a wide range of companies and industries.

In addition, it is also worthwhile noting that demand patterns in Business-to-Business environments (B2B) are all too often determined by the degree of heterogeneity of the client base (Bartezzaghi, Verganti, & Zotteri, 1999). Heterogeneous requests occur when the potential market consists of customers with considerably different sizes, e.g., a few large customers coexist with a number of small customers. (Similarly, in the MRO environment planned maintenance and breakdowns may also introduce differences in order inter-arrival time distributions.) The higher the heterogeneity of customers, the higher the demand lumpiness, since periods with high requests from a large customer alternate with periods with low or no requests at all from small customers. Alternatively, following a request from a large customer, it is unlikely that another demand will be received in the near future necessitating a delayed ordering mechanism on the part of the supplier. The potential correlation between customers' requests further induces lumpiness. Correlation may be due, amongst other reasons, to imitation and fashion, which induce similar behaviours in customers so that sudden peaks of demand may occur after periods of no requests.

Collective consumer behaviour may be modelled through what are often termed in the literature as 'censored Poisson' processes, whereby the pth event of a Poisson process is only recorded, resulting in inter-event Erlang (of order p) distributions (e.g., Chatfield & Goodhardt, 1973). The discussion conducted in this section also illustrates the compound nature of the demand and the need to take this into consideration, if a realistic inventory model is to be developed.

1.2. Research background

Order delays in a continuous review setting have not received sufficient attention in the literature. To the best of our knowledge, Katircioglu (1996), Moinzadeh (2001), Moinzadeh and Zhou (2008), Schultz (1987, 1989) and Axsäter and Viswanathan (2012) are the only authors who discuss this issue. Schultz (1987) considers $(S-1, S)$ policies and assumes for tractability that the probability of the sum of two demands being less than S is negligible, which is quite restrictive. He shows that a constant delay in placing an order can result in significant holding-cost reductions with little additional risk or cost of stockouts.

Schultz (1989) discusses a different, but again very restrictive setting. He assumes unit-sized transactions and only considers the $(s = 0, S = 1)$ policy. Furthermore, there is instantaneous emergency replenishment in case of shortages. Results are given for the optimal delay for customer inter-arrival distributions with increasing failure rates. Specific expressions for the optimal delay are given for several commonly used distributions, including the Erlang distribution.

Moinzadeh (2001) considers a somewhat more general setting, but still restricted to unit-sized transactions and $(S-1, S)$ policies. Each order is delayed by a constant period of time, independent of demand activities during that period. For general customer inter-arrival times, Moinzadeh (2001) develops an efficient heuristic for computing the policy parameters. He evaluates the performance of the heuristic via a numerical experiment for the cases with Erlang and Uniform customer inter-arrival times.

The studies by Katircioglu (1996) and Moinzadeh and Zhou (2008) are more general than those discussed so far in that they consider (a) unrestricted order levels $(s < S)$ and (b) more sophisticated policies that end a delay when a new demand occurs. However, both models still assume unit sized demands. They obtain similar optimality conditions, albeit through different sorts of analysis. Both also provide numerical results that indicate significant potential savings from order delays. Katircioglu (1996) proves that the optimal policy is of this type. Moinzadeh and Zhou (2008) extend their analysis and results to a two echelon setting with a single warehouse that delays orders and multiple retailers.

Axsäter and Viswanathan (2012) consider the case of a supplier who faces an Erlang demand process from a downstream customer with constant order sizes. They develop an algorithm to determine the optimal ordering time delay when the supplier controls its inventory according to a reorder point (R, nQ) installation stock policy and no information sharing takes place between the supplier and the customer. A numerical investigation shows substantial cost savings when the optimal time delay policy is used (instead of the installation stock policy without delay). These cost savings are also shown to be more substantial than those obtained when the installation stock policy without delay is used in conjunction with inventory information sharing between the customer and the supplier.

1.3. Contributions and organisation of the paper

In this paper, we provide the first analysis of (s, S) policies in a single echelon inventory system with order delays for compound demand processes. So, we drop the assumption that demands are unit-sized. As discussed before, this is an important generalisation since intermittent (spare parts) demand series, for which delaying orders is particularly suitable, are usually very lumpy (Boylan, Syntetos, & Karakostas, 2008). Related to this more general setting, we also consider more flexible delay policies where the maximum delay depends on the order quantity. Like in the studies of Katircioglu (1996), Moinzadeh and Zhou (2008) and Axsäter and Viswanathan (2012), an order is only delayed for this long if no demand happens before then.

For general customer inter-arrival times, we derive conditions that can be used to determine the optimal maximum delay times for any order quantity. This is done using a marginal cost analysis. The exact form of the optimality conditions depends on the specific type of customer inter-arrival distribution. For the purpose of our (numerical) analysis, Erlang distributed customer inter-arrival times will be assumed. The case of Erlang distributed customer inter-arrival times is obtained if demand originates from the lot sizing by a single customer experiencing Poisson demand. It has been considered by many other authors, including Liu and Shi (1999), Schultz and Johansen (1999), Strijbosch, Heuts, and van der Schoot (2000) and those mentioned before. The Erlang demand process is also a building block in analysing multi-echelon systems (Andersson, Axsäter, & Marklund, 1998; Axsäter, 2000; Berling & Marklund, 2006, 2013; Deuermeyer & Schwarz, 1981; Lee & Moinzadeh, 1987; Moinzadeh & Lee, 1986; Svoronos & Zipkin, 1988).

We will also consider the much more restrictive policy with a constant delay time, independent of the order quantity. This policy was also studied by Katircioglu (1996) and Moinzadeh and Zhou (2008) for systems with unit-sized demands, and indeed shown to be optimal for those systems. This is clearly not the case for compound demand processes, but the optimal policy of this type will be easier to implement (in non-computerised systems) and can serve as a benchmark for the performance of the flexible delay policy.

The remainder of this paper is organised as follows. In Section 2, we introduce notations and present the inventory system and policy in detail. We derive the general optimality conditions for determining the maximum delays in the flexible delay policies and subsequently we do the same for policies with a constant maximum delay time. The exact form of the optimality conditions for both types of policies is then provided assuming Erlang distributed customer inter-arrival times. In Section 3, we numerically study the effect of the order quantity on the maximum delay. Furthermore, we compare the costs of

Download English Version:

https://daneshyari.com/en/article/6896178

Download Persian Version:

https://daneshyari.com/article/6896178

Daneshyari.com