Newsvendor equations for production networks

Citation for published version (APA):

Document license:
TAVERNE

DOI:
10.1016/j.orl.2018.10.006

Document status and date:
Published: 01/11/2018

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 08. Jun. 2022
Newsvendor equations for production networks

S.W.F. Jansen *, Z. Atan, I.J.B.F. Adan, A.G. de Kok

Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Netherlands

A R T I C L E   I N F O

Article history:
Received 22 May 2018
Received in revised form 20 October 2018
Accepted 20 October 2018
Available online 29 October 2018

Keywords:
Production networks
Leadtime uncertainty
Newsvendor equations

A B S T R A C T

We consider production networks with stochastic activity leadtimes. When activities finish early holding costs are incurred and when end products are delivered late penalty costs are incurred. Objective is to find the activity start and finish times that minimize the total cost. We introduce the concept of a tardy path and derive the optimality equations for each node in the network. We show that under the optimal solution, for a set of nodes the tardiness probability satisfies the Newsvendor equations.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

We consider production systems that manufacture high-value, low-volume, customer-specific products, such as airplanes and lithography systems. The manufacturing process of such products is divided into multiple activities, which can only start when preceding activities are finished. High product complexity implies variability in activity leadtimes. In turn, random activity leadtimes imply randomness in the end product completion times and randomness in the customer delivery date. Several studies considered these production systems with the objective of determining the planned leadtimes of the activities. They addressed network structures such as two nodes [7], serial networks [3] and converging networks [1,2,5,6]. A common assumption in all these studies is that the network has a single final node, i.e., there is only one end product.

In this paper, we aim to generalize the results obtained so far to networks that have multiple end nodes. This generalization is motivated by practice. For capital-intensive expensive products, customers pay in several installments, where each installment is linked to the completion of an activity. For example, when building a house, the first installment is often paid when the foundation is finished. An example of a network representation for such a situation is depicted in Fig. 1(a). In this example, the customer pays installments after completion of activities represented by nodes 5, 3 and 1.

It is common practice that products for different customers are produced simultaneously. This is another motivation for our study. In fact, it can be beneficial to align identical activities for different products. For example, if the hull of a new vessel is built in Asia and the rest of the activities are performed in Europe, it makes sense to transport multiple hulls on a big vessel from Asia to Europe. Transporting an extra hull on the same vessel only marginally increases the total cost, while it can be very expensive and inefficient to transport each hull individually. An example of a network representation for such a system is depicted in Fig. 1(b). Node 10 is the shared transportation activity. After this activity, each project is completed separately. In previous studies, different cost structures have been considered, but they all rely on the same basic principle: each activity in the network adds value to the final product. This value is represented by the activity holding cost. If the production of an end product is finished earlier than planned, a holding cost is incurred until the due date. If the product is completed later than planned, an extra penalty cost needs to be paid to the customer. We extend this cost concept to networks with multiple end nodes. Since an activity can add value to multiple end products, we define holding cost for each of the end nodes relevant to the activity. Similar to networks with a single end node, there is a penalty cost for late completions of each end node. Our objective is to set the start times of all activities so that the total expected holding and penalty costs are minimized.

Previous studies determine activity planned leadtimes. However, planned leadtime solutions are ambiguous in networks with multiple end nodes. Therefore, the perspective in this paper is different: we describe the solution in terms of a planned start time for each node and a planned finish time for each end node. This formulation easily applies to networks with single or multiple end nodes.

We also introduce the concept of tardy paths. For a specific realization of activity leadtimes, a path is tardy if it leads to an end node that is behind schedule. This is different from the concept of critical paths introduced by [4]. While a critical path is solely based on planned leadtimes, a tardy path depends on the whole
plan, i.e. start and finish times, and also on the realization of the random leadtimes.

For an assembly system with a single end node, [5] show that under the optimal solution the probability of an activity delaying final delivery is proportional to the value it adds to the end product. In this paper, for given start and finish times and leadtime distributions, we derive an expression for the probability of a path being tardy under the optimal solution. We show that Newsvendor final delivery is proportional to the value it adds to the end product, i.e. start and finish times, and also on the realization of the random leadtimes.

The remainder of this paper is organized as follows. We formulate the model in Section 2. In Section 3, we introduce the concept of tardy paths. Structural results are presented in Section 4. We provide a numerical example in Section 5 and conclude in Section 6.

2. Model formulation

Let $G = (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph with $N$ nodes. $\mathcal{V} = \{1, 2, \ldots, N\}$ is the set of all nodes and $\mathcal{E}$ is the set of all directed edges. Each node represents an activity and each edge represents a precedence relation. The edge from node $i$ to node $j$ is denoted by $(i, j)$. We define $\mathcal{P}(i) := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ and $\mathcal{S}(i) := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$ as the sets of immediate predecessors and immediate successors of node $i$, respectively. In addition, $\mathcal{R} := \{i \in \mathcal{V} : \mathcal{P}(i) = \emptyset\}$ and $\mathcal{L} := \{i \in \mathcal{V} : \mathcal{S}(i) = \emptyset\}$ are the sets of root and leaf nodes, respectively. The set of all nodes that can reach node $i$ is denoted by $\mathcal{Y}(i)$. This set can be recursively defined as $\mathcal{Y}(i) := \{i\} \cup \mathcal{P}(i) \cup \bigcup_{j \in \mathcal{P}(i)} \mathcal{Y}(j)$, where $\mathcal{Y}(i) = \{i\}$ for all $i \in \mathcal{R}$. The set of all nodes that are reachable from node $i$ is $\mathcal{Z}(i)$, which can be defined as $\mathcal{Z}(i) := \{i\} \cup \mathcal{S}(i) \cup \bigcup_{j \in \mathcal{S}(i)} \mathcal{Z}(j)$, where $\mathcal{Z}(i) = \{i\}$ for all $i \in \mathcal{L}$. We assume that nodes are numbered such that for any two nodes, if $i \in \mathcal{Y}(j)$ then $i \geq j$. Finally, we define $\mathcal{W}(i, j)$ as the set of all paths from node $i \in \mathcal{V}$ to $j \in \mathcal{L}$. Note that $\mathcal{W}(i, j)$ might be empty.

Activities in the network have uncertain durations. The duration of the activity in node $i$ is modeled by a random variable $T_i$ that is continuous and exists on the entire domain $(0, \infty)$. Random variables can be dependent, however, the joint probability distribution of all random leadtimes should be continuous and it should exist on the domain $\mathbb{R}_+^N$. The planned start time of node $i$ is $t_i$. An activity can only start after all its predecessors are finished. If all predecessors of $i$ finish earlier than $t_i$, node $i$ starts at the planned start time. Due to the uncertainty in duration of predecessors, the actual start times of non-root nodes are uncertain. The random variable $A_i$ represents this actual start time and is defined as follows:

$$A_i = \max_{j \in \mathcal{P}(i)} \{A_j + T_j, t_i\}, \quad i \in \mathcal{V}.$$  

For any realization of the activity leadtimes, $A_i$ can be recursively computed for all nodes, starting from the root node with the highest index. Each leaf node $j$ also has a planned finish time $t_j$. This is the due date communicated to the customer. We define the actual finish time for each leaf node as:

$$B_j = \max\{t_j, A_j + T_j\}, \quad j \in \mathcal{L}.$$  

Note that if a product is completed before its planned finish time, it is delivered to the customer at its planned finish time and not earlier. Clearly, the random variables $A_i$ and $B_j$ depend on the vectors of planned start and finish times $\mathbf{t}$ and $\mathbf{t}'$, respectively. We do not explicitly indicate this dependence in the notation, except when it enhances readability.

When an activity is started, investments in resources need to be made. These investments are earned back when the product is delivered to the customer, i.e. after completion of a leaf node. During this period, holding costs are incurred, representing the interest paid on the investment. We define $h_j$ as the holding cost per time unit paid from the planned start time $t_j$ of node $i$ until the completion of leaf node $j$. We have $h_j > 0$ when $j \in \mathcal{L} \cap \mathcal{E}(i)$ and $h_j = 0$ otherwise. The unit echelon holding cost for node $i$ is calculated as $h_i = \sum_{j \in \mathcal{L} \cap \mathcal{E}(i)} h_j$. The unit local holding cost for leaf node $j$ is defined as $h_j = \sum_{i \in \mathcal{L} \cap \mathcal{E}(i)} h_j$. If leaf node $j$ is completed later than its planned finish time $t'_j$, a penalty cost $p_j$ is incurred per unit time late. This leads to the following expression for the expected cost of a production plan defined by the vectors of planned start times $\mathbf{t}$ and finish times $\mathbf{t}'$.

$$C(\mathbf{t}, \mathbf{t}') = \sum_{j \in \mathcal{L}} \left( \sum_{i \in \mathcal{L}} (t_j - \hat{t}_j) h_j + (h_j + p_j) \mathbb{E}[B_j - t_j'] \right)$$  

Since holding costs are incurred from the planned start time, there is no uncertainty in holding cost to be paid until the due date. The only uncertainty is in the actual completion times of the leaf nodes, which in turn depend on the completion times of all predecessor nodes. Our objective is to solve the optimization problem $(P)$ defined as $\min_{\mathbf{t}, \mathbf{t}'} C(\mathbf{t}, \mathbf{t}')$.

3. Tardy paths

Randomness of the activity leadtimes implies that the actual start and finish times might differ from the planned start and finish times. A realization of the leadtimes $T_1, \ldots, T_N$ is indicated by $\omega = (t_1, \ldots, t_N) \in \Omega = \mathbb{R}_+^N$. So $T_i(\omega) = t_i$ denotes a realization of $T_i$. For each realization of leadtimes, we can identify tardy paths. For a realization $\omega \in \Omega$, we say that a path from node $i$ to leaf node $j$ is tardy if there is no waiting time, or "slack" on that path. The formal definition is as follows:

**Definition 1.** For a realization $\omega \in \Omega$, a path from node $i$ to leaf node $j$ is tardy iff

1. Node $i$ starts at the planned start time: $A_i(\omega) = t_i$.
2. For each edge $(k, l)$ on that path, the actual finish time of node $k$ equals the actual start time of node $l$: $A_k(\omega) + T_k(\omega) = A_l(\omega)$.
3. Leaf node $j$ finishes later than planned: $B_j(\omega) > t'_j$.

The definition implies that there exists a tardy path to leaf node $j$ only if node $j$ is late, and if so, it is unique with probability 1, i.e. there is exactly one tardy path to node $j$. To find this path, one starts from node $j$ following the path with no slack, until a node is found that starts on time. Since there is a tardy path for each leaf node that is late, networks with multiple leaf nodes can have multiple tardy paths.

To explain the idea of tardy paths further, consider the networks in Fig. 1. For both networks, we show the production plan obtained from a candidate solution $(\mathbf{t'}, \mathbf{t}'')$ to the optimization problem $(P)$ together with a realization $\omega$ of this plan. Each start time $t_j'$ is denoted by a "flag" pointing to the right and each finish time $t_j''$ by a flag pointing to the left. The realization of each random variable $T_i(\omega)$ is denoted by a gray bar.

Fig. 2 shows the production plan for the network in Fig. 1(a). In this network, nodes 1, 3 and 5 are leaf nodes and thus have planned
Lemma 1. For all leaf nodes $j \in \mathcal{L}$,

$$\frac{\partial}{\partial t_j} \mathbb{E}[B_j(t, t')] = \begin{cases} \sum_{i \in \mathcal{Y}(j)} P(E_{ij}), & \text{if } k = j \\ 0, & \text{otherwise} \end{cases}$$

Proof. First we show that

$$\frac{\partial}{\partial t_j} \mathbb{E}[B_j(t, t') - t_j'] = \mathbb{E} \left[ \frac{\partial B_j(t, t')}{\partial t_j} \right]$$

or equivalently, that for every sequence $g_1, g_2, \ldots$ converging to 0,

$$\lim_{n \to \infty} \frac{\mathbb{E}[B_j(t + g_n e_i, t') - B_j(t, t')]}{g_n} = \mathbb{E} \left[ \frac{\partial B_j(t, t')}{\partial t_j} \right]$$

where $e_i$ is the unit vector with 1 at position $i$. To find the derivative, we rewrite $B_j(t, t')$ as follows. For every realization $\omega = (t_1, \ldots, t_n)$ we have

$$B_j(t, t')\omega) = \max \left\{ t_j', \max_{w \in \mathcal{W}(j)} \left( t_j + \sum_{l \in w} T_l(\omega) \right) \right\}$$

Hence the derivative $\frac{\partial B_j(t, t')\omega)}{\partial t_j}$ exists for almost all $\omega$ and

$$\frac{\partial B_j(t, t')\omega)}{\partial t_j} = \begin{cases} \mathbb{1}, & \text{if } B_j(t, t')\omega) = \arg\max_{w \in \mathcal{W}(j)} \left( t_j + \sum_{l \in w} T_l(\omega) \right) \\ 0, & \text{otherwise.} \end{cases}$$

For all $\omega$,

$$\left| B_j(t + g_n e_i, t')\omega) - B_j(t, t')\omega) \right| = 1,$$

so by bounded convergence we can conclude that (4) holds. Since the derivative with respect to $t_j$ is 1 if a path from $i$ to $j$ is tardy and 0 otherwise, we get

$$\frac{\partial}{\partial t_j} \mathbb{E}[B_j(t, t') - t_j'] = \mathbb{E} \left[ \frac{\partial B_j(t, t')}{\partial t_j} \right] = P(E_{ij}),$$

which proves (3). The proof of (2) proceeds along the same lines. □

Observe that the derivative is only a function of planned start and finish times of nodes that share the same leaf node.

Lemma 2. The cost function $C(t, t')$ is jointly convex in $(t, t')$.

Proof. Convexity of (1) immediately follows from convexity of (5) in $(t, t')$ for every $\omega$. □

This property guarantees that any local optimal production plan $(t^*, t^*)$ is also a global optimum. For the optimal production plan in an assembly network with a single leaf node [5] shows that the probability that a tardy path starts in node $i \in V$ is proportional to the value this activity adds to the final product. We show that this result also holds for specific nodes in production networks with multiple leaf nodes, depending on their location in the network. Before stating the optimality equations for the production plan, we define $\mathcal{U}$ as the set of nodes, which have the property that they can reach exactly one leaf node: $\mathcal{U} = \{ i : |\mathcal{L} \cap \mathcal{E}(i)| = 1 \}$. According to this definition, $\mathcal{U} = \{ 1, 2, 3, 5 \}$ and $\mathcal{U} = \{ 1, 2, 3, 4, 5, 6, 7, 8, 9 \}$ for the networks in Figs. 1(a) and 1(b), respectively. A more complicated network and its corresponding set $\mathcal{U}$ are shown in Fig. 4. So far, the networks studied in the literature have only one leaf node, for which it immediately follows that $\mathcal{U} = \emptyset$. We denote the optimal solution to (P) as $(t^*, t^*)$. Under the optimal solution the probability of a path being tardy is denoted by $P(E_{ij})$. The optimality equations for $(t^*, t^*)$ are formulated in Theorem 1.
Theorem 1. The optimal production plan \((t^o, t^*)\) of \((P)\) satisfies the following equations:

\[
\sum_{i \in \mathcal{U}} P(E_{i,j}^+) = \frac{h_j^c}{h_j^c + p_j}, \quad j \in \mathcal{L} \tag{6}
\]

\[
\sum_{j \in \mathcal{L}(i) \cap \mathcal{L}} (h_j^c + p_j)P(E_{i,j}^+) = h_i^c, \quad i \in \mathcal{V} \tag{7}
\]

\[
P(E_{i,j}^+) = \frac{h_{i,j}}{h_j^c + p_j}, \quad i \in \mathcal{U}, j \in \mathcal{L} \tag{8}
\]

Proof. To prove (6) we take the partial derivative of (1) with respect to \(t_{f,j}^*\). By Lemma 1,

\[
\frac{\partial \mathcal{C}(t^*, t^0)}{\partial t_{f,j}^*} = -\sum_{i \in \mathcal{U}(j)} h_{i,j} + (h_j^c + p_j) \sum_{i \in \mathcal{V}(j)} P(E_{i,j}^+), \quad j \in \mathcal{L}
\]

At optimality this derivative should vanish, so

\[
\sum_{i \in \mathcal{V}(j)} P(E_{i,j}^+) = \frac{h_j^c}{h_j^c + p_j}
\]

To prove (7) and (8) we take the derivative of (1) with respect to \(t_{f,k}^*\), yielding

\[
\frac{\partial \mathcal{C}(t^*, t^0)}{\partial t_{f,k}^*} = -h_k^c + \sum_{j \in \mathcal{L}(k) \cap \mathcal{L}} (h_j^c + p_j)P(E_{k,j}^+), \quad k \in \mathcal{V}.
\]

Since this derivative vanishes at optimality, we find

\[
\sum_{j \in \mathcal{L}(k) \cap \mathcal{L}} (h_j^c + p_j)P(E_{k,j}^+) = h_k^c, \quad k \in \mathcal{V}.
\]

If \(k \in \mathcal{U}\), then \(k\) can reach only one leaf node. Hence the above equation reduces to

\[
P(E_{k,j}^+) = \frac{h_k}{h_j^c + p_j}, \quad k \in \mathcal{U},
\]

which concludes the proof. \(\square\)

The interpretation of this theorem is as follows: (6) states that the classical Newsvendor equation holds for every leaf node. This equation depends only on the local holding cost \(h_j^c\) and penalty cost \(p_j\) of the corresponding leaf node and is independent from the local holding and penalty costs of other leaf nodes. The left-hand side of (7) is a summation over all leaf nodes reachable from \(i\). Each term in the summation denotes the probability that a path from \(i\) to \(j\) is tardy, multiplied by the sum of holding and penalty costs at that leaf node. This summation equals the echelon holding cost \(h_j^c\). If \(i \in \mathcal{U}\) the summation of (7) reduces to only one term and thus the equation can be simplified to (8) which is a closed form expression for the probability that a path is tardy. This equation states that under the optimal solution, the probability that a path from \(i\) to \(j\) is tardy, is proportional to value node \(i\) adds to leaf node \(j\), i.e., \(h_{i,j}\). A special case in which it is possible to derive a closed-form expression for \(P(E_{i,j}^+)\) is stated in the following lemma.

Lemma 3. If for \(i \notin \mathcal{U}\) and \(j \in \mathcal{L}\), it holds that \(\{i\} = \mathcal{Y}(j) \setminus \mathcal{U}\), then

\[
P(E_{i,j}^+) = \frac{h_{i,j}}{h_j^c + p_j}
\]

Proof. In case \(\{i\} = \mathcal{Y}(j) \setminus \mathcal{U}\), the tardy path probability can be written as

\[
P(E_{i,j}^+) = \sum_{k \in \mathcal{Y}(j) \setminus \mathcal{U}} P(E_{k,j}) - \sum_{k \in \mathcal{Y}(j) \setminus \mathcal{U}} P(E_{k,j})
\]

Using (6) and (8) we find that

\[
P(E_{i,j}^+) = \frac{h_j^c}{h_j^c + p_j} - \sum_{k \in \mathcal{Y}(j) \setminus \mathcal{U}} \frac{h_{i,j}}{h_j^c + p_j} = \frac{h_{i,j}}{h_j^c + p_j}. \quad \square
\]

Lemma 3 is useful for the network shown in Fig. 1(a). For this network, nodes 1 to 9 belong to the set \(\mathcal{U}\) and hence (8) holds for these nodes. Node 10 is not part of \(\mathcal{U}\), but the conditions in Lemma 3 are satisfied and hence (9) holds for this node.

The optimality equations can only be solved for special network structures. In general, the solution is complicated, since the expected lateness at each end node depends on the leadtime distributions of all its parents. The optimality equations are also useful to validate solutions obtained via approximate methods as the resulting tardy path probabilities should satisfy these equations (see Section 5). Furthermore, the optimality equations can be used to derive structural properties of the optimal solution. This is demonstrated below.

A strictly converging network is a network in which each node has exactly one successor, except for the only leaf node, node 1. Hence \(\mathcal{U} = \mathcal{V}, \mathcal{L} = \{1\}, \mathcal{Y}(1) = \mathcal{V}\) and thus (8) applies to each node in the network. The following monotonicity properties hold for the tardy path probabilities in strictly converging networks.

Lemma 4. For each node \(i \in \mathcal{V}\) in a strictly converging network, \(P(E_{1,i})\) is

(a) increasing in \(t_{f,i}^0\),
(b) decreasing in \(t_{f,i}^*\),
(c) decreasing in \(t_{f,k}^*\) for all \(k \in \mathcal{V} \setminus \{i\}\).

Proof.

(a) Consider a realization \(\omega \in E_{1,i}\), then the path from \(i\) to 1 is tardy and it remains so by increasing \(t_{f,i}^0\).

(b–c) Consider a realization \(\omega \notin E_{1,i}\), then the path from \(i\) to 1 is not tardy. It remains not tardy by increasing \(t_{f,i}^0\) or by increasing \(t_{f,k}^*\) for any \(k \in \mathcal{V} \setminus \{i\}\). \(\square\)

We can also look at properties of combinations of the tardy path probabilities.

Lemma 5. For any nonempty subset \(\mathcal{Q} \subseteq \mathcal{V}\) in a strictly converging network, \(\sum_{k \in \mathcal{Q}} P(E_{k,i})\) is increasing in each \(t_{f,k}^*\), \(i \in \mathcal{Q}\).

Proof. Let \(i \in \mathcal{Q}\). We need to show that the event \(\bigcup_{k \in \mathcal{Q} \setminus \{i\}} E_{k,i}\) increases as \(t_{f,i}^0\) increases. Consider a realization \(\omega \in \bigcup_{k \in \mathcal{Q} \setminus \{i\}} E_{k,i}\). If \(\omega \notin E_{k,i}\), then the path from \(i\) to 1 is tardy and it remains so by increasing \(t_{f,i}^0\). If \(\omega \in E_{k,i}\) with \(k \in \mathcal{Q} \setminus \{i\}\), then the path from \(k\) to

\[\text{...}\]

\[\text{...}\]
1 is tardy. By increasing $t_l^i$, either this path stays tardy, or the path from $i$ to 1 becomes tardy. In both cases, the realization stays in $E_{k1} \cup E_{k1}$.

Using the above properties, we derive monotonicity properties of the optimal solution with respect to the penalty cost parameter $p_1$.

**Lemma 6.** For each node $i \in V$ in a strictly converging network, the planned start time $t_l^i$ of the optimal solution $t_l^i$ is decreasing in $p_1$.

**Proof.** Let $p_1 > p_1$ and denote the corresponding optimal solutions by $t_l^i$ and $t_l^i$. Increasing $p_1$ decreases the Newsvendor fractile on the right-hand side of (6) and (8). Hence, to satisfy (6) for $p_1$, the probability $\sum_{i=1}^n P(E_{k1})$ should reduce. From Lemma 5 it follows that this probability is increasing in any $t_l^i$. Hence $t_l^i$ contains at least one start time $t_l^i < t_l^i$. Let $Q$ be the set of all nodes $k$ for which $t_l^k < t_l^k$. We now show that $Q = V$. Lemma 4(c) and 5 imply that the sum of tardy probabilities $\sum_{k \in V \setminus Q} P(E_{k1})$ for $t_l^i$ is greater than the one for $t_l^i$. Hence, there is at least one node $k \in V \setminus Q$ for which $P(E_{k1})$ for $t_l^i$ is greater than the one for $t_l^i$. But then it is impossible that for $t_l^i$, optimality equation (8) is satisfied for this node $k$. Hence, we conclude that $V \setminus Q = \emptyset$. □

An increase in the penalty cost leads to a higher optimal service level, i.e., more on-time completions. To achieve this level, each individual node is planned to start earlier. This clearly shows the trade-off between planned leadtimes and service level.

### 5. Numerical example

In this section, we provide a numerical example for the network in Fig. 4. As stated by [3] exact calculation of $C(t_l^i, t_l^i)$ is computationally intensive for networks with many nodes, specifically if nodes are in series. This is due to multi-dimensional integrals that need to be computed to obtain the expected lateness of each leaf node. To overcome this problem, we rely on a simulation. We generate a (large) number of samples $\omega_1, \ldots, \omega_n$ and estimate $C(t_l^i, t_l^i)$ by its sample average. Note that this sample average is also jointly convex (see the proof of Lemma 2). The set-up of our experiment is as follows. For each activity we use a Gamma distribution with shape parameter $k = 3$ and scale parameter $\theta = 4$. For the holding costs we have $h_{i,j} = 1$ if $j \in \mathcal{Y}(i)$ and 0 otherwise. For the penalty costs, we have $p_j = 10$ for all $j \in \mathcal{C}$. The sample size is set to $n = 10^6$. Without loss of generality we set $t_l^0 = 0$. The results are presented in Fig. 5. It seems counterintuitive that the start times of nodes with the same predecessor are not necessarily equal. For example, nodes 12 and 13 have the same predecessor (node 14), but have different start times. These nodes have different network locations and therefore the parameters in the optimality equations for their planned start times are different. In specific cases, start times are equal, for example for nodes 4 and 5, which is due to symmetry.

To validate the obtained solution we simulate the production plan using $10^6$ new samples of the activity lead times. We estimate the tardy path probabilities $P(E_{k1})$ and verify whether these probabilities satisfy Eqs. (6), (7), (8) and (9). The results are shown in Tables 1–3. In each table, the first column denotes the right-hand side (RHS) of an optimality equation, the second column denotes the simulated left-hand side (LHS) of an equation and the third column denotes the relative gap, defined as $\frac{|\text{LHS} - \text{RHS}|}{\text{LHS}} \times 100\%$.

Clearly, gaps are all within 3% and thus we conclude that the simulation results are sufficiently close to optimal. For Eq. (6) the gap is even smaller: it is at most 0.58%. For this example we use $10^6$ samples, which allows us to obtain a solution within seconds. The solution can be improved by increasing the number of samples used in the simulation.

### 6. Conclusion and future research directions

In this paper, we generalize the results available in the literature for the problem of setting planned leadtimes in assembly and serial production systems. For a large class of networks, we prove optimality equations, which have a Newsvendor structure for specific networks. In future research these optimality equations can be exploited to further develop and improve heuristics that solve the planning problem for large networks faster and more accurately. In addition, studying production networks with different cost structures, for example non-linear, may reveal new practical insights.
References


