Annals of Operations Research (2023) 331:899–921
https://doi.org/10.1007/s10479-022-05061-z

ORIGINAL RESEARCH

Analysis of second order properties of production–inventory systems with lost sales

Nha-Nghi de la Cruz[1] · **Hans Daduna**[2]

Accepted: 1 November 2022 / Published online: 16 November 2022

Abstract
We consider a single-item production–inventory system under a base stock policy for inventory control. We model the system as a closed Gordon–Newell network. The population size of the network is equal to the base stock level, which is the sum of the finished goods and work-in-process inventory. Each exogenous demand, which follows a Poisson process, releases a production order for a new unit and increases the amount of the work-in-process inventory. When there are no items in the finished goods inventory available, arriving demand is lost. The replenishment network operates with state dependent service rates, which we assume to be increasing and concave. First, we analyze the queue length behavior of a two node system and provide conditions under which the mean queue length at the production server is convex in the number of customers in the system. We prove that this leads to convexity of a standard cost function. Using Norton's theorem, we are able to generalize our results for arbitrarily large production–inventory systems.

Keywords Production–inventory theory · Lost sales · Second order properties · Queueing network theory · Norton's theorem

1 Introduction

We consider an integrated production–inventory system consisting of a finished goods inventory controlled by a base stock policy and an affiliated production network which serves as a replenishment system in a make-to-stock mode. When there is inventory, an arriving exogenous demand for a produced item from the system is satisfied immediately with exactly one item, and an order for producing a new item is instantly sent to the replenishment system. Arriving demand that finds the inventory depleted departs immediately without fulfillment and is lost. The control parameter for the inventory is the base stock level $z \geq 1$. That means whenever an item from stock is dispatched to an arriving demand, and the inventory

✉ Nha-Nghi de la Cruz
nha-nghi.de.la.cruz@uni-hamburg.de

[1] Faculty of Business Administration, Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

[2] Department of Mathematics, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

level is below this base stock level an order for the production of a new item is sent to the replenishment system.

We address the problem of determining a base stock level that minimizes a standard cost function. This cost function encompasses holding costs for items in the inventory, production costs for items in the replenishment system, and costs for unfulfilled demand (lost sales).

Due to the lost sales behavior of arriving demand, this integrated production–inventory network can be modeled as a closed network of queues with a fixed number of customers cycling in the system. Assuming the servers of the replenishment network to be exponential single servers and the stream of demands to be a Poisson process, Rubio and Wein (1996) show that the integrated system is a Gordon–Newell network (Gordon and Newell, 1967). Using queueing theory, the corresponding problem with backordering unsatisfied demand at integrated production–inventory systems under base stock control has already been successfully investigated for many years. However, almost only single replenishment servers have been subject of the investigations.

Until now, the case of production–inventory systems under base stock control with rejection of arriving demand due to depleted inventory has not been well understood, see Sections 4.2 and 4.3 in Rubio and Wein (1996), where only a very simple example could be solved analytically. We tackle this problem in the present article and consider a single item production–inventory system with lost sales. A sale is lost when there is no inventory available to serve an arriving demand (customer) immediately. We model our base stock system as a Gordon Newell network where a production order is released whenever demand is fulfilled. We analyze the cost function and focus on the queue length behavior of the system. Via numerical experiments, Rubio and Wein (1996) have found out that the cost function of a production–inventory system under base stock control with lost sales is in general not convex. For a production–inventory system where all replenishment stations and the finished goods inventory have the same traffic intensity (called "balanced system"), they prove that the cost function is a convex function of the base stock level. For a general replenishment network with an affiliated finished goods inventory, our main result states that if the holding costs dominate production costs the total cost function is convex in the base stock level under a few natural conditions.

Our paper is structured as follows: In Sect. 1.1, we provide an introduction into relevant literature on production–inventory systems. In Sect. 2, the detailed system is described, and the transformation into a closed network of exponential queues is given explicitly. In Sect. 3, the case of replenishment networks that consist of only a single station is analyzed. We demonstrate that the main technical problems can be solved by considering this fundamental setup. In a second step, we are able to reduce the general situation to this case. Details are presented in Sect. 4. In Sect. 5, we provide numerical evaluations for a two-station network, and Sect. 6 contains a short conclusion. Technical lemmata and lengthy proofs are collected in the "Appendix".

Notation and conventions: $\mathbb{N} = \{1, 2, \dots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$; increasing means nondecreasing and decreasing means nonincreasing.

1.1 Literature review

Analyzing production–inventory systems by using methods and models from queueing theory dates back to at least the 1950's, see "Queues, Inventories and Maintenance" (Morse, 1958). A great variety of different systems have been investigated since then and the available literature is overwhelming. We sketch a selection of such work with emphasis on papers related to our analysis of an inventory under base stock control with an attached production network. Since in general the replenishment servers or the replenishment networks have limited capacity, a base stock policy is a reasonable control scheme (Federgruen and Zipkin, 1986a, b) although not necessarily optimal for systems without backordering unsatisfied demand (Hill, 1999).

Fundamentals of "Continuous time stochastic demand models" with different reorder policies for a single inventory are surveyed by Lee and Nahmias (1993)[Section 5]. In Section 5.4, base stock policies under the heading of "Order-for-order inventory policies" are described, which are special cases of (s, S)-policies (with $s = S - 1$). The case of multi-stage inventories under base stock policies are dealt with by Axsäter (1993)[Section 2]. As an introductory survey of the history of "production–inventory systems", the "Notes" in Zipkin (2000) can be used, especially the "Notes" in Chapter 6 (p. 237) and Chapter 7 (p. 289).

Base stock policy and backordering of unsatisfied demand. Most of the queueing theory applications that analyze inventory systems tackle problems with backordering unsatisfied demand. Usually the early investigations consider either a single replenishment server or assume that reorders arrive after a random lead time. The latter can be modeled as an $M/G/\infty$ single station replenishment facility. Kaplan (1970) derives a dynamic programming formulation of the random lead time model where the outstanding orders cannot cross (lead times are therefore strongly correlated). Karmarkar (1987) investigates the interplay of lot sizing, manufacturing lead times and in-process inventories using standard queueing models to evaluate performance of manufacturing systems. Similar investigations in the realm of Flexible Manufacturing Systems (FMS) are numerous, see e.g. the survey by Karmarkar (1993). Different aspects of optimization criteria are "cost targets" where various risk measures are considered as foundation for management decisions in production–inventory systems and as a measure of quality. Recent works in this direction are by Li et al. (2020) (criterion: probability of loss and expected loss), Li and Arreola-Risa (2021) (criterion: total cost conditional value-at-risk).

Bradley and Glynn (2002) investigate the interplay of the base stock level and the capacity of a single replenishment server using heavy-traffic analysis of queueing systems.

Analysis of simple (open tandem) networks as replenishment systems have been conducted by Lee and Zipkin (1992, 1995). For production–inventory systems with a general production network of Jacksonian type (Chen and Yao, 2001)[Chapter 2] and with backordering under base stock control, Rubio and Wein (1996) prove that setting the optimal base stock resembles decision making in the newsvendor problem. For some special cases, they have determined structural properties of the optimal policies. These results have been extended and generalized by de la Cruz and Daduna (2016, 2019).

A problem different from that investigated in these articles but closely related emerges when the production servers in the replenishment system have attached local inventories. An example with explicit stationary distribution for an open tandem supply network is provided by Song and Zipkin (2009)[Section 3] with a modified base stock policy. In general, the joint decision for optimal local inventory control is not analytically tractable. In their review, Atan et al. (2017) state "…an important result for continuous-review single end product models with multiple demand classes is the optimality of state-dependent base-stock production/replenishment policy …".

Production–inventory systems with unreliable replenishment server are considered by Hu and Xiang (1993) for the case of backordering.

Related models are production–inventory systems under base stock control where the time to dispatch an item to the arriving demand is positive (usually random), for surveys see Krishnamoorthy et al. (2011) and Krishnamoorthy et al. (2021).

Base stock policy and lost sales of unsatisfied demand. Compared with the case of backordering unsatisfied demand, the lost sales behavior of demand that is not immediately fulfilled has found only limited interest. The first analysis of the topic seems to be by Karush (1957) who investigates an inventory in continuous time controlled by a base stock policy with random lead times (independent, identically distributed) for reorders. This can be considered as an $M/M/\infty$-replenishment server. Karush proves convexity of stock-out probability in the base stock level.

Related but different models are investigated e.g. in case of periodic review of inventory level (Janakiraman and Roundy, 2004) and in discrete time (Zipkin, 2008). Janakiraman and Roundy (2004) prove for the case of stochastic lead times which do not cross (i.e. an $M/M/\infty$-replenishment server with strongly correlated service times) convexity of the expected discounted sum of holding and lost sales costs in the base stock level. Zipkin (2008) proves structural properties of several ("plausible") policies and evaluates numerically some of the plausible policies. Johansen (2013) considers the counterexample of Hill (1999) and modifies the pure base stock policy in lost sales systems by introducing a lower bound on the time between placing replenishment orders.

The main reference of the system analyzed in this article is Rubio and Wein (1996) where an inventory under backordering and under lost sales for unsatisfied demand are considered. The replenishment system in both cases is a complex network of exponential service stations with random routing (processing networks in the terminology of Zipkin (2000)[Section 7.3.6]). In case of a tandem supply system under lost sales, Song and Zipkin (2009) show that for the case of a normal and an emergency supply source (node) the stationary distribution is accessible via generalized Gordon–Newell network models.

For general exponential networks with lost sales in the context of stock-out under base stock control, Rubio and Wein (1996) prove convexity of the cost function in the base stock level when the inventory and processing network is totally symmetric ("balanced" in their terminology). Thus far, more results similar to those for backordering by Rubio and Wein (1996), and de la Cruz and Daduna (2016, 2019) seem to be not available for the lost sales case.

Production–inventory systems with unreliable replenishment server are considered in Liu and Cao (1997) for the case of lost sales.

For production–inventory systems with positive service time, early results with product form steady state (similar to Rubio and Wein (1996) and the present article) are derived in Schwarz et al. (2006), Schwarz et al. (2007), and more recently in Krishnamoorthy and Viswanath (2013) and Otten et al. (2016). Toktay et al. (2000) illustrate a practical application of a closed queueing network to model a remanufacturing system with lost sales.

A comparison of different reorder policies, including base stock policy, for systems under lost sales are provided by Bijvank and Vis (2011, 2012).

2 The model

The production–inventory system in this article consists of a replenishment network of K service stations, indexed by $\{1, 2, \ldots, K\}$, and the finished goods inventory, indexed by $K + 1$. The base stock level is denoted by $z \geq 1$. An arriving demand that finds finished goods available at the inventory takes one item, triggers an order for production of a new unit in the replenishment network and leaves the system. If the inventory is depleted, an arriving demand will depart immediately without obtaining an item (lost sales).

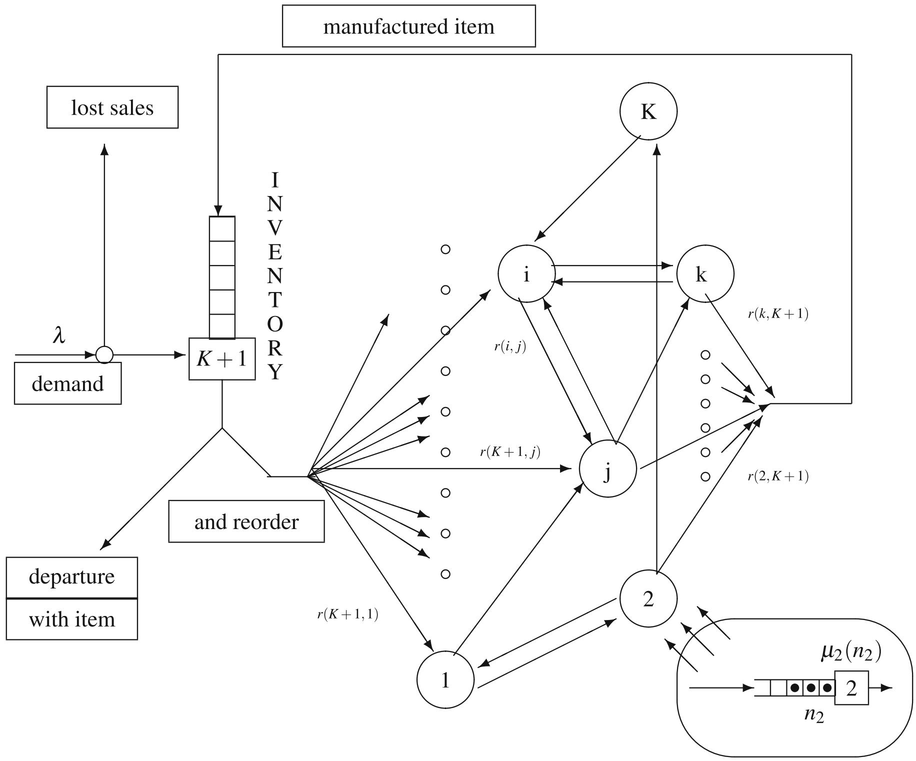

Fig. 1 Structure of the production–inventory system

To model the entire production–inventory system, we use a Gordon Newell network with $z \geq 1$ customers and $K + 1$ exponential single server stations $\overline{J} = \{1, \ldots, K, K + 1\}$. Stations $j = 1, \ldots, K$ represent the (replenishment-)production system, and the finished goods inventory is represented by station $K + 1$. Service intensities at nodes $j = 1, \ldots, K$ are allowed to be queue length dependent $\mu_j(n) > 0, n = 1, \ldots, z$. The population size z equals the base stock level. Routing of customers is governed by an irreducible stochastic matrix $R = (r(i, j) : i, j \in \{1, \ldots, K + 1\})$: With probability $r(K + 1, j)$, a new order for production of an unit is sent to station $j \in \{1, \ldots, K\}$. An order served at station j is sent to station $i \in \{1, \ldots, K\}$ with probability $r(j, i)$, and with probability $r(j, K + 1)$ it is sent as finished good to the inventory at station $K + 1$.

Along their paths through the network, the customers of the Gordon–Newell network therefore take different identities with respect to the production–inventory system. Customers at station $K + 1$ are items of the finished goods in stock. When leaving station $K + 1$ (triggered by an arrival of demand), a customer is transformed into an order which passes through the replenishment subsystem according to the routing matrix R and is processed at some subsequent stations out of $1, \ldots, K$. When reentering station $K + 1$, the customer is converted to a new finished goods item (see Fig. 1).

The lost sales property of the external demand, which arrives in a Poisson-λ stream, implies that an arrival that finds no customer at $K + 1$ (= stock-out) does not change the state of the network. An arrival that finds positive queue length at $K + 1$ (= finished goods stock

> 0) takes one item (= customer) from the queue, departs from the system, and triggers a replenishment order.

Hence by construction and the lost sales property, customers present at $K+1$ leave that node with exponential-λ interdeparture times (= service times) as long as customers are present there. Therefore, station $K+1$ operates with a state independent service rate $\mu_{K+1}(n) =: \mu_{K+1} = \lambda, n \geq 1$.

Note that whenever the total population of the network resides in the finished goods station $K+1$ there is no ongoing replenishment activity in the production sub-system $\{1, \ldots, K\}$, and if there are fewer than z "customers" present at $K+1$ there are ongoing production processes at some of the stations $1, \ldots, K$.

Imposing the usual (conditional) independence assumptions on the stochastic behavior of the systems implies that the joint queue length vector for the system is a continuous time ergodic Markov process $X = ((X_j(t), j = 1, ..., K+1) : t \geq 0)$ on state space

$$S(z, \overline{J}) := \{(n_1, \ldots, n_{K+1}) \in \mathbb{N}_0^{K+1} : n_1 + \cdots + n_{K+1} = z\}. \tag{1}$$

Here, $X_{K+1}(t)$ is the size of the on-hand inventory, while $X_j(t)$ is the number of replenishment orders waiting or in service at station j at time $t \geq 0$, $j \in \{1, \ldots, K\}$. The dynamics of the system is governed by the transition intensity matrix $Q = (q(x, y) : x, y \in S(z, \overline{J}))$ as follows. For a generic state $(n_1, \ldots n_i, \ldots n_j, \ldots n_{K+1}) \in S(z, \overline{J})$, we have:

$$\begin{aligned}
&q((n_1, \ldots n_i, \ldots n_j, \ldots n_{K+1}), (n_1, \ldots n_i - 1, \ldots n_j + 1, \ldots n_{K+1})\\
&\quad = \mu_i(n_i) \cdot r(i, j), \quad n_i > 0,\\
&q((n_1, \ldots n_j, \ldots n_{K+1}), (n_1, \ldots n_j + 1, \ldots n_{K+1} - 1)\\
&\quad = \lambda \cdot r(K+1, j), \quad n_{K+1} > 0,\\
&q((n_1, \ldots n_j, \ldots n_{K+1}), (n_1, \ldots n_j - 1, \ldots n_{K+1} + 1)\\
&\quad = \mu_j(n_j) \cdot r(j, K+1), \quad n_j > 0.
\end{aligned}$$

The stationary distribution π_z of X exists and is the (unique) probability solution of the global balance equation $\pi \cdot Q = 0$ with

$$\begin{aligned}
&\pi_z(n_1, \ldots, n_{K+1})\\
&= \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j}{\mu_j(k_j)} \right) \left(\frac{\eta_{K+1}}{\mu_{K+1}} \right)^{n_{K+1}} G(z, \overline{J})^{-1}, (n_1, \ldots, n_{K+1}) \in S(z, \overline{J}),
\end{aligned} \tag{2}$$

where $G(z, \overline{J})$ is the normalization constant and $\eta = (\eta_j : j = 1, \ldots, K+1)$ is the probability solution of the traffic equation $\eta = \eta \cdot R$.
We denote henceforth by $(X_j, j = 1, ..., K+1)$ a random vector distributed according to the stationary distribution (2) of X.

The stationary throughput of the network with z customers, i.e. the mean total number of departures per time unit over all nodes $1, \ldots, K+1$ is $TH(z) = G(z-1, \overline{J})/G(z, \overline{J})$, see Chen and Yao (2001)[(2.14) in Section 2.5]. Therefore, the mean total number of departures per time unit from node j (throughput at j) is $TH_j(z) = \eta_j \cdot TH(z)$.

The cost structure of our problem is as follows: Work-in-process (WIP) cost per item is c (per time unit), finished goods inventory cost per item is h (per time unit), and cost ℓ is incurred for every lost sale. It follows from the definition of the throughput that the steady state mean number of lost sales per time unit under base stock level z is

$$\lambda - TH_{K+1}(z). \tag{3}$$

The global steady state cost function is (with $E_z(\cdot)$ the expectation under base stock level z)

$$C(z) = hE_z[X_{K+1}] + c\sum_{j=1}^{K} E_z[X_j] + \ell\,(\lambda - TH_{K+1}(z))\,. \tag{4}$$

For the companion problem with backordering in Rubio and Wein (1996), it is proved that the optimal base stock z^* can be computed explicitly and the solution is of the form of the news-vendor problem's solution. A similar procedure seems not possible for the lost sales system considered here. Rubio and Wein notice that the cost function is convex if $h = c$. For the "balanced" case in which all $K + 1$ single server stations have state independent service rates $\mu_j(n_j) = \mu_j, n_j = 1, \ldots, z$ and the same utilization $\rho_j = \eta_j/\mu_j = \rho$ for $j = 1, \ldots, K + 1$, they deliver a closed form expression for the cost-optimal z (Rubio and Wein, 1996)[Formula (14)]. Moreover, they state that for $h \geq c$ the total cost function is in general not convex, see Rubio and Wein (1996)[Section 4.2]. We therefore will disclose structural properties of the cost function (4) and prove that under some natural conditions **it is indeed convex** if $h \geq c$.

3 Single station replenishment network, $K = 1$

We start our analysis with the simplest case: an inventory with a replenishment network that consists of only one station. A similar setting has already been dealt with by Karush (1957), but with a somewhat different replenishment station. As shown in Rubio and Wein (1996)[4.2], this is a two-station Gordon–Newell network. We consider a slightly more general network which will enable us to approach the case of a general processing network in the next section.

In this section, we study the queue length behavior of a Gordon–Newell network with two nodes $\overline{J} := \{1, 2\}$ where the service rate of node 1, which represents the processing network, has queue length dependent service rates.

There are in total $z > 0$ customers cycling in the network. Node $j = 1$ has state dependent service intensities $\mu_1(n)$, $n \leq z$, which are increasing and concave in n. Node $j = 2$ has exponential service times with rate μ_2. (Recall that in the production–inventory system represented by this Gordon–Newell network holds $\mu_2 := \lambda$.)

The routing matrix $R = (r(i, j) : i, j \in \{1, 2\})$ is determined by $r(2, 1) = 1$ and $r(1, 2) = 1 - p$ for $p \in [0, 1)$. This indicates that an order is sent immediately to the replenishment server and that production of an item needs rework with probability $r(1, 1) = p$.

The cost function for this system is:

$$\begin{aligned} C(z) &= cE_z[X_1] + hE_z[X_2] + \ell\,(\lambda - \eta_2 \cdot TH(z)) \\ &= hz + (c - h)E_z[X_1] + \ell\,(\lambda - \eta_2 \cdot TH(z)) \end{aligned} \tag{5}$$

The following theorem presents the main result of Section 3:

Theorem 1 *In the production–inventory system with a single replenishment station, the cost function* (5) *is convex in the base stock level if the following holds:*
(i) The service rate function $\mu_1(\cdot)$ *is increasing and concave on {1,2,...}.*
(ii) It holds

$$\frac{\eta_1}{\mu_1(1)} \leq \frac{\eta_2}{\mu_2}. \tag{6}$$

(iii) The costs rates fulfill $h \geq c$.

Proof For a Gordon–Newell network with local service rates that are increasing and concave in the local queue lengths, the throughput $TH(z)$ is increasing and concave in the population size z (Shanthikumar and Yao, 1988). Hence, it suffices to show that $E_z[X_1]$ is concave in z which is the result of Proposition 1 below. □

Some comments on Theorem 1 and the natural conditions imposed on the production–inventory system are in order here. (a) The natural conditions (i) and (ii) constitute a setting which guarantees that the cost function is well-behaved. The interpretation of these conditions is as follows: The result in Shanthikumar and Yao (1988) indicates that (i) is a minimal condition. It means that higher load enforces the production controller to provide more production capacity until the capacity limit is reached. The meaning of condition (ii) is best understood in case of no rework in the production system (no feedback in queueing network terms), because then $\eta_1 = \eta_2 = 1/2$ holds. In this case, condition (ii) is simply $\mu_2 \leq \mu_1(1)$ and states that the service capacity of the production system is adapted to the arrival stream such that on average there is enough capacity to satisfy the arriving demand of intensity $\lambda = \mu_2$. The lost sales are then due to second order effects: the variability of inter arrival times and service times. In terms of closed network theory: for state-independent service rates at node 1, condition (6) indicates that node 2 is the bottleneck of the system. This interpretation also applies to the general case when using an intuitive but more complex definition of a bottleneck.

(b) Holding costs include all costs involved in storing unused inventory, such as costs for warehouse space and also insurance and opportunity costs. Condition (iii) states that holding costs dominate production costs. This property is discussed in Rubio and Wein (1996)[Section 4.2, p. 266]. Under this condition, they conjecture that for production networks with state-independent service rates μ_j the cost function is quasi-convex, but not convex in general for $h \geq c$. In case of $h = c$, they propose that $C(z)$ is convex.

The result of the next proposition is purely queueing-theoretical because it relies only on the conditions (i) and (ii). Moreover, it seems to be of independent interest because it provides second order properties for the mean queue lengths in a two-station network.

Proposition 1 *Consider the two-node Gordon–Newell network with $z \geq 1$ customers and $\eta = (\eta_1, \eta_2)$ as the solution of the traffic equation. Assume that the service rate at node 2 is constant μ_2 and that the service rate $\mu_1(n)$, $n \in \mathbb{N}$, at node 1 is increasing in n. Let (X_1, X_2) be distributed according to the stationary distribution (7).*
If $\eta_1/\mu_1(1) \leq \eta_2/\mu_2$ holds, then the expected queue length $E_z[X_1]$ at station 1 is a concave function of the population size z.

The proof is postponed to "Appendix 1". Here, we add a remark indicating relations to a well-known fact from queueing theory which has been already used in the literature of production–inventory systems (Karush, 1957). To prove Proposition 1, it will be convenient to rewrite the stationary distribution (2) of the two-station production–inventory system as follows. The stationary distribution π_z of the joint queue lengths vector (X_1, X_2) can be rewritten for $n_1 = n, n_2 = z - n,\ 0 \leq n \leq z$, as

$$\begin{aligned}\pi_z(n, z-n) &= \left(\prod_{k=1}^{n} \frac{\eta_1}{\mu_1(k)}\right)\left(\frac{\eta_2}{\mu_2}\right)^{z-n} \left(\sum_{l=0}^{z}\left(\prod_{k=1}^{l} \frac{\eta_1}{\mu_1(k)}\right)\left(\frac{\eta_2}{\mu_2}\right)^{z-l}\right)^{-1} \\ &= \left(\prod_{k=1}^{n} \frac{\mu_2 \cdot (\eta_1/\eta_2)}{\mu_1(k)}\right) \bar{G}(z, \{1,2\})^{-1}, \end{aligned} \tag{7}$$

with $\bar{G}(z, \{1, 2\}) = \sum_{m=0}^{z} \left(\prod_{k=1}^{m} \frac{\mu_2 \cdot (\eta_1/\eta_2)}{\mu_1(k)} \right)$. Formula (7) indicates that the stationary distribution of the two-station Gordon–Newell network equals the stationary distribution of an exponential single server queue with finite waiting room z, state-dependent service rates $\mu_1(\cdot)$ and arrival intensity $\mu_2 \cdot (\eta_1/\eta_2)$ which satifies the natural condition $\mu_2 \cdot (\eta_1/\eta_2) < \mu_1(1)$. This has been already utilized by Karush (1957) when applying Palm's formula for evaluating the performance of the production–inventory system.

Considering the case of state-independent service rates $\mu_1(\cdot) = \mu_1$ at node 1, (6) reads

$$\frac{\eta_1}{\mu_1} \leq \frac{\eta_2}{\mu_2}, \tag{8}$$

which means that station 2 is the bottleneck of the system.

As a result, Proposition 1 says that the expected queue length at the non-bottleneck station 1 is a concave function of the population size in the network, which is the base stock level in our motivating system. On the other hand, we see that at the bottleneck node 2 the expected queue length $E_z[X_2] = z - E_z[X_1]$ is convex in z. Consequently, if the equality in (8) holds the expected queue length at both stations will be the same and linear in z. This is the case of a "balanced" production–inventory system where convexity of the cost function can be derived analytically by differentiating the cost function. Indeed, this generalizes to the multi-station network (Rubio and Wein, 1996) [Section 4.2].

Example 1 We consider the production–inventory system where an reorder has to undergo only a random lead time which is exponentially distributed with mean ν^{-1} and all lead times are assumed to be independent. This is the first model investigated by Karush (1957) and it can be easily seen that the replenishment network can be modeled as an infinite server queue, i.e. $\mu_1(n) = n \cdot \nu, n = 1, 2, \ldots, z$. Karush proves that the stock-out probability $\pi_z(z, 0)$ is convex in z. This implies that the mean number of lost sales is convex in z. Note that $\mu_1(\cdot)$ is increasing and concave. Thus, Theorem 1 applies to this system as well.

In a second approach, he proves that this property is maintained even under the assumption that the replenishment lead time is a mixture of generalized Erlang distributions, which he realizes as a network of a finite number of different branches of multistage tandem networks, each stage containing an infinite server. As Karush indicates, the class of mixtures of generalized Erlangian distributions is a very versatile class of distributions on $[0, \infty)$. Indeed, this class is dense in the class of all distributions on $[0, \infty)$ (Schassberger, 1973)[Section I. 6].

The important observation is that the stock-out probability does not reflect the complexity of the lead time distribution because it only depends on the mean lead time ν^{-1}. This result dates back to Palm (1938) and is in fact a property of $M/G/z/z$ queues, i.e. the Erlang loss system. Such properties are now labeled as "insensitivity properties" of queueing systems, see e.g. König et al. (1974) and Taylor (2011). For more details on inventory systems see Zipkin (2000)[Chapter 7].

4 Multi-station replenishment network, $K \geq 1$

In this section, we consider the general production–inventory problem as described in Sect. 2 where the system is modeled as a Gordon–Newell network with node set $\overline{J} := \{1, ..., K + 1\}$ and cost function (4). This problem can be reduced to the one investigated in Sect. 3 by exploiting the fact $E_z[X_{K+1}] = z - \sum_{j=1}^{K} E_z[X_j]$ such that equation (4) reduces to

$$C(z) = hz + (c-h)E_z[\sum_{j=1}^{K} X_j] + \ell\,(\lambda - \eta_{K+1} \cdot TH(z))\,. \tag{9}$$

(Recall that in the associated production–inventory system it holds $\mu_{K+1} := \lambda$.)
The main idea is to apply Norton's Theorem for Gordon–Newell networks (Chandy et al., 1975) which says:
For the $K+1$-node Gordon–Newell network of Sect. 2 with z customers, we can construct a reduced two-node Gordon–Newell network consisting of

• a node $\widetilde{K+1}$ which has the same service rate as the node $K+1$ in the original network (in our setting with a constant service rate $\mu_{\widetilde{K+1}} := \mu_{K+1} = \lambda$), and
• a single "composite queue" $\widetilde{1}$ substituting the nodes 1, ..., K from the original network with queue length dependent service rates $\phi(n)$, $n = 0, 1, 2, \ldots$ (to be determined below), such that the steady state marginal distributions of the nodes $K+1$ and $\widetilde{K+1}$ in the two networks are identical although the rest of the network is very different.

This transforms the cost function (9) into

$$\begin{aligned} C(z) &= hz + (c-h)\underbrace{(z - E_z[X_{K+1}])}_{\text{original network}} + \ell\,(\lambda - \eta_{K+1} \cdot TH(z)) \\ &= hz + (c-h)\underbrace{(z - E_z[X_{\widetilde{K+1}}])}_{\text{using Norton's Theorem}} + \ell\,(\lambda - \eta_{K+1} \cdot TH(z)) \\ &= hz + (c-h)\underbrace{E_z[X_{\widetilde{1}}]}_{\text{composite queue } \widetilde{1}} + \ell\,(\lambda - \eta_{K+1} \cdot TH(z))\,, \end{aligned} \tag{10}$$

which is the analog of the cost function (5).
With that, we are prepared to state and prove our main theorem:

Theorem 2 *Consider the production–inventory network with base stock level $z \geq 1$ and with Poisson-λ demand process under lost sales from Sect. 2 and the associated Gordon–Newell network with exponential single server nodes $\overline{J} = \{1, \ldots, K, K+1\}$ and $(\eta_1, \ldots, \eta_{K+1})$ as probability solution of the traffic equation. Node $K+1$ has constant service rate $\mu_{K+1} = \lambda$ while the other nodes may have state dependent service rates $\mu_j(\cdot)$.*
The cost function (10) is convex if the following holds:
$(\widetilde{i}\,)$ The service rate functions $\mu_j(\cdot)$, $j = 1, \ldots, K$ are increasing and concave.
$(\widetilde{ii}\,)$ It holds

$$\sum_{j=1}^{K} \frac{\eta_j}{\mu_j(1)} \leq \frac{\eta_{K+1}}{\mu_{K+1}}. \tag{11}$$

$(\widetilde{iii}\,)$ The costs fulfill $h \geq c$.

Proof The first observation is: from $(\widetilde{i}\,)$ it follows that the throughput $TH(z)$ of the original network in (10) is concave (Shanthikumar and Yao, 1988). Consequently, for $h = c$ the proof is finished. We therefore consider henceforth the case $h > c$.

We will apply Norton's Theorem (Chandy et al., 1975) to the associated Gordon–Newell network from Sect. 2 and obtain a two-node network with node set $\widetilde{J} = \{\widetilde{1}, \widetilde{K+1}\}$ as described above. The service rate at $\widetilde{K+1}$ is $\mu_{\widetilde{K+1}} = \mu_{K+1} = \lambda$ and the routing matrix $\widetilde{R}$

is selected such that the solution of the traffic equation $\widetilde{\eta} \cdot \widetilde{R} = \widetilde{\eta}$ has the probability solution $\widetilde{\eta} = (\widetilde{\eta}_{\widetilde{1}}, \widetilde{\eta}_{\widetilde{K+1}}) = (1 - \eta_{K+1}, \eta_{K+1})$. (This is always possible.)

The service rates $\phi(\cdot)$ at node $\widetilde{1}$ are queue length dependent with $\phi(0) = 0$ and

$$\phi(h) = \frac{\sum_{n_1+\cdots+n_K=h-1} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j/(1-\eta_{K+1})}{\mu_j(k_j)} \right)}{\sum_{n_1+\cdots+n_K=h} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j/(1-\eta_{K+1})}{\mu_j(k_j)} \right)}, \quad h > 0\,. \tag{12}$$

According to Norton's Theorem, the stationary distribution $\widetilde{\pi}_z$ of the substitute network is with normalization constant $\widetilde{G}(z, \{\widetilde{1}, \widetilde{K+1}\})$

$$\widetilde{\pi}_z(n, z-n) = \left(\frac{\eta_{K+1}}{\mu_{K+1}} \right)^n \left(\prod_{h=1}^{z-n} \frac{1 - \eta_{K+1}}{\phi(h)} \right) \cdot \widetilde{G}(z, \{\widetilde{1}, \widetilde{K+1}\})^{-1}, \quad n = 0, 1, \ldots, z. \tag{13}$$

With this reduction, we are now in the framework of Theorem 1. To apply this theorem, we need to show that the following holds:
(a) The marginal steady state queue lengths at node $K+1$ of the original system and node $\widetilde{K+1}$ in the Norton's Theorem substitute are the same.
(b) $(1 - \eta_{K+1})/\mu_{\widetilde{1}}(1) \le \eta_{K+1}/\mu_{\widetilde{K+1}}$. (Analog to (ii) of Theorem 1.)
(c) The service rate function $\phi(\cdot)$ is increasing and concave. (Analog to (i) of Theorem 1.)
If we verify (a),(b), (c), then from Proposition 1 it follows that $E_z[X_{\widetilde{1}}] = E_z[\sum_{j=1}^{K} X_j]$ is concave, and we conclude that $C(z)$ in (10) and thus in (9) is convex.
Verification of (a) The marginal steady state queue length distribution $\widetilde{\pi}_z^{\widetilde{K+1}}$ at node $\widetilde{K+1}$ in the Norton's Theorem substitute is $\widetilde{\pi}_z^{\widetilde{K+1}}(n) = \widetilde{\pi}_z(n, z-n)$ for $n \ge 0$. By using the definition of $\phi(\cdot)$ and simplifying the expression, we obtain from (13)

$$\begin{aligned}
&\widetilde{\pi}_z^{\widetilde{K+1}}(n) \\
&\quad = \left(\prod_{\ell=1}^{n} \frac{\eta_{K+1}}{\mu_{K+1}(\ell)} \right) (1 - \eta_{K+1})^{z-n} \\
&\qquad \sum_{n_1+\cdots+n_K=z-n} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j/(1 - \eta_{K+1})}{\mu_j(k_j)} \right) \widetilde{G}(z, \{\widetilde{1}, \widetilde{K+1}\})^{-1} \\
&\quad = \underbrace{\left(\prod_{\ell=1}^{n} \frac{\eta_{K+1}}{\mu_{K+1}(\ell)} \right) \sum_{n_1+\cdots+n_K=z-n} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j}{\mu_j(k_j)} \right)}_{(\star)} \widetilde{G}(z, \{\widetilde{1}, \widetilde{K+1}\})^{-1}.
\end{aligned}$$

Because $(\star)$ is the non-normalized marginal stationary distribution of node $K+1$ in the original network, we conclude

$$\widetilde{G}(z, \{\widetilde{1}, \widetilde{K+1}\}) = G(z, \{1, \ldots, K, K+1\}),$$

where $G(z, \{1, \ldots, K, K+1\})$ is the normalization constant in (2). Thus, we have proved that the marginal distribution of the nodes $K+1$ and $\widetilde{K+1}$ are the same: $\widetilde{\pi}_z^{\widetilde{K+1}} = \pi_z^{K+1}$.
Verification of (b) By definition, we have

$$\mu_{\widetilde{1}}(1) = \phi(1) = \frac{1}{\sum_{j=1}^{K} \frac{\eta_j/(1-\eta_{K+1})}{\mu_j(1)}} = \frac{1 - \eta_{K+1}}{\sum_{j=1}^{K} \frac{\eta_j}{\mu_j(1)}}.$$

So,

$$\frac{1-\eta_{K+1}}{\phi(1)} = \sum_{j=1}^{K} \frac{\eta_j}{\mu_j(1)} \overset{(\widetilde{ii})}{\leq} \frac{\eta_{K+1}}{\mu_{K+1}}$$

completes this part of the verification.

Verification of (c) According to Norton's Theorem, the service rate function at the composite queue $\widetilde{1}$ is $\phi(0) = 0$ and for $h > 0$

$$\phi(h) = \frac{\sum_{n_1+\cdots+n_K=h-1} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j/(1-\eta_{K+1})}{\mu_j(k_j)}\right)}{\sum_{n_1+\cdots+n_K=h} \prod_{j=1}^{K} \left(\prod_{k_j=1}^{n_j} \frac{\eta_j/(1-\eta_{K+1})}{\mu_j(k_j)}\right)}.$$

A short reflection reveals that this is the throughput in a Gordon–Newell network with node set {1,...,K}, a routing (which can be selected irreducible) with stationary distribution $(\eta_j / \sum_{i=1}^{K} \eta_i \,:\, j = 1, \ldots, K)$ of the routing process (solution of the relevant traffic equation), and queue length dependent service rates $\mu_j(\cdot)$.

Because these service rate functions are by $(\widetilde{ii})$ increasing and concave, $\phi(\cdot)$ is increasing and concave (Shanthikumar and Yao, 1988). This verifies (c). □

Example 2 In Example 1, we have mentioned that for the case of processing networks consisting of a single $\cdot/M/\infty$ system which represents independent identically distributed lead times the result of Palm (1938) states that the stock-out probability is invariant under changing the shape of the lead time distribution as long as the mean ν^{-1} remains fixed. Having the K-station replenishment now at our disposal, we can reproduce the invariance result for the stock-out probability immediately for networks of infinite server queues and suitably defined Markovian random routing of the reorders before the lead time expires.

5 Numerical experiments

We have performed a series of numerical experiments, mainly to investigate the structure of cost functions of production–inventory systems which are not covered by our theoretical results. For our computations, we only consider queueing networks without feedback, i.e. $\eta_1 = \eta_2 = 1/2$. We also set the lost sales costs $\ell = 30$ to clearly illustrate the convexity of the expected total costs function in the figures.

First, we consider a closed Gordon Newell network with $\bar{J} = 1, 2$ we have dealt with in Theorem 1. Taking parameters from Rubio and Wein, we set the WIP cost per item $c = 1$, and the finished goods inventory cost per item $h = 2$ (Rubio and Wein, 1996)[Section 3, p. 265]. The inventory (node 2) is the bottleneck of the system and generates higher costs than production at station $j = 1$ does (holding costs dominate production costs). As anticipated, the expected total costs are convex for state dependent increasing and concave service rates at station $j = 1$ with $\frac{\eta_1}{\mu_1(1)} \leq \frac{\eta_2}{\mu_2}$ and $h > c$ (see Fig. 2 for a typical behavior of the cost function). Although the lost sales cost is rather high, the optimal base stock level is astonishingly low. This is explicable by the fact that node 2 (the inventory) is the bottleneck of the system: On average, there is more replenishment capacity available than demand arrives. In Fig. 3, we can observe the concavity of the expected queue length at station $j = 1$ and the convexity of the expected queue length at station $j = 2$, proved in Proposition 1.

Secondly, we investigate the same system in more detail when the production costs (at station 1) dominate the holding costs at the bottleneck (the inventory). Surprisingly, in these

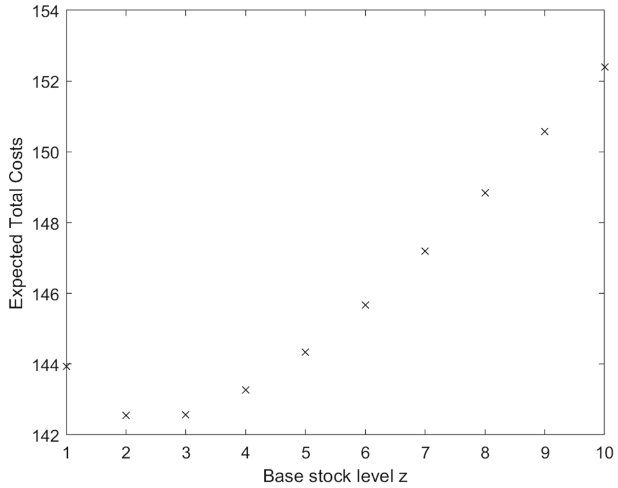

Fig. 2 Expected total costs: $\mu_1(n) = \ln(n) + \mu_2 + 0.1$, $\mu_2 = 5$, $\eta_1 = \eta_2 = 0.5$, $c = 1 < h = 2$, $\ell = 30$

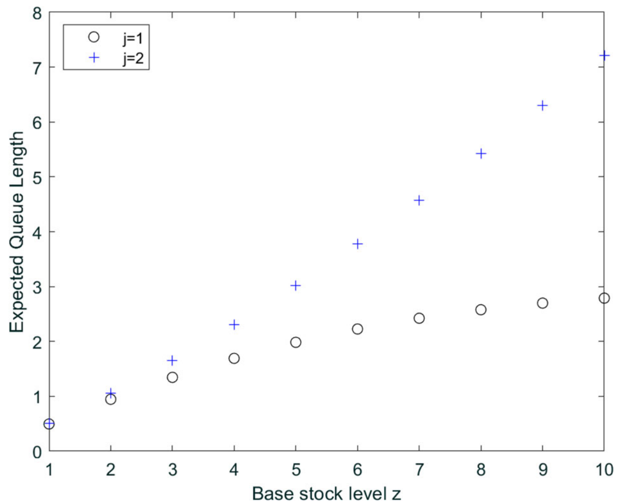

Fig. 3 Expected queue lengths (parameters as in Fig. 2)

experiments the expected total costs are convex as long as the cost ratio $\frac{c}{h}$ is moderately higher than 1 (see Fig. 4 with $\frac{c}{h} \leq 6$). In these cases, the optimal base stock level z is decreasing in c. But for large cost ratios of $\frac{c}{h}$, the expected total cost function becomes increasing and concave in z (see Fig. 5 with $\frac{c}{h} \geq 20$). In this situation, the optimal base stock level is $z = 1$ in all cases we have investigated.

For completeness, we consider an example where station 1 with state-independent reproduction rate $\mu_1 < \mu_2$ is the bottleneck and the production costs dominate the holding costs ($\frac{c}{h} > 1$). As it can be seen in Fig. 6, the expected total costs are convex for this case. Furthermore, the expected queue length at station $j = 1$ is convex and the expected queue length at station $j = 2$ is concave (see Fig. 7). Note that this result is covered by Theorem 1 as well (by renumbering the stations).

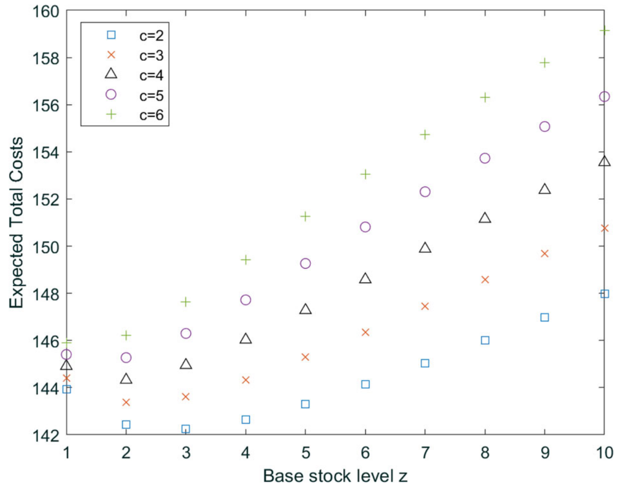

Fig. 4 Expected total costs: $\mu_1(n) = \ln(n) + \mu_2 + 0.1$, $\mu_2 = 5$, $\eta_1 = \eta_2 = 0.5$, $h = 1$, $\ell = 30$

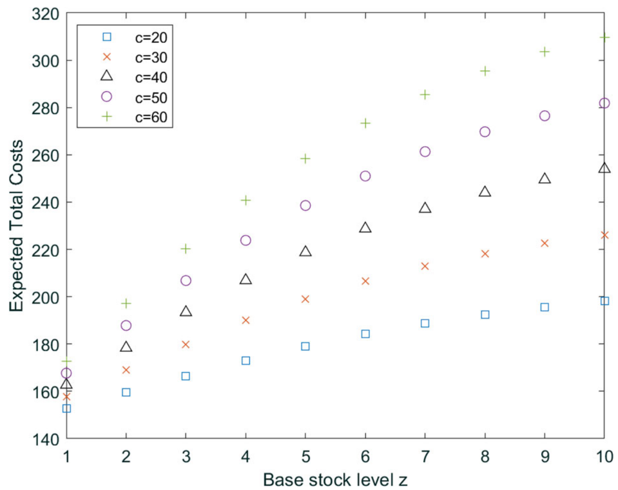

Fig. 5 Expected total costs: $\mu_1(n) = \ln(n) + \mu_2 + 0.1$, $\mu_2 = 5$, $\eta_1 = \eta_2 = 0.5$, $h = 1$, $\ell = 30$

6 Conclusion

For a production–inventory system with a complex processing network as replenishment system, we have proved convexity of a standard cost function. Although the system could be transformed into a standard closed Gordon–Newell network, which is in general easy to understand with respect to steady state distribution and expected queue lengths (first order properties), we have demonstrated that proving the second order properties of interest is quite challenging. Our numerical results clearly indicate that further research in the field of replenishment systems with state dependent service rates is needed. Our ongoing research encompasses the case of cost functions where production costs dominate holding costs and especially proving similar properties for production–inventory systems with positive dis-

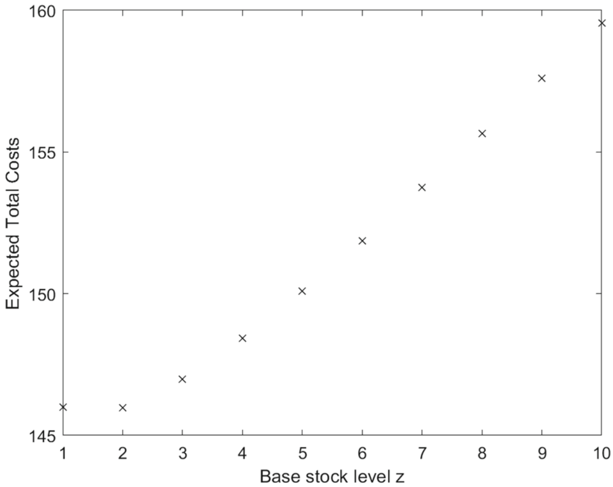

Fig. 6 Expected total costs: $\mu_1 = 3 < \mu_2 = 5, \eta_1 = \eta_2 = 0.5, c = 2 > h = 1, \ell = 30$

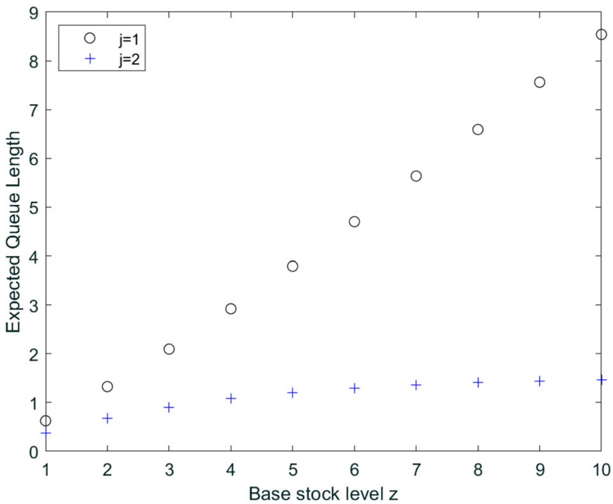

Fig. 7 Expected queue lengths (parameters as in Fig. 6)

patching times at the inventory. The case of multi-stage inventories under base-stock policy at every stage is another direction of our research.

Acknowledgements We thank the three reviewers for their helpful comments on the first version of the paper which helped us to improve the article.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interst N. de la Cruz and H. Daduna declares that he has no conflict of interest.

A Technical lemmata and omitted proofs

For the following calculations, we use the abbreviations $\frac{\mu_2\cdot(\eta_1/\eta_2)}{\mu_1(n)} =: x_n, n \in \mathbb{N}$ and $\bar{G}(z, \{1, 2\}) = \sum_{m=0}^{z} \left(\prod_{k=1}^{m} x_k\right) =: \bar{G}(z, 2)$ which yields:

$$\pi_z(n, z-n) = \left(\prod_{k=1}^{n} x_k\right) \bar{G}(z, 2)^{-1} \tag{14}$$

$$\bar{G}(z+\ell, 2) = \bar{G}(z, 2) + \sum_{h=1}^{\ell} \left(\prod_{k=1}^{z+h} x_k\right) \tag{15}$$

$$1 \geq x_1 \geq x_2 \geq \cdots \geq x_n \geq \cdots > 0 \tag{16}$$

We need two preparatory lemmata to finally prove Proposition 1. For the proof of Proposition 1 and also for the proofs of Lemma (1) and (2), we only utilize (14)-(16).

Lemma 1 and its proof

Lemma 1 *For the two-node network of Theorem 1, the following inequalities are equivalent:*

$$E_{z+2}[X_1] + E_z[X_1] \leq 2E_{z+1}[X_1]$$

$$\sum_{i=0}^{z} \bar{G}(i, 2) \left((1 - x_{z+2})\bar{G}(z, 2) + \prod_{k=1}^{z+1} x_k\right) + \sum_{i=0}^{z-1} \bar{G}(i, 2) \prod_{k=1}^{z+2} x_k \geq x_{z+2}\bar{G}(z, 2)\bar{G}(z, 2)$$

Proof

$$E_{z+2}[X_1] + E_z[X_1] \leq 2E_{z+1}[X_1]$$

$$\overset{(14)}{\iff} \sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k\right) \bar{G}(z+2, 2)^{-1} + \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k\right) \bar{G}(z, 2)^{-1}$$

$$\leq 2 \sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k\right) \bar{G}(z+1, 2)^{-1}$$

$$\iff \left(\sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k\right)\right) \bar{G}(z+1, 2)\bar{G}(z, 2) + \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k\right)\right) \bar{G}(z+2, 2)\bar{G}(z+1, 2)$$

$$\leq 2 \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k\right)\right) \bar{G}(z+2, 2)\bar{G}(z, 2)$$

$$
\begin{aligned}
&\overset{(15)}{\Longleftrightarrow} \bar{G}(z,2)^2 \left[\sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k \right) + \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right] \\
&\quad + \bar{G}(z,2) \left[\left(\sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k + \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(2 \prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \right] \\
&\quad + \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \prod_{k=1}^{z+1} x_k \\
&\quad \le 2 \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\bar{G}(z,2)^2 + \bar{G}(z,2) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \right) \\
&\Longleftrightarrow \bar{G}(z,2)^2 \left[2 \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) - \sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k \right) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right] \\
&\quad + \bar{G}(z,2) \left[2 \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) - \left(\sum_{m=1}^{z+2} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k \right. \\
&\quad \left. - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(2 \prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \right] \\
&\quad - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \prod_{k=1}^{z+1} x_k \ge 0 \\
&\Longleftrightarrow \bar{G}(z,2)^2 \left[\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) - \sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) - (z+2) \prod_{k=1}^{z+2} x_k \right. \\
&\quad \left. + (z+1) \prod_{k=1}^{z+1} x_k + \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right] \\
&\quad + \bar{G}(z,2) \left[(z+1) \prod_{k=1}^{z+1} x_k \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) + \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \right. \\
&\quad - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k \\
&\quad + \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k - \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k \\
&\quad \left. - (z+2) \prod_{k=1}^{z+1} x_k \prod_{k=1}^{z+2} x_k + \left(\sum_{m=1}^{z+1} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k \right] \\
&\quad - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \prod_{k=1}^{z+1} x_k \ge 0 \\
&\Longleftrightarrow \bar{G}(z,2) \left[\bar{G}(z,2) \left((z+1) \prod_{k=1}^{z+1} x_k - (z+2) \prod_{k=1}^{z+2} x_k \right) + (z+1) \prod_{k=1}^{z+1} x_k \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \right. \\
&\quad \left. - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k + \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - \prod_{k=1}^{z+1} x_k \prod_{k=1}^{z+2} x_k \right]
\end{aligned}
$$

$$- \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \underbrace{\prod_{k=1}^{z+1} x_k}_{>0 \text{ by } (16)} \geq 0$$

$$\iff \bar{G}(z,2) \left[\bar{G}(z,2) \left(z+1-(z+2)x_{z+2} \right) + (z+1) \prod_{k=1}^{z+1} x_k + z \prod_{k=1}^{z+2} x_k - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right.$$
$$\left. + \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) x_{z+2} \right] - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \geq 0$$

$$\iff \bar{G}(z,2) \left[\bar{G}(z,2) \left((1-x_{z+2})(z+1) - x_{z+2} \right) + \prod_{k=1}^{z+1} x_k - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) (1-x_{z+2}) \right]$$
$$+ \left(z\bar{G}(z,2) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left(\prod_{k=1}^{z+1} x_k + \prod_{k=1}^{z+2} x_k \right) \geq 0$$

$$\iff \bar{G}(z,2)\bar{G}(z,2)(1-x_{z+2})(z+1) + \bar{G}(z,2) \prod_{k=1}^{z+1} x_k + z\bar{G}(z,2) \prod_{k=1}^{z+1} x_k$$
$$- \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) (1-x_{z+2})\bar{G}(z,2) - \left(\sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+1} x_k$$
$$+ \left(z\bar{G}(z,2) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \geq 0$$

$$\iff \left((z+1)\bar{G}(z,2) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left((1-x_{z+2})\bar{G}(z,2) + \prod_{k=1}^{z+1} x_k \right)$$
$$+ \left(z\bar{G}(z,2) - \sum_{m=1}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \geq 0$$

$$\iff \left(\bar{G}(z,2) + z \sum_{m=0}^{z} \left(\prod_{k=1}^{m} x_k \right) - \sum_{m=0}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \left((1-x_{z+2})\bar{G}(z,2) + \prod_{k=1}^{z+1} x_k \right)$$
$$+ \left(z \sum_{m=0}^{z} \left(\prod_{k=1}^{m} x_k \right) - \sum_{m=0}^{z} m \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \geq 0$$

$$\iff \left(\bar{G}(z,2) + \sum_{m=0}^{z} (z-m) \left(\prod_{k=1}^{m} x_k \right) \right) \left((1-x_{z+2})\bar{G}(z,2) + \prod_{k=1}^{z+1} x_k \right)$$
$$+ \left(\sum_{m=0}^{z} (z-m) \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \geq 0$$

$$\iff \left(\bar{G}(z,2) + \sum_{i=0}^{z-1} \sum_{m=0}^{i} \left(\prod_{k=1}^{m} x_k \right) \right) \left((1-x_{z+2})\bar{G}(z,2) + \prod_{k=1}^{z+1} x_k \right)$$
$$+ \left(\sum_{i=0}^{z-1} \sum_{m=0}^{i} \left(\prod_{k=1}^{m} x_k \right) \right) \prod_{k=1}^{z+2} x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \geq 0$$

$$\iff \sum_{i=0}^{z}\bar{G}(i,2)\left((1-x_{z+2})\bar{G}(z,2)+\prod_{k=1}^{z+1}x_k\right)+\sum_{i=0}^{z-1}\bar{G}(i,2)\prod_{k=1}^{z+2}x_k - x_{z+2}\bar{G}(z,2)\bar{G}(z,2)\geq 0$$

□

Lemma 2 and its proof

Lemma 2 *It holds for $z\in\mathbb{N}$:*

$$\bar{G}(z,2)\bar{G}(z-1,2)=\sum_{j=0}^{z-1}\bar{G}(j,2)\left[\prod_{k=1}^{j+1}x_k+\prod_{k=1}^{j}x_k\right] \tag{17}$$

Proof The proof is by induction with $A(z)$ denoting the statement (17).
Base case $(A(1))$:

$$\bar{G}(1,2)\bar{G}(0,2)=\left(\sum_{m=0}^{1}\left(\prod_{k=1}^{m}x_k\right)\right)1=(1+x_1)1=\sum_{j=0}^{0}\bar{G}(j,2)\left[\prod_{k=1}^{j+1}x_k+\prod_{k=1}^{j}x_k\right]$$

Thus, $A(1)$ is true.
Induction step $(A(z)\Rightarrow A(z+1))$: For $z+1\in\mathbb{N}$ holds

$$\begin{aligned}\bar{G}(z+1,2)\bar{G}(z,2)&=\bar{G}(z,2)\left[\bar{G}(z-1,2)+\prod_{k=1}^{z}x_k+\prod_{k=1}^{z+1}x_k\right]\\&=\bar{G}(z,2)\left[\prod_{k=1}^{z}x_k+\prod_{k=1}^{z+1}x_k\right]+\bar{G}(z,2)\bar{G}(z-1,2)\\&\overset{A(z)}{=}\bar{G}(z,2)\left[\prod_{k=1}^{z}x_k+\prod_{k=1}^{z+1}x_k\right]+\sum_{j=0}^{z-1}\bar{G}(j,2)\left[\prod_{k=1}^{j+1}x_k+\prod_{k=1}^{j}x_k\right]\\&=\sum_{j=0}^{z}\bar{G}(j,2)\left[\prod_{k=1}^{j+1}x_k+\prod_{k=1}^{j}x_k\right]\end{aligned}$$

□

Proof of Proposition 1

The expected queue length at station 1 is concave if $E_{z+2}[X_1]+E_z[X_1]\leq 2E_{z+1}[X_1]$. By Lemma 1, it suffices to show that the relation $B(z):\iff$

$$\sum_{i=0}^{z}\bar{G}(i,2)\left((1-x_{z+2})\bar{G}(z,2)+\prod_{k=1}^{z+1}x_k\right)+\sum_{i=0}^{z-1}\bar{G}(i,2)\prod_{k=1}^{z+2}x_k\geq x_{z+2}\bar{G}(z,2)\bar{G}(z,2) \tag{18}$$

holds for $z\in\mathbb{N}_0$. The proof is by induction.
Base case $(B(0))$: Since we have $1\geq x_1\geq x_2$, it holds $(1-x_2)+x_1\geq x_2$. Thus, $B(0)$ is true.

Induction step $(B(z) \Rightarrow (B(z+1))$: Using Lemma 2 and the induction hypothesis, it follows for the left side of $B(z+1)$:

$$
\begin{aligned}
&\sum_{i=0}^{z+1} \bar{G}(i,2)\left((1-x_{z+3})\bar{G}(z+1,2)+\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad=\sum_{i=0}^{z+1} \bar{G}(i,2)(1-x_{z+3})\bar{G}(z+1,2)+\sum_{i=0}^{z+1} \bar{G}(i,2)\left(\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad=\sum_{i=0}^{z} \bar{G}(i,2)(1-x_{z+3})\bar{G}(z+1,2)+\bar{G}(z+1,2)\bar{G}(z+1,2)(1-x_{z+3}) \\
&\qquad+\sum_{i=0}^{z-1} \bar{G}(i,2)\left(\prod_{k=1}^{z+2} x_k\right)+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad=\sum_{i=0}^{z} \bar{G}(i,2)\left((1-x_{z+3})\bar{G}(z,2)+(1-x_{z+3})\left(\prod_{k=1}^{z+1} x_k\right)\right)+\bar{G}(z+1,2)\bar{G}(z+1,2)(1-x_{z+3}) \\
&\qquad+\sum_{i=0}^{z-1} \bar{G}(i,2)\left(\prod_{k=1}^{z+2} x_k\right)+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad=\sum_{i=0}^{z} \bar{G}(i,2)\left(\underbrace{(1-x_{z+3})}_{\geq 1-x_{z+2}}\bar{G}(z,2)+\prod_{k=1}^{z+1} x_k\right)-x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z} \bar{G}(i,2) \\
&\qquad+\bar{G}(z+1,2)\bar{G}(z+1,2)(1-x_{z+3}) \\
&\qquad+\sum_{i=0}^{z-1} \bar{G}(i,2)\left(\prod_{k=1}^{z+2} x_k\right)+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad\geq \underbrace{\sum_{i=0}^{z} \bar{G}(i,2)\left((1-x_{z+2})\bar{G}(z,2)+\prod_{k=1}^{z+1} x_k\right)+\sum_{i=0}^{z-1} \bar{G}(i,2)\left(\prod_{k=1}^{z+2} x_k\right)}_{\text{left side of } B(z)}-x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z} \bar{G}(i,2) \\
&\qquad+\bar{G}(z+1,2)\bar{G}(z+1,2)(1-x_{z+3})+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(\prod_{k=1}^{z+2} x_k\right)+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad\geq \underbrace{\overbrace{x_{z+2}}^{\geq x_{z+3}}\bar{G}(z,2)\bar{G}(z,2)}_{\text{right side of } B(z)}+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(\overbrace{x_{z+2}}^{\geq x_{z+3}}\prod_{k=1}^{z+1} x_k\right)-x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z} \bar{G}(i,2) \\
&\qquad+\bar{G}(z+1,2)\bar{G}(z+1,2)(1-x_{z+3})+\sum_{i=0}^{z} \bar{G}(i,2)\prod_{k=1}^{z+3} x_k \\
&\quad\geq x_{z+3}\bar{G}(z,2)\bar{G}(z,2)+(\bar{G}(z,2)+\bar{G}(z+1,2))\left(x_{z+3}\prod_{k=1}^{z+1} x_k\right) \\
&\qquad+\left(\left(\prod_{k=1}^{z+1} x_k\right)\bar{G}(z+1,2)+\underbrace{\bar{G}(z,2)\bar{G}(z+1,2)}_{\text{apply } Lemma\ 2}\right)(1-x_{z+3})
\end{aligned}
$$

$$
\begin{aligned}
&- x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z}\bar{G}(i,2)(1-x_{z+2})\\
&= x_{z+3}\left(\bar{G}(z,2)\bar{G}(z+1,2)+\left(\prod_{k=1}^{z+1} x_k\right)\bar{G}(z+1,2)\right)\\
&\quad+\left(\left(\prod_{k=1}^{z+1} x_k\right)\bar{G}(z+1,2)+\sum_{i=0}^{z}\bar{G}(i,2)\left(\prod_{k=1}^{i+1} x_k+\prod_{k=1}^{i} x_k\right)\right)(1-x_{z+3})\\
&\quad- x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z}\bar{G}(i,2)(1-x_{z+2})\\
&= x_{z+3}\bar{G}(z+1,2)\bar{G}(z+1,2)\\
&\quad+\left(\sum_{i=0}^{z+1}\bar{G}(i,2)\left(\prod_{k=1}^{i} x_k\right)+\underbrace{\sum_{i=0}^{z}\bar{G}(i,2)\left(\prod_{k=1}^{i+1} x_k\right)}_{(\star)}\right)(1-x_{z+3})\\
&\quad\underbrace{- x_{z+3}\left(\prod_{k=1}^{z+1} x_k\right)\sum_{i=0}^{z}\bar{G}(i,2)}_{(\star\star)}(1-x_{z+2}) \overset{(\star)\geq(\star\star)}{\geq} x_{z+3}\bar{G}(z+1,2)\bar{G}(z+1,2)
\end{aligned}
$$

This proves $B(z+1)$ and thus the expected queue length at node 1 is concave in the population size z. □

References

Atan, Z., Ahmadi, T., Stegehuis, C., de Kok, T., & Adan, I. (2017). Assemble-to-order systems: A review. *European Journal of Operational Research, 261*(3), 866–879.

Axsäter, S. (1993). Continuous review policies for multi-level inventory systems with stochastic demand. In S. Graves, A. Rinnooy Kan, & P. Zipkin (Eds.), *Logistics of production and inventory, Handbooks in operations research and management science* (Chap. 4, Vol. 4, pp. 175–198). North-Holland.

Bijvank, M., & Vis, F. (2011). Lost sales inventory theory: A review. *European Journal of Operational Research, 215,* 1–13.

Bijvank, M., & Vis, F. (2012). Lost sales inventory systems with a service level criterion. *European Journal of Operational Research, 220,* 610–618.

Bradley, J., & Glynn, P. W. (2002). Managing capacity and inventory jointly in manufacturing systems. *Management Science, 48*(2), 273–288.

Chandy, K. M., Herzog, U., & Woo, L. (1975). Parametric analysis of queuing networks. *IBM Journal of Research and Development, 19*(1), 36–42.

Chen, H., & Yao, D. D. (2001). *Fundamentals of queueing networks: Performance, asymptotics, and optimization. Applications of Mathematics* (Vol. 4). Springer.

de la Cruz, N. N., & Daduna, H. (2016). Monotonicity of base stock policies. *Operations Research Letters, 44*(2), 186–190.

de la Cruz, N. N., & Daduna, H. (2019). Optimal capacity allocation in a production-inventory system with base stock. *Annals of Operations Research, 277*(2), 329–344.

Federgruen, A., & Zipkin, P. (1986). An inventory model with limited production capacity and uncertain demand, I: The average cost citerion. *Mathematics of Operations Research, 11,* 193–207.

Federgruen, A., & Zipkin, P. (1986). An inventory model with limited production capacity and uncertain demand, II: The discounted cost citerion. *Mathematics of Operations Research, 11,* 208–215.

Gordon, W. J., & Newell, G. F. (1967). Closed queuing systems with exponential servers. *Operations Research, 15*(2), 254–265.

Hill, R. (1999). On the suboptimality of (S-1, S) lost sales inventory policies. *International Journal of Production Economics, 59,* 387–393.
Hu, J. Q., & Xiang, D. (1993). The queueing equivalence to a manufacturing system with failures. *IEEE Transactions on Automatic Control, 38*(3), 499–502.
Janakiraman, G., & Roundy, R. (2004). Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Operations Research, 52*(5), 795–803.
Johansen, S. (2013). Modified base-stock policies for continuous-review, lost-sales inventory models with Poisson demand and a fixed lead time. *International Journal of Production Economics, 143,* 379–384.
Kaplan, R. S. (1970). A dynamic inventory model with stochastic lead times. *Management Science, 16*(7), 491–507.
Karmarkar, U. S. (1987). Lot sizes, lead times and in-process inventories. *Management Science, 33*(3), 409–418.
Karmarkar, U. S. (1993). Manufacturing lead times, order release and capacity loading. In S. Graves, A. Rinnooy Kan, & P. Zipkin (Eds.), *Logistics of production and inventory, Handbooks in operations research and management science* (Chap. 6, Vol. 4, pp. 287–329). North-Holland.
Karush, W. (1957). A queuing model for an inventory problem. *Operations Research, 5*(5), 693–703.
König, D., Matthes, K., & Nawrotzki, K. (1974). Unempfindlichkeitseigenschaften von Bedienungsprozessen. In B. I. Gnedenko & N. Kowalenko (Eds.), *Einführung in die Bedienungstheorie* (pp. 358–450). Akademie-Verlag.
Krishnamoorthy, A., Lakshmy, B., & Manikandan, R. (2011). A survey on inventory models with positive service time. *Opsearch, 48*(2), 153–169.
Krishnamoorthy, A., Shajin, D., & Narayanan, W. (2021). Inventory with positive service time: A survey. In V. Anisimov & N. Limnios (Eds.), *Advanced trends in queueing theory Series of books "Mathematics and Statistics" sciences* (Vol. 2, pp. 201–238). ISTE & Wiley.
Krishnamoorthy, A., & Viswanath, N. C. (2013). Stochastic decomposition in production inventory with service time. *European Journal of Operational Research, 228*(2), 358–366.
Lee, H., & Nahmias, S. (1993). Single-product, single-location models. In S. Graves, A. Rinnooy Kan, & P. Zipkin (Eds.), *Logistics of production and inventory, Handbooks in operations research and management science* (Chap. 1., Vol. 4, pp. 3–58). North-Holland.
Lee, Y. J., & Zipkin, P. (1992). Tandem queues with planned inventories. *Operations Research, 40*(5), 936–947.
Lee, Y. J., & Zipkin, P. (1995). Processing networks with inventories: Sequential refinement systems. *Operations Research, 43*(6), 1025–1036.
Li, B., & Arreola-Risa, A. (2021). On minimizing downside risk in make-to-stock, risk-averse firms. *Naval Research Logistics (NRL), 68*(2), 199–213.
Li, B., Ji, Q., & Arreola-Risa, A. (2020). Optimizing a production-inventory system under a cost target. *Computers and Operations Research, 123,* 105015.
Liu, B., & Cao, J. (1997). Production control of an unreliable manufacturing system under the assumption of no backlog. *Mathematical Methods of Operations Research, 46*(1), 103–117.
Morse, P. (1958). *Queues, inventories and maintenance*. Wiley.
Otten, S., Krenzler, R., & Daduna, H. (2016). Models for integrated production-inventory systems: Steady state and cost analysis. *International Journal of Production Research, 54*(20), 6174–6191.
Palm, C. (1938). Analysis of the Erlang traffic formula for busy-signal arrangements. *Ericsson Technics, 5*(9), 39–58.
Rubio, R., & Wein, L. M. (1996). Setting base stock levels using product-form queueing networks. *Management Science, 42*(2), 259–268.
Schassberger, R. (1973). *Warteschlangen*. Springer.
Schwarz, M., Sauer, C., Daduna, H., Kulik, R., & Szekli, R. (2006). M/M/1 queueing systems with inventory. *Queueing Systems, 54*(1), 55–78.
Schwarz, M., Wichelhaus, C., & Daduna, H. (2007). Product form models for queueing networks with an inventory. *Stochastic Models, 23*(4), 627–663.
Shanthikumar, J. G., & Yao, D. D. (1988). Second-order properties of the throughput of a closed queueing network. *Mathematics of Operations Research, 13*(3), 524–534.
Song, J. S., & Zipkin, P. (2009). Inventories with multiple supply sources and networks of queues with overflow bypass. *Management Science, 55*(3), 362–372.
Taylor, P. (2011). Intensitivity in stochastic models. International series in operations research and management science. In R. Boucherie & N. van Dijk (Eds.), *Queueing networks. A fundamental approach* (Chap. 3, Vol. 154, pp. 121–140). Springer.
Toktay, L. B., Wein, L. M., & Zenios, S. A. (2000). Inventory management of remanufacturable products. *Management Science, 46*(11), 1412–1426.
Zipkin, P. H. (2000). *Foundations of inventory management*. Mc Graw-Hill.

Zipkin, P. (2008). Old and new methods for lost-sales inventory systems. *Operations Research, 56*(5), 1256–63.