Available online at www.sciencedirect.com
ScienceDirect
Int. J. Production Economics 112 (2008) 869–884
international journal of production economics
www.elsevier.com/locate/ijpe

Service performance of two-echelon supply chains under linear rationing

A.G. Lagodimos[a,*], S. Koukoumialos[b]

[a]Department of Business Administration, University of Piraeus, 80 Karaoli and Dimitriou Street, GR-185 34 Piraeus, Greece
[b]Department of Business Administration, University of the Aegean, 8 Michalon Street, GR-82 100 Chios, Greece

Received 10 October 2006; accepted 12 July 2007
Available online 28 August 2007

Abstract

Consideration is given to the operation of simple two-echelon divergent supply chains, where one central stock-point feeds several end stock-points. Periodic review echelon order-up-to policies are used to control the chain. Customer demand is imposed at end stock-points and, if unsatisfied, is backordered. In the event of central stock-point shortages, available material has to be rationed. We study a class of practical rationing rules, known as linear rationing (LR), which distributes shortage to end stock-points according to fixed fractions. After developing an alternative definition for the LR class rationing function, we formally show that popular rules, such as fair share (FS) and consistent appropriate share (CAS), belong to the LR class. We also propose a new LR rule that directly extends FS scope and applicability. For normal demand processes, we then develop closed-form customer service models, applicable to any LR rule. Modeling accuracy critically depends on the satisfaction of the standard balance assumption. Using a distribution-free model for a surrogate measure of the balance probability, we discuss the expected balance behavior of LR rules. Monte Carlo simulations reveal the balance assumption to be robust, at least for normally distributed demands, under fairly general conditions. Finally, numerical comparisons of the new LR rule with an existing one show encouraging results for the new rule.

Keywords: Multi-echelon; Stochastic; Inventory; Service level; Base stock

1. Introduction

For nearly three decades, multi-echelon supply chains have constituted a focal research area. As a result, models for the control of supply chain of several forms and operating disciplines are now available. Due to the shear volume and variety of these models, surveys of varying scope or focus often appear (Inderfurth, 1994; Van Houtum et al., 1996; Diks et al., 1996; de Kok and Fransoo, 2003; Mula et al., 2006). In this paper, we revisit a very simple two-echelon chain, where one central stock-point feeds several end stock-points, aiming at developing closed form service performance models. The chain is controlled by a periodic review of echelon order-up-to policies and unsatisfied external demand is backordered. In the event of material shortages at the central stock-point, all available material is rationed to partially satisfy end stock-points requisitions. We study a general class of rationing rules, known as *linear rationing* (LR).

*Corresponding author. Tel.: +30 210 4142 197.
E-mail address: alagod@unipi.gr (A.G. Lagodimos).

0925-5273/$ - see front matter
doi:10.1016/j.ijpe.2007.07.007

Since rationing is central in this paper, some clarifications are needed.

Rationing may be viewed as a special case of the so-called allocation problem, where a feeding stock-point distributes material to several successors. For multi-period ordering cycles, allocation decisions need not allocate all available material, so allocation quantities (and timing) may serve as control variables (Jonsson and Silver, 1987; McGavin et al., 1993; Cao and Silver, 2005). In contrast, rationing allocates all available material and existing rationing rules use the following simplified logic. Considering the feeding stock-point echelon net inventory (i.e. on hand plus echelon inventory position of all successors) all available for allocation, inventory position targets for all successors *after rationing* are determined. But, since previous decisions have already committed part of this inventory to specific successors, targets need not be feasible. Hence, the so-called balanced inventories (balance) assumption is introduced, stating that targets are always achieved. Based now on the targets determination logic, rationing rules may be classified into: dynamic and fixed (or practical). Dynamic rules view targets setting as a (usually cost) local optimization problem, whose solution depends on the feeding stock-point echelon net inventory (Federgruen, 1993; van Houtum et al., 1996; Axsater, 2003). In contrast, for fixed rules targets are set using a predetermined rationing function, independent of the feeding stock-point inventory. We now present some of the major results on practical rationing rules, where this paper is focused.

The best known practical rule, *fair share* (FS) rationing, was proposed (together with the concepts of balance and echelon inventory now taken as standard) by Clark and Scarf (1960). As introduced, FS rationing aims at equalizing end stock-point's stock-out probability. Using FS and a discounted cost objective (with linear holding and backorders costs), they extended an optimal base stocks evaluation procedure from serial to chains with a divergent last echelon. Eppen and Schrage (1981) studied a two-echelon distribution chain where the central stock-point (depot) does not hold stock. Considering FS, they modeled system inventories and average cost under different control disciplines. For base stock control and the cost structure of Clark and Scarf, FS is the optimal rationing policy. A basic assumption of this study is that retailers face identical normal demands and have identical lead times. By relaxing this assumption, Bollapragada et al. (1999) repeated the analysis and showed that all respective results still apply. Two other studies also considered two-echelon chains under base stock with FS, but allowed the depot to hold stock. van Donselaar and Wijngaard (1986) gave an exact analytical model for the stock-out probability for the case of identical retailers. Lagodimos (1992) allowed for non-identical retailers and proposed exact models and approximations for three standard service measures.

A key property of FS is backorders minimization under normal demands (Jonsson and Silver, 1987). However, by imposing identical stock-out probability at all end stock-points, FS also limits system controllability. This motivated de Kok (1990) to propose *consistent appropriate share* (CAS) rationing. CAS rationing aims at keeping the projected net inventory (i.e. inventory position minus expected lead-time demand) fraction of each end stock-point over system-wide projected net inventory fixed. By jointly determining rationing fractions and respective order-up-to levels, a chain may achieve any individual end stock-point's service objectives. In essence, CAS is a generalization of FS whose rationing fractions are effectively fixed by the lead-time demand at end stock-points (de Kok et al., 1994). In his original work, de Kok studied CAS for two-echelon chains with a stockless depot and gave exact and heuristic algorithms to determine the system parameters to achieve specific fill-rate targets. De Kok et al. (1994) allowed the depot to hold stock and heuristically evaluated the rationing fractions in the context of a fill-rate constrained cost minimization problem. Verrijdt and de Kok (1996) refined these heuristics, while Verrijdt and de Kok (1995) extended the analysis to general divergent chains with stocks only at end stock-points. Shortcomings of CAS, however, limit its applicability. As shown by simulation (Verrijdt and de Kok, 1996; van der Heijden et al., 1997), especially for low fill-rate targets, CAS may cause imbalances, resulting in deviations from required performance. Further, computational problems may prevail due to negative projected net inventory values (see Diks and de Kok, 1996, for a modification of CAS to deal with this problem).

An important contribution in the development of practical rationing rules is the LR class of rules by van der Heijden (1997). Instead of rationing the projected net inventory (as FS and CAS), LR rules ration system-wide shortage. As a result, rationing aims at maintaining the fraction of the shortage at a particular end stock-point to the system-wide shortage at a fixed value. Due to the number of variables involved, with LR rules the shortage rationing fractions and order-up-to levels need not be jointly determined (as for CAS) to obtain a required performance. Hence, provided that rationing fractions are externally given, calculations are much simplified. Van der Heijden proposed determining the rationing fractions so as to minimize a measure of average imbalance, thus introducing the *balanced stock* (BS) rule. An approximate Lagrangean computation procedure was given, strictly valid for two-echelon chains with a stockless depot under normal demand. Heuristic closed-form solutions for the BS rule rationing fractions determination were later proposed by van der Heijden et al. (1997), while de Kok and Fransoo (2003) established an exact solution. In addition to its excellent balance performance (van der Heijden et al., 1997), the BS rule can be easily adapted to control any divergent multi-echelon chain. Considering such systems and using this BS feature, van der Heijden (2000) proposed a heuristic to determine respective order-up-to levels so as to minimize holding costs while satisfying specific fill-rate targets at end stock-points.

We close this review with the *priority rationing* (PR) rule by Lagodimos (1992). Using a priority list, PR calls for the complete satisfaction of all successor stock-point orders in the sequence listed until available material is exhausted. For two-echelon chains using PR, Lagodimos provided analytical service models valid under an assumption equivalent to balance. Although easy to use and to extend to general divergent chains, few results for PR have been reported (see Lagodimos and Anderson, 1993, for safety stock positioning and Zhang, 2003, for a single-period model application).

We now make a general remark that effectively motivated the present paper. In the research presented, modeling assumptions and approaches vary according to the rationing rule considered. Studies of FS and PR assume normal demands, while those of CAS and BS Erlang or γ-distributed demands. This directly affects the models developed. Although the studies of FS and PR usually arrive at detailed analytical models, those for CAS and BS are general, requiring either numerical integration or special approximation techniques (as the approximation of the fill rate by a gamma-distribution in most studies by de Kok and co-workers). Thus, despite the statement by Diks and de Kok (1999) that FS and CAS belong to the LR class, no unified results are available.

This paper aims to bridge somewhat the gap between these approaches. Using an additional set of parameters, we give an alternative definition of the LR class rationing function. This allows a unified treatment of any rule in the LR class, the formal identification of FS and CAS with this class and the natural extension of FS through a new rule that overcomes its major limitation. It also forms the basis for developing analytical models for both the service performance and the balance probability under normal demand. The models (given in closed form, thus eliminating the need for approximations) are applicable for any LR rule and can be used for the effective control of two-echelon supply chains. Finally, comparative results for the proposed new rule with BS rationing (effectively, the first BS study under normal demand processes) are provided and their implications on rationing rule selection discussed.

2. Modeling preliminaries

In this section, we introduce the notation and the operating assumptions of the supply chain studied. We also present the modeling approach used, giving the general expressions underlying the system dynamic operation.

2.1. Notation

Other than some non-dimensional ratios presented in Section 4.1, the following notation is used throughout this paper:

$d_i(t)$	end stock-point i demand at t
$D_i(t_1,t_2)$	end stock-point i demand in $[t_1, t_2]$; $D_i(t_1,t_2)=\sum_{k=t_2}^{t_2} d_i(k)$
$D_0(t_1,t_2)$	system-wide demand in $[t_1, t_2]$; $D_0(t_1,t_2)=\sum_{j=1}^{N} D_j(t_1,t_2)$
$I_1^i(t)$	net inventory of end stock-point i after demand realization at t; $I_1^i(t)=U_1^i(t)-d_i(t)$
$J_1^i(t)$	inventory position of end stock-point i after ordering at t
$J_1(t)$	the quantity $J_1(t)=\sum_{j=1}^{N} J_1^j(t)$
L	lead time of central stock-point
l_i	lead time of end stock-point i
N	total number of end stock-points

$U_1^i(t)$	net inventory of end stock-point i before demand realization at t
$U_0(t)$	net inventory of central stock-point before demand realization at t
S_1^i	order-up-to-level for end stock-point i
S_1	the quantity $S_1 = \sum_{j=1}^{N} S_1^j$
S_0	echelon order up to level for central stock-point
$\{a_j\}_{j=1}^N$	set of rationing factors of an LR rule
$\{f_j\}_{j=1}^N$	set of rationing fractions of an LR rule
Δ	the quantity $\Delta = S_0 - S_1$
α_i, β_i, γ_i	measures of customer service for end stock-point i
μ_i, σ_i	mean and standard deviation of end stock-point i period demand

2.2. Operating assumptions

The supply chain studied is depicted in Fig. 1. It can be interpreted either as a manufacturing chain, where several end products are produced from a common part or as a distribution chain where a depot serves several retailers. Irrespectively, the chain is controlled by periodic review echelon order-up-to policies, with a common review period of one time unit. We use the following assumptions (standard in most research in the area):

1. Period demands at any end stock-point i form a stream of stationary independent random variables, realized at the end of each period.
2. Demand at any end stock-point i not satisfied from stock is backordered.
3. Central stock-point orders (to external suppliers) are always fully satisfied, while those of end stock-points (to the central stock-point) may be partially satisfied according to availability.
4. There are no capacity constraints and lead times are fixed.
5. No emergency transshipments between end stock-points take place.

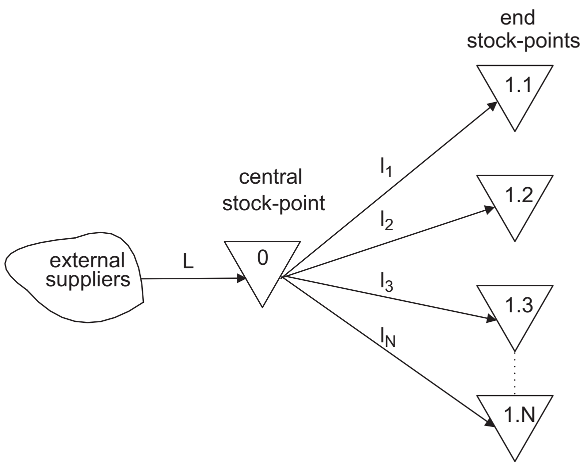

Fig. 1. Schematic representation of the two-echelon supply chain.

We also need to clarify the sequence of events within any time period. (1) All stock-points receive the replenishment orders placed respective lead-time periods earlier. (2) All stock-points review their inventory status and place replenishment orders. (3) The central stock-point ships the supplies (just ordered) to end stock-points (either the quantity requested or an appropriate ration in case of shortage). (4) Demand at the end stock-points is realized.

2.3. Modeling approach

Since the service measures we consider are functions of specific inventory variables, we need to determine their stochastic behavior over time. Following the standard approach (Lagodimos, 1992; Verrijdt and de Kok, 1996) for two-echelon systems, we study the system dynamics in an arbitrary time interval $[t-L,\ t+l_i]$. Since at the beginning of period $t-L$ the central stock-point echelon inventory position is raised to S_0, we can determine (at period t) its echelon net inventory as well as the sum of all end stock-point's inventory positions:

$$U_0(t) = S_0 - D_0(t-L, t-1) \text{ and}$$
$$J_1(t) = \min[U_0(t), S_1]. \quad (1)$$

While $J_1(t)$ as above represents the aggregate end echelon inventory exactly, we need the inventory position of individual end stock-points, which depends on the rationing rule. Considering any fixed rationing rule, with a rationing function q_i for end stock-point i, we can disaggregate to obtain

$$J_1^i(t) = q_i[J_1(t)], \quad \text{where}$$
$$\sum_{j=1}^{N} q_j[J_1(t)] = J_1(t). \quad (2)$$

Under the balance assumption, the inventory position of any end stock-point i given by (2) can always be achieved and the system is effectively decomposed to N (virtually independent) serial two-echelon chains. So we can directly determine the corresponding end stock-point net inventory (at the start and end) of period $t+l_i$

$$U_1^i(t+l_i) = J_1^i(t) - D(t, t+l_i-1) \text{ and}$$
$$I_1^i(t+l_i) = J_1^i(t) - D_i(t, t+l_i). \tag{3}$$

Note that the above analysis applies for any demand distribution. Provided a specific distribution is given and a particular rationing function specified, the distribution of all inventory variables can be determined and used for the development of performance models. It also covers chains where the central stock-point may either hold stock or not. In fact, as defined (see Section 2.1), $S_0 = \Delta + S_1$. Observe that, if $\Delta = 0$, then from (1) we always have that $J_1(t) = U_0(t)$ and the chain degenerates to one with a stockless central stock-point (see Diks et al., 1996, for a discussion).

3. Linear rationing class

We now give an alternative form for the LR class rationing function that allows comparison with existing fixed rationing rules. We also discuss the relation of these rules with the LR class and present a new rule.

3.1. The rationing function

As originally defined (see van der Heijden, 1997), the LR class rationing decisions aim at distributing total central stock-point shortage to individual end stock-points. Specifically, after rationing at some period t, the respective inventory position of any end stock-point i is given by

$$J_1^i(t) = S_1^i - f_i \max[0, D_0(t-L, t-1) - \Delta],$$
$$\text{where} \quad \sum_{j=1}^{N} f_j = 1. \tag{4}$$

The second term above represents system-wide shortage and each f_i the fixed fraction of the shortage allocated to end stock-point i. Considering that other fixed rationing rules are defined in terms of the available end echelon inventory (precisely the projected net inventory) and not the shortage, comparisons are not directly possible. The following observation helps overcome this problem.

Proposition 3.1. *The rationing function of any rule in the LR class, for any end stock-point i, can be equivalently represented as*

$$q_i[J_1(t)] = f_i\left[J_1(t) - \sum_{j=1}^{N}(l_j+1)\mu_j\right] + (l_i+1)\mu_i - a_i,$$
$$\text{where} \quad \sum_{j=1}^{N} a_j = 0.$$

Moreover, any rationing rule with a rationing function as above belongs to the LR class.

Proof. See Appendix A. □

In contrast to the original definition, the above clearly distinguishes the rationing function of the LR class from system order-up-to levels. For this purpose, an additional set of parameters (refered hereafter as *rationing factors*) is introduced that, together with the rationing fractions, define any specific rule. Thus, there are $2N$ parameters associated with any rule in the LR class: N rationing fractions $\{f_j\}_{j=1}^N$ and N rationing factors $\{a_j\}_{j=1}^N$. Provided that the values of all these parameters are determined or a specific evaluation method decided, a specific rule of the class is fully defined.

We also note that the rationing function of any LR rule is now clearly associated with the projected net inventory notion (see de Kok et al., 1994, and Diks et al., 1996, for a comprehensive discussion of this inventory variable). To see this, notice that (for any stock-point i and period t) the corresponding rationing fraction is given by

$$f_i = \frac{J_1^i(t) - (l_i+1)\mu_i + a_i}{J_1(t) - \sum_{j=1}^{N}(l_j+1)\mu_j}, \tag{5}$$

where the denominator is the system-wide projected net inventory and the nominator is the projected net inventory of stock-point i (modified by a_i). Thus, this form of the LR class rationing function allows direct comparisons with other existing fixed rationing rules.

3.2. Existing rules

We now examine existing rationing rules and discuss their relation with the LR class in terms of respective rationing fractions $\{f_j\}_{j=1}^N$ and factors $\{a_j\}_{j=1}^N$. For the rules not usually considered as members of the class, we simply reduce their rationing function as known in its standard form (see Diks et al., 1996) to that in Proposition 3.1.

Details with respect to these rules and associated research are given in Section 1.

3.2.1. The FS rule

This rule is a member of the LR class, with parameters:

$$f_i = \frac{\sigma_i\sqrt{l_i+1}}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}} \quad \text{and} \quad a_i = 0 \quad \text{for all } i = 1,\ldots,N.$$

Both sets of parameters are fixed to give identical stock-out probabilities at all end stock-points under normal demand (see Section 1), values that minimize expected backorders. As a result, FS is very specialized and cannot be used when this is not the desired performance target.

3.2.2. The CAS rule

This rule is also a member of the LR class, with rationing factors arbitrarily set to

$$a_i = 0 \quad \text{for all } i = 1,\ldots,N.$$

The rationing fractions $\{f_j\}_{j=1}^N$, on the other hand, are left free to be evaluated according to the desired system service performance. Clearly, CAS is a generalization of FS and reduces to FS if an identical end stock-point stock-out probability target is desired (also under normal demand). It is the arbitrary setting of the CAS rationing factors which explains both the negative rationing fractions and the unsatisfactory balance performance reported for this rule (see Sections 1 and 5).

3.2.3. The BS rule

This rule is a known member of the LR class, with rationing fractions fixed to

$$f_i = \frac{\mu_i^2}{2\sum_{j=1}^{N}\mu_j^2} + \frac{\sigma_i^2}{2\sum_{j=1}^{N}\sigma_j^2} \quad \text{for all } i = 1,\ldots,N.$$

The rationing factors $\{a_j\}_{j=1}^N$ are left free to be evaluated according to the desired system service performance. The above expression for $\{f_j\}_{j=1}^N$ was proposed by de Kok and Fransoo (2003) as an exact expression minimizing the average imbalance in two-echelon chains with a stockless depot (see van der Heijden, 1997, for a procedure and van der Heijden et al., 1997, for heuristic $\{f_j\}_{j=1}^N$ expressions).

3.3. A new rule

Following the above analysis, a new LR rule, referred as *augmented fair share* (AFS) rationing, emerges. In essence, the AFS rule forms a natural extension of FS, overcoming its major limitation. We now define AFS in terms of its rationing function parameters. Specifically, the AFS rationing fractions are fixed (as for the FS rule) to

$$f_i = \frac{\sigma_i\sqrt{l_i+1}}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}} \quad \text{for all } i = 1,\ldots,N.$$

Respective rationing factors $\{a_j\}_{j=1}^N$ are left free to be evaluated according to the desired system service performance.

As defined, AFS follows the BS rule logic (i.e. fixed rationing fraction with free rationing factors), a fact that facilitates respective parameters evaluation. More important, however, is the inherent property of AFS to reduce to FS when identical end stock-point's stock-out probabilities is the performance target (under normal demands). In fact, it can easily be shown (by contradiction) that, in this case, the AFS rationing factors become $a_i = 0$ for all i. Thus, AFS not only allows full chain control but also carries the optimality properties of FS under the conditions the latter is defined. As a result, it can replace the latter in all applications.

4. Service performance

Using the previous results, we develop customer service performance models for the chain under study assuming normal demands. The models are then given in closed form, allowing their evaluation using standard software or handbooks.

4.1. Inventory distributions

We need first to determine the probability distribution of the net inventory variables $U_1^i(t+l_i)$ and $I_1^i(t+l_i)$. Using (3) and (5), we directly obtain

$$I_1^i(t+l_i) = f_i\left\{\min[S_1+\mathbf{\Delta}-D_0(t-L,t-1),S_1] - \sum_{j=1}^{N}(l_j+1)\mu_j\right\} + (l_i+1)\mu_i - a_i - D_i(t,t+l_i),$$

$$U_1^i(t+l_i) = f_i\left\{\min[S_1+\mathbf{\Delta}-D_0(t-L,t-1),S_1] - \sum_{j=1}^{N}(l_j+1)\mu_j\right\} + (l_i+1)\mu_i - a_i - D_i(t,t+l_i-1). \quad (6)$$

The random variables in each pair $[D_0(t-L,t-1), D_i(t,t+l_i-1)]$ and $[D_0(t-L,t-1), D_i(t,t+l_i)]$, representing demands for unrelated time periods, are mutually independent. Assuming now normal period demands uncorrelated in time, we have

$$D_0(t-L,t-1) \sim N\left[L\sum_{j=1}^{N}\mu_i, L\sum_{j=1}^{N}\sigma_j^2\right],$$
$$D_i(t,t+l_i-1) \sim N[l_i\mu_i, l_i\sigma_i^2]$$
$$\text{and } D_i(t,t+l_i) \sim N[(l_i+1)\mu_i, (l_i+1)\sigma_i^2].$$

It is convenient to use non-dimensional ratios in the analysis. These reduce the variables involved and give the models in a general form independent of specific chain parameters (see Lagodimos, 1992, 1993 for a discussion and physical interpretation). Thus, for each end stock-point i, we use the ratios

$$Z_i = \frac{S_1^i - (l_i+1)\mu_i}{\sigma_i\sqrt{l_i+1}},\quad M_i = \frac{\mu_i}{\sigma_i\sqrt{l_i+1}},\quad V_i = \frac{\sqrt{l_i}}{\sqrt{l_i+1}},$$
$$T_i = \frac{a_i}{\sigma_i\sqrt{l_i+1}},\quad W_i = \frac{\sigma_i\sqrt{l_i+1}}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}},$$
$$\text{and}\quad r_i = f_i/W_i. \tag{7a}$$

Note that r_i is a transformed version of the allocation fraction f_i and that for the FS or AFS rule we have $r_i = 1$ for all i. We also use the following ratios representing the overall system setting:

$$Z = \frac{S_1 - \sum_{j=1}^{N}(l_j+1)\mu_j}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}},\quad Z_\Delta = \frac{\mathbf{\Delta} - L\sum_{j=1}^{N}\mu_i}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}},$$
$$V = \frac{\sqrt{L\sum_{j=1}^{N}\sigma_j^2}}{\sum_{j=1}^{N}\sigma_j\sqrt{l_j+1}}. \tag{7b}$$

It is easy to show that, from the definitions of Z and Z_i and (5), the following relations always hold:

$$Z = \sum_{j=1}^{N} W_j Z_j \text{ and } Z_i = Zr_i - T_i \quad \text{for all } i = 1,\ldots,N. \tag{8}$$

Introducing the ratios from (7) in expressions (6) and standardizing the normal variables involved, after some algebra, we obtain

$$I_1^i = \sigma\sqrt{l_i+1}\min(Z_i - x, Z_i + Z_\mathbf{\Delta}r_i - Vr_iy - x),$$
$$U_1^i = \sigma_i\sqrt{l_i}\min\left(\frac{Z_i+M_i}{V_i} - z, \frac{Z_i+M_i+Z_\mathbf{\Delta}r_i}{V_i} - \frac{V}{V_i}r_iy - z\right), \tag{9}$$

where x, z and y are normal N(0,1). Since both inventory variables are stationary, respective time indices above (and from now on) are omitted. Note that (9) are very similar to the respective net inventory expressions for serial two-echelon chains (see Lagodimos, 1993; Lagodimos et al., 1995), forming the basis for the analysis that follows.

4.2. Service models

We now present closed-form models for three customer service measures (α, β and γ according to the typology by Schneider, 1981) applicable to any LR rule. Models are presented with reference to any arbitrary end stock-point i.

4.2.1. The α measure

This measure represents the fraction of periods where no stock-out occurs and so demand can be directly satisfied on request, constituting the discrete-time version of the *ready rate* (P_3 measure in Silver et al., 1998). Note that α evaluation forms an integral part of solution algorithms for the cost-optimal control of supply chains (Diks and de Kok, 1999; van Houtum et al., 1996). In general

$$\alpha_i = Pr(I_1^i \geqslant 0).$$

Introducing I_1^i from (9) in the above, we obtain

$$\alpha_i = Pr\{x \leqslant Z_i \text{ and } x + Vr_iy \leqslant Z_i + Z_\mathbf{\Delta}r_i\}. \tag{10}$$

The above represents an integral over the circular bivariate normal probability plan (x, y; 0) and cannot be analytically evaluated, requiring either numerical integration or approximations (see Lagodimos, 1993 for an approximation). However, a closed-form expression may be obtained using the simple transformation in Lagodimos et al. (1995). Define a new variable

$$w = (x + Vr_iy)/\sqrt{A},\quad \text{where } A = (Vr_i)^2 + 1. \tag{11}$$

It can be directly shown that $w \sim N(0,1)$ and that x with w are now correlated and form a standardized bivariate normal variable $(x, w; 1/\sqrt{A})$. Direct substitution in (10) gives

$$\alpha_i = Pr\{x \le Z_i \text{ and } w \le (Z_i + Z_\mathbf{\Delta}r_i)\big/\sqrt{A}\}$$
$$\alpha_i = \Psi\left(Z_i, \frac{Z_i + Z_\mathbf{\Delta}r_i}{\sqrt{A}}; \frac{1}{\sqrt{A}}\right), \tag{12}$$

where $\Psi(.,.;\rho)$ is the standardized bivariate normal probability function with correlation coefficient ρ. This may be found in standard handbooks (Abramowitz and Stegun, 1965) and is also available in mathematical software packages.

4.2.2. The γ measure

This measure expresses the average backorders as a fraction of the average period demand. For not seriously under-stocked chains, this is a good approximation for the fill rate (see below). Since it directly relates with backorders cost, γ indirectly enters cost-optimization algorithms (Diks and de Kok, 1998; Axsater, 2003). In general

$$\gamma_i = 1 - \frac{E\{-I_1^{i-}\}}{\mu_i},$$

where $b^- = \min(b, 0)$. Clearly, $E[-I_1^{i-}]$ represents average backorders of stock-point i. Introducing I_1^i from (9) and the ratios from (7) in the above, we finally obtain

$$(1-\gamma_i)M_i = E\{x - Z_i | \mathscr{A}\} Pr\{\mathscr{A}_1\} + E\{x + Vr_i y - Z_i - Z_\Delta r_i | \mathscr{A}_2\} Pr\{\mathscr{A}_2\},$$

where $\mathscr{A}_1$ and $\mathscr{A}_2$ are the events:

$$\mathscr{A}_1 = \{x > Z_i \text{ and } Vy < Z_\Delta\}$$

and

$$\mathscr{A}_2 = \{x + Vr_i y > Z_i + Z_\Delta r_i \text{ and } Vy \geqslant Z_\Delta\}.$$

An expression identical to the above (with different random variables coefficients) corresponds to two-echelon serial systems (see Lagodimos, 1993) for which a closed-form model was presented. Thus, by maintaining analogies between the respective variables we can write

$$\begin{aligned}(1-\gamma_i)M_i = {} & [1-\Phi(Z_i)]\left[Vr_i\varphi\left(\frac{Z_\Delta}{V}\right) + Z_\Delta r_i \Phi\left(\frac{Z_\Delta}{V}\right)\right] \\ & + \Phi\left(\frac{Z_\Delta}{V}\right)\varphi(Z_i) \\ & + \sqrt{A}\varphi\left(\frac{Z_i + Z_\Delta r_i}{\sqrt{A}}\right) \\ & \times \left[1 - \Phi\left(\frac{Z_\Delta r_i + Z_i(1-A)}{Vr_i\sqrt{A}}\right)\right] \\ & - (Z_i + Z_\Delta r_i)(1-\alpha_i), \qquad (13)\end{aligned}$$

where φ(.) and Φ(.) are the univariate standardized normal probability density and function, respectively, A is given by (11) and α_i is the service measure in (12). Hence, γ_i is evaluated.

4.2.3. The β measure (fill rate)

This measure represents the fraction of average demand satisfied immediately from stock (corresponds to the P_2 measure in Silver et al., 1998) and is the service measure mostly used in multi-echelon supply chains research (see references in Section 1). In general,

$$\beta_i = 1 - \frac{E\{-I_1^{i-}\} - E\{-U_1^{i-}\}}{\mu_i}.$$

The numerator simply represents the additional average backorders occurring at the last period of a cycle. Clearly, the above is equivalent to

$$(1-\beta_i)M_i = (1-\gamma_i)M_i - \frac{E\{-U_1^{i-}\}}{\sigma_i\sqrt{l_i+1}}, \qquad (14)$$

so we only need an expression for $E\{-U_1^{i-}\}$. Using (9), we can write

$$\begin{aligned}\frac{E\{-U_1^{i-}\}}{\sigma_i\sqrt{l_i+1}} = {} & V_i E\left\{\frac{Z_i + M_i}{V_i}\middle| \mathscr{B}_1\right\} Pr\{\mathscr{B}_1\} \\ & + V_i E\left\{z + \frac{Vr_i}{V_i}y - \frac{Z_i + M_i + Z_\Delta r_i}{V_i}\middle| \mathscr{B}_2\right\} \\ & \times Pr\{\mathscr{B}_2\},\end{aligned}$$

where $\mathscr{B}_1$ and $\mathscr{B}_2$ are the events:

$$\mathscr{B}_1 = \left\{z > \frac{Z_i + M_i}{V_i} \text{ and } Vy < Z_\Delta\right\}$$

$$\mathscr{B}_2 = \left\{z + \frac{Vr_i}{V_i}y > \frac{Z_i + M_i + Z_\Delta r_i}{V_i} \text{ and } Vy \geqslant Z_\Delta\right\}.$$

As expected, this is identical in form to the respective expression for $(1-\gamma)M_i$ and can be analyzed in the same manner. Leaving aside the algebra involved, we finally obtain

$$\begin{aligned}\frac{E\{-U_1^{i-}\}}{\sigma_i\sqrt{l_i+1}} = {} & \left[1 - \Phi\left(\frac{Z_i + M_i}{V_i}\right)\right] \\ & \times \left[Vr_i\varphi\left(\frac{Z_\Delta}{V}\right) + Z_\Delta r_i \Phi\left(\frac{Z_\Delta}{V}\right)\right] \\ & + V_i\Phi\left(\frac{Z_\Delta}{V}\right)\varphi\left(\frac{Z_i + M_i}{V_i}\right) \\ & + V_i\sqrt{B}\varphi\left(\frac{Z_i + M_i + Z_\Delta r_i}{V_i\sqrt{B}}\right) \\ & \times \left[1 - \Phi\left(\frac{Z_\Delta r_i + (Z_i + M_i)(1-B)}{Vr_i\sqrt{B}}\right)\right] \\ & - (Z_i + M_i + Z_\Delta r_i) \\ & \times \left[1 - \Psi\left(\frac{Z_i + M_i}{V_i}, \frac{Z_i + M_i + Z_\Delta r_i}{V_i\sqrt{B}}; \frac{1}{\sqrt{B}}\right)\right], \qquad (15)\end{aligned}$$

where $B = (Vr_i/V_i)^2 + 1$.

Thus, replacing the above in (14), β_i can be directly evaluated. Observe that, in general, $\beta_i \leqslant \gamma_i$; thus, γ_i overestimates (somewhat) the fill rate when used to approximate the latter.

5. The balance assumption

We now present an exact model of a surrogate measure for the satisfaction of the balance assumption, applicable for any LR rule, and explore its behavior under various settings. We also discuss the expected balance performance of specific LR rules.

5.1. Probability model

The accuracy of the proposed service models critically depends on the satisfaction of the balance assumption. Therefore, we need a measure of the balance probability, P_B say, at an arbitrary period. While it is generally hard to model P_B directly (see Zipkin, 1984, for a theoretical treatment of balance), Eppen and Schrage (1981) suggested a surrogate measure $\tilde{P}_B$; the probability of balance at some period given balance at the previous period. This measure was subsequently modeled under various operating assumptions and rationing rules (e.g. Eppen and Schrage, 1981, stockless central stock-point with FS, Verrijdt and de Kok, 1996, the same with CAS, Lagodimos, 1992, allowing central stock with FS).

Using an approach analogous to Eppen and Schrage and Lagodimos, we now present a distribution-free model for $\tilde{P}_B$, which helps understanding the effects of system parameters and allows direct estimation of this surrogate measure of balance.

Proposition 5.1. *The surrogate balance probability for any rule of the LR class is given by*

$$\begin{aligned}\tilde{P}_B = Pr\{&\min_{1\leqslant j\leqslant N}[\frac{d_j(t-1)}{f_j}]\geqslant \min[D_0(t-1,t-1).\\&-D_0(t-L-1,t-L-1),\ \ .D_0(t-L,t-1)-\mathbf{\Delta}]\\&\textit{and } D_0(t-L,t-1)>\mathbf{\Delta}\}\\&+Pr\{D_0(t-L,t-1)\leqslant\mathbf{\Delta}\}.\end{aligned}$$

Proof. See Appendix A. □

On the basis of the above, two factors directly related with system control may affect the balance: Δ and the rationing function. Δ regulates central stock availability, which is minimized for $\Delta = 0$, where the chain degenerates to one with a stockless central stock-point (see Section 2.3). In this case, the second term in $\tilde{P}_B$ vanishes, so balance depends on the first term only. This is when balance is most critical. As Δ increases, the second term becomes dominant and the impact of balance diminishes.

Turning to the rationing function, observe that only the rationing fractions $\{f_j\}_{j=1}^N$ affect balance. Therefore, LR rules with fixed rationing fractions (i.e. BS, FS and AFS) are expected to have a uniform balance performance independent of any service targets (as demonstrated in the simulations by van der Heijden et al., 1997, for the BS rule). In fact, AFS and FS (having identical rationing fractions) are expected to perform identically. Moreover, rules with variable rationing fractions (such as CAS), determined on the basis of specific service targets, will manifest a service-related balance performance. Due to the extreme values that the rationing fractions may need to obtain, balance may also be poor (see $\tilde{P}_B$ estimates and simulations in Verrijdt and de Kok, 1995, 1996, and simulations in van der Heijden et al., 1997).

5.2. Numerical investigation

To assess the impact of both the environmental conditions (system setting) and the rationing rule on balance, a numerical investigation was performed using Monte Carlo simulation. We studied a system with N end stock-points (without loss of generality, N is even), divided into two equally sized groups, $\mathcal{GA}$ and $\mathcal{GB}$, each having end stock-points with identically distributed period demands, where:

$$\mathcal{GA} = \left\{i : \mu_i = \mu_A \text{ and } \sigma_i = \sigma_A \text{ for } i \leqslant \frac{N}{2}\right\}$$

$$\mathcal{GB} = \left\{i : \mu_i = \mu_B \text{ and } \sigma_i = \sigma_B \text{ for } i \leqslant \frac{N}{2}\right\}.$$

In order to capture the rationing rule effect, we used the ratio f_A/f_B, where f_A and f_B are the rationing fractions for end stock-points in $\mathcal{GA}$ and $\mathcal{GB}$, respectively. Clearly, from (4), it always holds that $(f_A+f_B)N/2 = 1$, so any f_A/f_B value also determines f_A and f_B for given N. Therefore, by varying f_A/f_B, the effects of the rationing rule on balance may be understood.

As shown in Appendix B, for the particular investigation setting described above, the following non-dimensional factors generally influence $\tilde{P}_B$:

$$L, N, f_A, f_B,\ \ \theta = \frac{\mathbf{\Delta}}{E\{D_0\}},$$

$$\mathrm{CV}_A = \frac{\sigma_A}{\mu_A},\ \ \mathrm{CV}_B = \frac{\sigma_B}{\mu_B},\ \text{ and }\ b = \frac{\mu_B}{\mu_A}.$$

Note that CV_A and CV_B represent the *variation coefficients* of the period demands in the respective group, factors usually considered in such studies. The factor θ represents the central stock-point stock availability expressed as a multiple of the respective mean lead-time demand, while b is the ratio of the period demand means for the two groups (we use the convention, $b \geqslant 1$).

Tables 1 and 2 give representative Monte Carlo results for various system settings. All were obtained using $\tilde{P}_{\mathrm{B}}$ in its non-dimensional form in Appendix B using run lengths of 10,000 periods. Specifically, Table 1 presents symmetric cases, where $\mathcal{GA}$ and $\mathcal{GB}$ have identical characteristics, while Table 2 presents non-symmetric cases, where variation coefficients and/or demand means vary between $\mathcal{GA}$ and $\mathcal{GB}$. Clearly, normal demand processes are strictly represented by variation coefficient values less than 0.33 (that ensure strictly non-negative period demands).

Considering first the impact of the environmental conditions, both tables reveal high balance probabilities. This is especially true for the symmetric cases, where even for $\mathrm{CV_A} = \mathrm{CV_B} = 0.5$ (that do not represent normal demands) $\tilde{P}_{\mathrm{B}}$ remains high. In non-symmetric cases, the $\tilde{P}_{\mathrm{B}}$ values (although high in absolute terms) are generally lower than their symmetric counterparts. This is more pronounced when the higher variation coefficient is associated with $\mathcal{GB}$ (group with the highest mean). In general, we can envisage non-symmetric cases (with high values of b) where $\tilde{P}_{\mathrm{B}}$ could become lower than the values in Table 2. As expected from other studies (Eppen and Schrage, 1981; Lagodimos, 1992; verrijdt and de Kok, 1996), $\tilde{P}_{\mathrm{B}}$ drops with increased number of end stock-points N and central stock-point lead-time L, while it drastically increases with increased θ (central inventory availability). These trends are clearly observed in both tables.

The effects of the rationing rule on balance are identified through different $f_{\mathrm{A}}/f_{\mathrm{B}}$ values. Leaving aside the differing behavior of $\tilde{P}_{\mathrm{B}}$ for $f_{\mathrm{A}}/f_{\mathrm{B}}$ in Tables 1 and 2 (see bellow), observe that (in all cases) the rationing rule becomes critical only for extreme (with respect to the observed maximum) values. Otherwise, the effect of this factor is practically small. Equally important, however, is the observed joint effects of $f_{\mathrm{A}}/f_{\mathrm{B}}$ with N, where the influence of $f_{\mathrm{A}}/f_{\mathrm{B}}$ drastically diminishes with increased N (in all cases). Thus, for high number of end stock-points N, the choice of the rationing rule used seems not to seriously affect balance. This, not previously reported result, can be explained by noting that, for increased N, the absolute difference of individual stock-points rationing fractions is

Table 1
Monte Carlo estimates for $\tilde{P}_{\mathrm{B}}$ with identical demand parameters for $\mathcal{GA}$ and $\mathcal{GB}$

θ	L	N	$\mathrm{CV_A} = \mathrm{CV_B} = 0.5$ and $b = 1$					$\mathrm{CV_A} = \mathrm{CV_B} = 0.33$ and $b = 1$				
			$f_{\mathrm{A}}/f_{\mathrm{B}}$					$f_{\mathrm{A}}/f_{\mathrm{B}}$				
			10/90	30/70	50/50	70/30	90/10	10/90	30/70	50/50	70/30	90/10
0	2	2	88.5	93.5	95.6	93.5	88.4	97.1	99.1	99.8	99.1	97.0
		6	81.7	85.7	87.2	84.9	80.8	97.8	99.0	99.3	98.8	97.8
		10	74.6	78.0	79.4	77.9	73.6	97.7	98.6	98.8	98.4	97.5
	6	2	87.9	93.5	95.4	93.5	88.7	97.1	99.4	99.7	99.2	97.4
		6	81.9	85.7	86.9	85.4	81.2	97.8	98.8	99.2	99.0	97.6
		10	74.5	78.1	79.1	78.7	74.4	97.8	98.4	98.7	98.3	97.7
	10	2	88.5	93.4	95.4	93.9	88.4	97.3	99.0	99.8	98.9	97.4
		6	81.5	85.9	86.8	85.3	82.2	97.8	98.7	99.3	99.0	97.6
		10	75.0	77.9	79.8	78.0	74.5	97.3	98.4	98.6	98.3	97.5
1	2	2	97.3	98.9	99.4	98.7	97.3	99.2	99.9	100.0	99.9	99.2
		6	94.3	95.4	96.0	95.3	94.2	99.3	99.7	99.7	99.7	99.4
		10	90.5	91.6	92.2	92.1	90.6	99.3	99.4	99.4	99.5	99.2
	6	2	96.0	98.1	98.8	98.1	96.5	99.8	99.9	100.0	99.9	99.8
		6	92.6	94.2	94.7	94.2	92.5	99.5	99.8	99.9	99.7	99.6
		10	89.2	90.6	90.7	90.2	89.2	99.5	99.6	99.5	99.6	99.4
	10	2	96.1	97.6	98.5	97.9	95.5	99.0	99.7	100.0	99.8	99.2
		6	92.2	93.7	94.5	94.0	91.8	99.0	99.6	99.8	99.6	99.3
		10	88.8	90.8	91.0	90.3	89.0	98.9	99.4	99.5	99.3	98.8

Table 2
Monte Carlo estimates for $\tilde{P}_{\mathrm{B}}$ with non-identical demand parameters for $\mathcal{GA}$ and $\mathcal{GB}$

θ	L	N	$CV_A = 0.33$, $CV_B = 0.25$ and $b = 1$					$CV_A = 0.33$, $CV_B = 0.25$ and $b = 2$					$CV_A = 0.25$, $CV_B = 0.33$ and $b = 2$				
			f_A/f_B					f_A/f_B					f_A/f_B				
			10/90	30/70	50/50	70/30	90/10	10/90	30/70	50/50	70/30	90/10	10/90	30/70	50/50	70/30	90/10
0	2	2	98.1	99.5	100.0	99.8	98.8	99.9	99.8	98.9	96.3	92.6	99.8	99.8	97.6	92.8	87.7
		6	99.0	99.5	99.7	99.5	98.8	99.6	99.7	99.1	97.5	95.2	99.6	99.7	99.1	97.0	93.5
		10	99.0	99.5	99.4	99.1	98.7	99.3	99.3	98.8	97.5	95.6	99.2	99.4	99.2	97.9	95.2
	6	2	98.2	99.5	99.9	99.8	98.8	99.9	99.8	99.1	96.3	91.9	99.8	99.8	97.6	93.1	86.8
		6	99.1	99.7	99.6	99.5	98.8	99.8	99.5	99.1	97.3	94.5	99.5	99.6	99.2	97.0	93.6
		10	99.1	99.4	99.4	99.2	98.8	99.5	99.3	98.8	97.7	95.6	99.2	99.6	99.2	97.7	95.2
	10	2	98.1	99.4	99.9	99.8	98.9	99.9	99.9	99.1	96.2	92.1	99.8	99.8	97.8	92.9	87.1
		6	99.2	99.7	99.8	99.4	98.8	99.8	99.5	99.0	97.7	94.9	99.6	99.7	99.2	96.7	93.0
		10	99.2	99.5	99.4	99.0	98.9	99.4	99.3	98.7	97.5	95.3	99.2	99.4	99.3	97.8	95.0
1	2	2	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.9	99.7	100.0	100.0	100.0	100.0	100.0
		6	99.9	99.9	99.9	99.9	99.9	100.0	100.0	99.9	99.7	99.3	100.0	100.0	100.0	100.0	99.9
		10	99.7	99.9	99.9	99.8	99.7	100.0	99.8	99.8	99.6	99.1	100.0	100.0	100.0	100.0	99.9
	6	2	99.5	99.9	100.0	100.0	99.7	100.0	100.0	99.6	98.8	96.7	100.0	100.0	99.9	99.6	99.2
		6	99.8	99.9	99.9	99.8	99.8	99.8	99.9	99.6	98.8	98.1	100.0	100.0	100.0	99.6	99.2
		10	99.8	99.9	99.8	99.8	99.6	99.8	99.7	99.5	99.1	98.1	99.9	99.9	99.8	99.7	99.2
	10	2	99.3	99.9	100.0	99.9	99.7	100.0	99.9	99.7	98.8	97.3	100.0	99.9	99.3	97.0	94.5
		6	99.7	99.9	99.9	99.7	99.7	99.9	99.9	99.7	99.0	97.7	99.8	99.9	99.8	98.9	97.3
		10	99.7	99.8	99.9	99.8	99.5	99.8	99.7	99.5	98.9	98.3	99.7	99.8	99.8	99.2	98.3

reduced, since $(f_A + f_B)N/2 = 1$. Hence, the rationing rule effect on balance diminishes accordingly.

We still need to explain the variation of $\tilde{P}_B$ with f_A/f_B. In the symmetric case (Table 1), $\tilde{P}_B$ behaves as a concave function, with a maximum at $f_A/f_B = 50/50$. This directly relates with system symmetry, where we intuitively expect $f_A = f_B = 1/N$ to be optimal with respect to $\tilde{P}_B$. For the non-symmetric case in Table 2, however, $\tilde{P}_B$ behaves as a decreasing function with a maximum for $f_A/f_B = 10/90$. Although we have not derived any analytical results, recall the BS rule whose rationing fractions minimize a related balance measure (see Section 3.2). Direct application of this rule gives f_A and f_B values very close to the observed optima in Table 2. For example, for $\{N = 2,\ CV_A = 0.33,\ CV_B = 0.25,\ b = 2\}$ we find the BS rationing fractions to be $f_A = 0.162$ and $f_B = 0.838$. Observe the maximum in the table that occurs for $f_A/f_B = 10/90 = 0.11$. Thus, provided a specific rationing rule does not impose f_A/f_B values at the extreme opposite direction, it will invariably result in acceptable $\tilde{P}_B$ results.

The implications of the above analysis on service models accuracy with respect to balance should be clear. Provided demand processes satisfy normality (under which the models were derived) and do not differ drastically between individual end stock-points, the models are expected to be accurate. This holds for any LR rule whose rationing fractions do not obtain extreme values. Thus, it generally applies to all LR rules, with the only possible exception that of CAS for chains with very few end stock-points and extreme service targets. Even then, however, when $\tilde{P}_B$ can be low, respective models could still remain accurate (see simulation results in van Donselaar and Wijngaard, 1987, for a practical validation of this point).

6. Comparative results

In order to obtain some indication of the efficiency of different LR rules, numerical comparisons were performed for the BS and AFS rules. Specifically, for given end stock-point α_i service targets, we compared the resulting system performance for both rules under identical conditions. Note that, on the basis of an existing Newsboy-style result (see Diks and de Kok, 1998), supply chain control variables need to satisfy specific α_i targets (evaluated from the respective cost coefficients) at the cost-optimal setting. Computations (with the MatLab-v6.1 standard software) were performed using a simple procedure by van der Heijden (1997), adapted to the non-dimensional representation of the models developed here and the use of α_i targets instead of fill rate targets (see Appendix C).

An environment similar to that in Section 5.2 was used for comparison. We considered a chain with end stock-points forming two equally sized groups, $\mathcal{GA}$ and $\mathcal{GB}$, each having end stock-points with identical demand processes. For each group we also imposed identical lead-time structures and service targets. Both the symmetric and the non-symmetric (in its two variations) demand structures were considered under conditions of approximate normality. In order to ensure observable differences for the two rules, the system inputs were chosen accordingly. Specifically, the following inputs were fixed:

$$\Delta = 0, \quad L = 5, l_i = 10 \quad \text{for } i \in \mathcal{GA} \text{ and } l_i = 2 \quad \text{for } i \in \mathcal{GB}.$$

Note that $\Delta = 0$ corresponds to supply chains with no centrally hold inventory, where the effects of the rationing rule are most severe. The selected lead-time structures (combined with the demand parameters) lead to observable rationing fractions differences between the BS and AFS rules. In order to ensure results under comparable balance conditions, respective $\tilde{P}_B$ estimates for each setting were also made (via Monte Carlo simulation).

Two indicators were used to measure rationing rules efficiency: total average backorders (B_{TOT}) and total average end-echelon on hand inventory (I_{TOT}), both calculated at the end of an arbitrary period. We simply note that both indicators are elements of the cost model (with linear holding and backorders costs) dominating multi-echelon research (van Houtum et al., 1996; Axsater, 2003). In fact, the aggregate form used here corresponds to the case of undifferentiated holding and shortage costs for end stock-points. Under the conditions we study here, these are given by

$$B_{TOT} = \sum_{j=1}^{N} (1-\gamma_j) M_j \sigma_j \sqrt{l_j + 1} \text{ and}$$

$$I_{TOT} = B_{TOT} + S_1 - \sum_{j=1}^{N} (L + l_j + 1)\mu_j.$$

The B_{TOT} expression is general and can be evaluated directly using (13). The expression for I_{TOT} only applies for $\Delta = 0$. Note that its second term corresponds to the unconditional average $\sum_1^j E\{I_1^j\}$, which is directly derived from (6).

Table 3 presents the computational results obtained (subscripts A and B in the table indicate data and results associated with groups $\mathcal{GA}$ and $\mathcal{GB}$, respectively). These correspond to four different sets of α_A/α_B group service targets combinations, providing a total of 24 distinct cases. Other than the respective service targets, each case is identified by the following inputs: period demand parameters (mean and standard deviation) for each group and total number of end stock-points in the chain N. The results corresponding to each case (for both end stock-point groups) are presented for the BS and the AFS rule separately. For each rule, these comprise of the rationing function parameters (rationing factors and fractions), the estimated balance probability, the resulting γ service levels and the two efficiency indicators. To facilitate comparisons, the I_{TOT} and B_{TOT} values that correspond to the rule manifesting the best respective efficiency in each case are underlined.

From the results in Table 3, we can directly observe that the factor dominating the rationing rules relative efficiency is α_A/α_B. For identical service targets for both groups (i.e. $\alpha_A = \alpha_B$), the AFS rule (that, as discussed in Section 3.3, reduces to FS in these cases) always outperforms the BS rule for both I_{TOT} and B_{TOT}. For non-identical α_A/α_B targets, however, respective results are inconclusive, with AFS and BS outperforming each other in terms of either I_{TOT} or B_{TOT} (never both), depending on the relative values of α_A/α_B. This behavior is consistent irrespective of either the end stock-point groups demand process specifications or the total number of end stock-points in the supply chain. It is interesting to note that, in all cases examined, the estimated balance probabilities (observe all $\tilde{P}_B$ values in the table) were very high for either rationing rule, therefore imbalance is negligible and does not affect the observed behavior.

The implication of these limited results is fairly clear. Except for identical α-service targets for all end stock-points (corresponding to cases with undifferentiated holding and backorders costs) where AFS is always optimal (Bollapragada et al., 1999), the relative efficiency of AFS and BS depends on the environmental conditions under consideration. This is primarily influenced by the magnitude of the observed α-service targets differences (reflecting holding and backorders costs differentials between end stock-points), as well as by the respective lead times and (to a much smaller extent, if at all) by the demand process parameters or by the

Table 3
Comparative results for the BS and AFS rules under different service targets and environmental conditions.

α_A/α_B	$\mu_A(\sigma_A)/\mu_B(\sigma_B)$	N	BS rule							AFS rule						
			f_A/f_B	a_A/a_B	S_1^A/S_1^B	γ_A/γ_B	$\tilde{P}_B$	I_{TOT}	B_{TOT}	f_A/f_B	a_α/a_B	S_1^A/S_1^B	γ_A/γ_B	$\tilde{P}_B$	I_{TOT}	B_{TOT}
0.95/0.95	1000(350)/2000(500)	2	0.264/0.736	−894.1/894.1	16966/19214	0.975/0.986	99.9	4233.1	53.1	0.573/0.427	0/0	21893/14127	0.971/0.989	98.2	4070	51
		6	0.264/0.736	−973.7/973.7	16906/18748	0.975/0.989	99.4	11101.3	139.3	0.573/0.427	0/0	21639/13938	0.974/0.990	98.2	10867.3	136.3
	1000(250)/1000(350)	2	0.419/0.581	−312.9/312.9	16705/10167	0.981/0.983	99.9	2908.5	36.5	0.578/0.422	0/0	18418/8424	0.979/0.985	99.9	2878.1	36.1
		6	0.419/0.581	−349.8/349.8	16606/9941	0.982/0.986	99.5	7738	97	0.578/0.422	0/0	18239/8293	0.981/0.986	99.5	7692.5	96.5
	1000(350)/1000(350)	2	0.5/0.5	−382.5/382.5	18115/9350	0.973/0.983	99.7	3509.1	44.1	0.657/0.343	0/0	19822/7607	0.971/0.985	99.1	3472.5	43.5
		6	0.5/0.5	−426.6/426.6	17980/9127	0.975/0.986	98.7	9439.5	118.5	0.657/0.343	0/0	19600/7491	0.974/0.987	98.5	9390.7	117.7
0.75/0.75	1000(350)/2000(500)	2	0.264/0.736	−366.7/366.7	15786/17928	0.819/0.901	99.9	2093,1	379.1	0.573/0.427	0/0	20535/13113	0.791/0.922	98.2	2012.4	364.4
		6	0.264/0.736	−399.3/399.3	15762/17736	0.824/0.922	99.4	5488.2	994.2	0.573/0.427	0/0	20431/13036	0.814/0.931	98.2	5374	973
	1000(250)/1000(350)	2	0.419/0.581	−128.2/128.2	15811/9367	0.863/0.877	99.9	1438.4	260.4	0.578/0.422	0/0	17450/7716	0.851/0.891	99.9	1423.7	257.7
		6	0.419/0.581	−143.4/143.4	15770/9274	0.872/0.898	99.5	3824.4	692.4	0.578/0.422	0/0	17376/7662	0.867/0.903	99.5	3802.7	688.7
	1000(350)/1000(350)	2	0.5/0.5	−156.9/156.9	16867/8554	0.808/0.878	99.7	1735.3	314.3	0.657/0.343	0/0	18493/6913	0.796/0.893	99.1	1717	311
		6	0.5/0.5	−174.9/174.9	16812/8462	0.820/0.898	98.7	4667.6	845.6	0.657/0.343	0/0	18402/6866	0.816/0.904	98.5	4644.7	840.7
0.95/0.75	1000(350)/2000(500)	2	0.264/0.736	−1.2/1.2	16966/17928	0.975/0.901	99.9	3117.2	223.2	0.573/0.427	−580.2/580.2	21893/13113	0.971/0.922	98.2	3190.9	184.9
		6	0.264/0.736	−1241.1/1241.1	16906/17736	0.975/0.922	99.4	8466.4	540.4	0.573/0.427	−516.4/516.4	21639/13036	0.974/0.931	98.2	8518.8	493.8
	1000(250)/1000(350)	2	0.419/0.581	−648/648	16705/9367	0.981/0.877	99.9	2214.2	142.2	0.578/0.422	−409/409	18418/7716	0.979/0.891	99.9	2263.7	129.7
		6	0.419/0.581	−629/629	16606/9274	0.982/0.898	99.5	6001.1	361.1	0.578/0.422	−364.3/364.3	18239/7662	0.981/0.903	99.5	6049.6	346.6
	1000(350)/1000(350)	2	0.5/0.5	−780.8/780.8	18115/8554	0.973/0.878	99.7	2818.4	149.4	0.657/0.343	−456/456	19822/6913	0.971/0.893	99.1	2870.3	135.3
		6	0.5/0.5	−759.1/759.1	17980/8462	0.975/0.898	98.7	7708.3	382.3	0.657/0.343	−410.9/410.9	19600/6866	0.974/0.904	98.5	7763.8	365.8
0.75/0.95	1000(350)/2000(500)	2	0.264/0.736	−26.5/26.5	15786/19214	0.819/0.986	99.9	3209.1	209.1	0.573/0.427	580.2/−580.2	20535/14127	0.791/0.989	98.2	2892.5	230.5
		6	0.264/0.736	−131.9/131.9	15762/18748	0.824/0.989	99.4	8123.1	593.1	0.573/0.427	516.8/−516.8	20430/13938	0.814/0.990	98.2	7720.9	616.9
	1000(250)/1000(350)	2	0.419/0.581	206.9/−206.9	15811/10167	0.863/0.983	99.9	2132.7	154.7	0.578/0.422	409/−409	17450/8424	0.851/0.985	99.9	2038.1	164.1
		6	0.419/0.581	135.8/−135.8	15770/9941	0.872/0.986	99.5	5561.3	428.3	0.578/0.422	364.3/−364.3	17376/8293	0.867/0.986	99.5	5445.6	438.6
	1000(350)/1000(350)	2	0.5/0.5	241.4/−241.4	16867/9350	0.808/0.983	99.7	2426	209	0.657/0.343	456/−456	18493/7607	0.796/0.985	99.1	2319.2	219.2
		6	0.5/0.5	157.6/−157.6	16812/9127	0.820/0.986	98.7	6398.8	581.8	0.657/0.343	410.9/−410.9	18402/7491	0.816/0.987	98.5	6271.7	592.7

actual number of end stock-points in the supply chain. Focused further research on this issue is certainly required.

7. Conclusions and further research

This paper has analyzed simple divergent two-echelon supply chains in view of modeling their service performance for the linear rationing (LR) class of practical rationing rules. By introducing an additional set of parameters (the rationing factors), we proposed an alternative definition of the LR class rationing function that allows the treatment of the rationing rule as a distinct control entity (separate from order-up to levels). We formally showed that (as previously suggested) the CAS and FS rules are members of the LR class, determined the rationing parameters for each rule and discussed their limitations. Due to inherent deficiencies, our analysis suggests CAS as a practically inefficient rule, of little value for future applications.

We also proposed AFS, a new LR rule, which constitutes the natural extension of FS. AFS not only ensures system controllability (for any performance objective) but also reduces to FS under the conditions where the use of the latter is desirable (i.e. to ensure identical stock-out probabilities at end stock-points). As a result, AFS can fully replace FS in all applications. The limited comparative results obtained for BS and AFS and the observed efficiency of the latter (under normal demands) clearly identify AFS as a rule worth of further consideration.

Considering three standard measures of customer service, we proposed analytical models for each measure, strictly valid for normal demand processes. The models apply for any LR rule and are given in closed form, hence can be evaluated exactly using standard software and handbooks (thus eliminating the need for approximations). The models accuracy, however, critically depends on the satisfaction of the standard balance assumption. After modeling a surrogate measure of the balance probability, we explored the severity of this assumption using Monte Carlo simulation. The results clearly revealed the balance assumption to be robust under fairly general conditions. Therefore, provided that end stock-point demands can be accurately modeled as normal processes (see Tang and Grubbstrom, 2006, for a discussion of this common modeling assumption) and do not differ drastically from each other, the proposed service models may be safely used in practical applications.

Appendix A. Omitted proofs

This appendix presents the proofs of the two propositions introduced in this paper.

A.1. Proof of Proposition 3.1

For the first part, we simply need to show that (4), implies the above. Let us define

$$a_i = f_i\left[S_1 - \sum_{j=1}^{N}(l_j+1)\mu_j\right] - S_1^i + (l_i+1)\mu_i. \tag{A.1}$$

Since $\sum_{j=1}^{N} f_i = 1$, summing both terms of the above for all stock-points i, the r.h.s. of the summation becomes zero. Clearly, this implies that $\sum_{j=1}^{N} a_j = 0$. Solving now (A.1) for S_1^i and replacing in (4), we obtain

$$J_1^i(t) = f_i\left[S_1 - \sum_{j=1}^{N}(l_j+1)\mu_j\right] - f_i \max[0, D_0(t-L, t-1) - \mathbf{\Delta}] + (l_i+1)\mu_i - a_i,$$

which, after some algebra, becomes

$$J_1^i(t) = f_i\left\{\min[S_1 + \mathbf{\Delta} - D_0(t-L, t-1), S_1] - \sum_{j=1}^{N}(l_j+1)\mu_j\right\} + (l_i+1)\mu_i - a_i.$$

But, from the definition of Δ and (1) we see that $\min[S_1 + \Delta - D_0(t-L, t-1), S_1] = J_1(t)$, so the proof of the first part is complete. For the second part, since each of the above steps is reversible, the poof follows directly.

A.2. Proof of Proposition 5.1

Consider any end stock-point i for two consecutive periods $t-1$ and t. From material equilibrium and the definitions in Section 2.1, we always have

$$J_1^i(t) = J_1^i(t-1) - d_i(t-1) + R_i(t), \tag{A.2}$$

where $R_i(t)$ is the positive allocation at the start of period t. Let us assume now that we have balance at the start of period $t-1$. Then we can directly evaluate the $R_i(t)$ necessary in order also to have balance after allocation in period t:

$$R_i(t) = f_i[J_1(t) - J_1(t-1)] + d_i(t-1). \tag{A.3}$$

This is obtained from (A.2), using (5) for $J_1^i(t-1)$ and $J_1^i(t)$. In order for this to be feasible (under the system operating assumptions) we simply need that $R_i(t)\geqslant 0$ for all $i=1,\ldots,N$. Equivalently, from (A.3):

$$J_1(t-1)-J_1(t)\leqslant \min_{1\leqslant j\leqslant N}\left[\frac{d_j(t-1)}{f_j}\right].$$

Therefore, we can directly determine probability $\tilde{P}_B$:

$$\begin{aligned}\tilde{P}_B &= Pr\left\{J_1(t-1)-J_1(t)\leqslant \min_{1\leqslant j\leqslant N}\left[\frac{d_j(t-1)}{f_j}\right]\right\}\\ &= Pr\left\{J_1(t-1)-J_1(t)\leqslant \min_{1\leqslant j\leqslant N}\left[\frac{d_j(t-1)}{f_j}\right] \text{ and } U_0(t)<S_1\right\}\\ &\quad + Pr\{U_0(t)\geqslant S_1\}. \end{aligned} \tag{A.4}$$

The second equality above is obtained by conditioning on the value of $U_0(t)$ and considering that when $U_0(t)\geqslant S_1$ we have always balance (since all S_1^i are by definition balanced). Introducing expressions (1) for $J_1(t-1)$, $J_1(t)$ and $U_0(t)$ in (A.4), after some algebra the required expression prevails.

Appendix B. Non-dimensional balance probability model

This appendix presents the non-dimensional $\tilde{P}_B$ model, corresponding to two equally sized stock-point groups with identical demand characteristics, as used to test the balance assumption.

Let us introduce the variables (subscripts A and B refer to variables in groups $\mathcal{GA}$ and $\mathcal{GB}$, respectively):

$$\mathrm{CV_A}=\sigma_A/\mu_A \text{ and } \mathrm{CV_B}=\sigma_B/\mu_B$$
$$b=\mu_B/\mu_A,$$
$$\theta=\boldsymbol{\Delta}/E\{D_0\}.$$

Using these variables, $\tilde{P}_B$ can be expressed as

$$\tilde{P}_B = Pr\{C\geqslant \min(A,B) \text{ and } A>0\} + Pr\{A\leqslant 0\},$$

where

$$\begin{aligned} A &= \frac{LN}{2\mathrm{CV}_A}(1-\theta)(1+b)+\sum_{j\in\mathcal{GA}} v_j+\sum_{i=1}^{L-1}\sum_{j\in\mathcal{GA}} p_{ij}\\ &\quad + b\frac{\mathrm{CV_B}}{\mathrm{CV_A}}\left(\sum_{j\in\mathcal{GB}} v_j+\sum_{i=1}^{L-1}\sum_{j\in\mathcal{GB}} p_{ij}\right),\end{aligned}$$

$$B=\sum_{j\in\mathcal{GA}}(v_j-w_j)+b\frac{\mathrm{CV_B}}{\mathrm{CV_A}}\sum_{j\in\mathcal{GB}}(v_j-w_j),$$

$$C=\frac{1}{\mathrm{CV_A}}\min\left\{\min_{k\in\mathcal{GA}}\left(\frac{1+\mathrm{CV_A}v_k}{f_A}\right),\ b\min_{k\in\mathcal{GB}}\left(\frac{1+\mathrm{CV_B}v_k}{f_B}\right)\right\}.$$

The variables w_j, v_j and p_{ij} are independent normal $N(0,1)$ and are obtained by standardizing the original variables in the $\tilde{P}_B$ expression given by Proposition 5.1. Note that when, $f_A=f_B=1/N$, $\mathrm{CV_A}=\mathrm{CV_B}$ and $b=1$ this expression reduces to that in Lagodimos (1992). In addition, when also $\Delta=0$, it reduces to that in Eppen and Schrage (1981).

Appendix C. Computational procedure

This appendix presents the computational procedure used for the numerical comparisons of different rationing rules under individual end stock-points α_i targets.

The procedure applies for any LR rule with predetermined rationing fractions and computes the system control parameters (i.e. order-up-to levels and LR rule rationing factors) for given Δ. Inputs to the procedure are: supply chain structure (N, L and l_i for all i), period demand processes (μ_i, σ_i for all i), the rationing fractions $\{f_i\}_{i=1}^N$ of the rule under study, α_i targets for all i and Δ. It comprises the following steps:

1. For each i, apply (7a) to determine W_i and r_i. Also apply (7b) to determine V and Z_Δ.
2. For each i, consider the α_i model in (12), which for fixed α_i represents an equation with Z_i being the only unknown variable. Apply any numerical technique (e.g. bisection) to solve for Z_i.
3. Use the N evaluated Z_i values to solve the system of $N+1$ linear equations in (8), first to obtain Z and then each T_i value in turn (for all i).
4. For each i, use the definitions of Z_i and T_i in (7a) to directly determine the order up to level S_1^i and the LR rule rationing factor a_i.

Having determined all control parameters, we can directly use (14) and (15) to evaluate both β_i and γ_i. Note that the procedure is easily adapted for using β or γ service targets.

References

Abramowitz, M., Stegun, A., 1965. Handbook of Mathematical Functions. Dover Publications, New York.

Axsater, S., 2003. Supply chain operations: Serial and distribution systems. In: de Kok, A.G., Graves, S.C. (Eds.), Supply Chain Management: Design, Coordination and Operation. Elsevier, Amsterdam, pp. 525–556.

Bollapragada, S., Akella, R., Srinivasan, R., 1999. Centalized ordering and allocation policies in a two-echelon system with non-identical warehouses. European Journal of Operational Research 106, 74–81.

Cao, D.B., Silver, E.A., 2005. A dynamic allocation heuristic for centralized safety stock. Naval Research Logistics 52, 513–526.

Clark, A.J., Scarf, H., 1960. Optimal policies for a multi-echelon inventory problem. Management Science 6, 475–490.

de Kok, A.G., 1990. Hierarchical production planning for consumer goods. European Journal of Operational Research 45, 55–69.

de Kok, T.G., Fransoo, J.C., 2003. Planning supply chain operations: Definition and comparison of planning concepts. In: de Kok, A.G., Graves, S.C. (Eds.), Supply Chain Management: Design, Coordination and Operation. Elsevier, Amsterdam, pp. 597–675.

de Kok, A.G., Lagodimos, A.G., Seidel, H.P., 1994. Stock allocation in a two-echelon distribution network under service constraints. Research Report DBK/LBS/94-03, Department of Industrial Engineering and Management Science, Eindhoven University of Technology, Netherlands.

Diks, E.B., de Kok, A.G., 1996. Controlling a divergent 2-echelon network with transhipments using the consistent appropriate share rationing policy. International Journal of Production Economics 45, 369–379.

Diks, E.B., de Kok, A.G., 1998. Optimal control of a divergent N-echelon inventory system. European Journal of Operational Research 111, 75–97.

Diks, E.B., de Kok, A.G., 1999. Computational results for the control of a divergent N-echelon inventory system. International Journal of Production Economics 59, 327–336.

Diks, E.B., de Kok, A.G., Lagodimos, A.G., 1996. Multi-echelon systems: A service measure perspective. European Journal of Operational Research 95, 241–263.

Eppen, G., Schrage, L., 1981. Centralized ordering policies in a multi-warehouse system with lead times and random demand. In: Schwarz, L.B. (Ed.), Multi-Level Production/Inventory Control Systems: Theory and Practice. North-Holland, Amsterdam, pp. 51–68.

Federgruen, A., 1993. Centralized planning models for multi-echelon inventory systems under uncertainty. In: Graves, S.C., Rinnoy Kan, A.H.G., Zipkin, P.H. (Eds.), Logistics of Production and Inventory. Elsevier, Amsterdam, pp. 133–173.

Inderfurth, K., 1994. Safety stocks in multistage divergent inventory systems: A review. International Journal of Production Economics 35, 321–329.

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

Lagodimos, A.G., 1992. Multi-echelon service models for inventory systems under different rationing policies. International Journal of Production Research 30, 939–958.

Lagodimos, A.G., 1993. Models for evaluating the performance of serial and assembly MRP systems. European Journal of Operational Research 68, 49–68.

Lagodimos, A.G., Anderson, E.J., 1993. Optimal positioning of safety stocks in MRP. International Journal of Production Research 31, 1797–1813.

Lagodimos, A.G., de Kok, A.G., Verrijdt, J.H.C.M., 1995. The robusteness of multi-echelon service models under autocorrelated demands. Journal of the Operational Research Society 46, 92–103.

McGavin, E.J., Schwarz, L.B., Ward, J.E., 1993. Two-interval inventory-allocation policies in a one-warehouse N-identical retailer distribution system. Management Science 39, 1092–1107.

Mula, J., Poler, R., Garcia-Sabater, J.P., Lario, F.C., 2006. Models for production planning under uncertainty: A review. International Journal of Production Economics 103, 271–285.

Schneider, H., 1981. Effect of service levels on order-points and order-levels in inventory models. International Journal of Production Research 19, 615–631.

Silver, E.A., Pyke, D.F., Peterson, R., 1998. Inventory Management and Production Planning and Scheduling. Willey, New York.

Tang, O., Grubbstrom, R.W., 2006. On using higher-order moments for stochastic inventory systems. International Journal of Production Economics 104, 454–461.

van der Heijden, M.C., 1997. Supply rationing in multi-echelon divergent systems. European Journal of Operational Research 101, 532–549.

van der Heijden, M., 2000. Near cost-optimal inventory control policies for divergent networks under fill rate constraints. International Journal of Production Economics 63, 161–179.

van der Heijden, M.C., Diks, E.B., de Kok, A.G., 1997. Stock allocation in general multi-echelon distribution systems with (R,S) order-up-to policies. International Journal of Production Economics 49, 157–174.

van Donselaar, K., Wijngaard, J., 1986. Practical application of the echelon approach in a system with divergent product structure. Lecture Notes in Economics and Mathematical Systems 266, 182–196.

van Donselaar, K., Wijngaard, J., 1987. Commonality and safety stocks. Engineering Costs and Production Economics 12, 197–204.

van Houtum, G.J., Inderfurth, K., Zijm, W.H.M., 1996. Materials coordination in stochastic multi-echelon systems. European Journal of Operational Research 95, 1–23.

Verrijdt, J.H.C.M., de Kok, A.G., 1995. Distribution planning for a divergent N-echelon network without intermediate stocks under service restrictions. International Journal of Production Economics 38, 225–243.

Verrijdt, J.H.C.M., de Kok, A.G., 1996. Distribution planning for a divergent depotless two-echelon network under service restrictions. European Journal of Operational Research 89, 341–354.

Zhang, J., 2003. Managing multi-customer service level requirements with a simple rationing policy. Operations Research Letters 31, 477–482.

Zipkin, P., 1984. On the imbalance of inventories in multi-echelon systems. Mathematics of Operations Research 28, 445–475.