The stochastic economic lot scheduling problem: a survey

Citation for published version (APA):

Document status and date:
Published: 01/01/2005

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: 23. Aug. 2020
The Stochastic Economic Lot Scheduling Problem: A Survey

E.M.M. Winands¹,², I.J.B.F. Adan¹ and G.J. van Houtum²

¹Department of Mathematics and Computer Science
²Department of Technology Management
Technische Universiteit Eindhoven
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{e.m.m.winands, i.j.b.f.adan, g.j.v.houtum@tue.nl}

January 4, 2005

Abstract

We consider the production of multiple standardized products on a single machine with limited capacity and set-up times under random demands and random production times, i.e., the so-called stochastic economic lot scheduling problem (SELS). The main task for the production manager in this setting is the construction of a production plan for the machine that minimizes the total costs, i.e., the sum of holding, backlogging and set-up costs. Based on the critical elements of such a production plan, we give a classification and extensive overview of the research on the SELSP together with an indication of open research areas.

Keywords: inventory, production, finite capacity, multiple products, set-up times, set-up costs.
1 Introduction

Production plants usually have one stage that generates most of the added value and utilizes the most expensive installation, the so-called bottleneck. Production on the bottleneck often has to be done under very tight capacity constraints, as described by Fransoo [18, 19]. Bourland [8] indicates that in practice the bottleneck is often one of the stages that produce products used in the assembly of finished goods. The scheduling decisions of the plant are in general dominated by this bottleneck, i.e., the production plans of upstream and downstream stages are deduced from the plan deployed at the bottleneck, as indicated by Bourland [8]. Scheduling production of multiple products on a single machine under tight capacity constraints is one of the classic problems in operations research. There are many variations of multi-product single-machine scheduling problems, but we may classify them by the following three characteristics:

1. \textit{Presence or absence of set-up times and/or costs.} The most important impact of set-ups on the production plan is that the products need to be produced in batches, since otherwise costly capacity is wasted on set-ups. Furthermore, set-ups make it impossible to be completely responsive to the demand, as argued by Bourland [8].

2. \textit{Customized or standardized products.} Since customized products can only be produced when there is a request for an order, these products have to be produced in a \textit{make-to-order} fashion. In case of standardized products one may choose for a \textit{make-to-stock} production policy, because such products do not have to be produced to customer specifications. It is obvious that standardized products give more freedom in deciding when to make which product and in what quantity.

3. \textit{Stochastic or deterministic environment.} In a completely deterministic environment one can confine oneself to a cyclic production plan. However, when the company has to be responsive to a stochastic environment, i.e., stochastic demand, production times or setup times, such a cyclic schedule will not suffice anymore.
The Venn diagram in Figure 1 depicts the eight subproblems, which are created by the above characterization. Of course, this diagram only provides a global classification of multi-product single-machine scheduling problems. One can easily think of examples of so-called hybrid systems (see, e.g., Adan and Van der Wal [1] and Federgruen and Katalan [17] for the combination of make-to-order and make-to-stock).

The present paper considers the production of multiple standardized products on a single machine with limited capacity and possibly random set-up times under random demands and possibly random production times, depicted as subproblem 8 in Figure 1: the so-called stochastic economic lot scheduling problem (SELS). In this problem, a production plan is needed which describes for each possible state of the system whether to continue production of the current product, whether to switch to another product or whether to idle the machine. The primary goal of such a production plan is minimization of the total costs, that is, the sum of holding, backlogging and set-up costs. As emphasized by Karmarkar [27], set-up costs should not be used as surrogate for set-up times in the modeling of scheduling problems. Costs do not completely capture the problematic nature associated with set-up times, i.e., the loss of production capacity each time a set-up is performed. We consider both set-up costs and times, where set-up costs represent the actual costs due to, e.g., loss of material and, hence, they are not used as surrogate for set-up times.

The present paper does not comprise the other subproblems shown in Figure 1. Suggested books for readers who would like to pursue their knowledge of these subproblems are Silver et al. [37] and Zipkin [47]. Nevertheless, in Section 2 of the present paper we describe the relationship between the SELSP and subproblem 4, which is usually called the (deterministic)
economic lot scheduling problem (ELSP). This relation is described for a couple of reasons. First of all, subproblem 4 is closely related to the SELSP, i.e., it can be seen as its deterministic counterpart. Secondly, in developing a production strategy in a stochastic environment the solutions for the ELSP are frequently used as basis for the stochastic solution. Thirdly, in practice a deterministic solution is often wrongfully implemented, without any adaptations, in a stochastic environment.

The SELSP is an interesting problem both from a theoretical and a practical point of view. The theoretical justification lays in the fact that the SELSP is generally regarded as a challenging problem; the finite production capacity has to be dynamically distributed among the products in order to be reactive to the stochastic demands, productions and set-ups (see Sox et al. [41]). The presence of set-up times in combination with the stochastic environment are the key complicating factors of the problem. On the one hand, one seeks for short cycle lengths, and thus frequent production opportunities for the various products, in order to be able to react to the stochasticity in the system. On the other hand, short cycle lengths will increase the set-up frequency, which has a negative influence on the amount of capacity available for production. Consequently, this effect will hinder the timely fulfillment of demand. The trade-off between these contrasting effects forms a challenging issue from a theoretical point of view. Allahverdi et al. [2] recently wrote a survey on deterministic scheduling problems involving set-ups, in which they indicate stochastic scheduling problems with set-ups as an avenue of future research.

The SELSP is a common problem in practice, e.g., in glass and paper production, injection molding, metal stamping and semi-continuous chemical processes, but also in bulk production of consumer products such as detergents and beers. Some specific applications described in the open literature are a laminate manufacturing plant (see Anupindi and Tayur [7]), a glass-containers manufacturing company (see Fransoo et al. [19]), a large consumer products manufacturer (see Gascon et al. [22]), a producer of plastic bumpers for cars (see Grasman et al. [24]) and an aerospace component supplier (see Sox and Muckstadt [40]).

The aim of the present paper is twofold. First of all, it gives an overview of the research on the SELSP along with a comprehensive list of references. Secondly, several areas for future
The present section gives a detailed problem description of the SELSP (Subsection 2.1) and gives the relationship between the SELSP and its deterministic counterpart (Subsection 2.2).

2.1 Model

We consider a system with one single production capacity for multiple products, in which there is an infinite stock space for each product and raw material is always available. Demands for the various products arrive according to stationary and mutually independent stochastic processes. Demand that cannot be satisfied directly from stock is either lost or backlogged until the product becomes available after production. The individual products are produced in a make-to-stock fashion with possibly stochastic production times. A possibly stochastic
set-up time occurs \textit{before} the start of the production of a product. Motivated by the nowadays efficient control of the production process, often the assumption is made that the production and set-up times are deterministic. The set-ups are, furthermore, independent of the demand processes, production times and other set-up times. Finally, only one product can be produced at a time. The main objective of the SELSP is to minimize the total expected costs, i.e., the sum of set-up, holding and backlogging costs, per unit of time over a planning horizon, which can either be finite or infinite. Besides the total costs, other quantities of interest could, for example, be the fraction of time that is lost due to set-ups, the fill rate (the fraction of demand satisfied directly from stock), the average stock level or the waiting time of customers.

\textbf{Remark 2.1.} In some cases set-ups depend only on the product to be produced, so-called \textit{sequence-independent} set-ups, while in other cases the set-ups are dependent both on the product to be produced and the preceding product, so-called \textit{sequence-dependent} set-ups. When set-up times or costs are strongly sequence-dependent, it is advisable to introduce \textit{product families}, in such a way that between the families set-ups are significant, but that they are negligible within a family. By doing so, one may reduce the costs and capacity waste induced by set-ups and make the problem more manageable. \hfill \Box

\textbf{Remark 2.2.} In many practical situations products are produced on a multi-stage production line. If this line allows no interchangeability of products on other lines, if no intermediate storage is possible and if production times are short, such a line can be regarded as a single machine and, thus, the discussion of the present paper fully applies (see Fransoo et al. [19]). \hfill \Box

\section{2.2 Relation with the deterministic ELSP}

The deterministic counterpart of the SELSP, the so-called ELSP, has received lots of attention in the literature over the past decades (for surveys see, e.g., Elmaghraby [11] and Salomon [36]). In the literature two approaches can be distinguished for the ELSP (see again, e.g., Elmaghraby [11]), which has been proven to be NP-hard (see Hsu [26]): analytical approaches with optimal solutions for restricted problems and heuristic approaches with good solutions.
for the general problem. Both approaches derive a rigid cyclic schedule, which will be strictly followed until the end of the planning horizon, as described by Gascon et al. [22]. Unfortunately, the solutions of the ELSP can only be applied in an ideal plant, where machines are perfectly reliable, set-up and production rates are constant, raw material and tools are always available, demand is known and initial inventories are on level, as argued by Gallego [20]. This is a utopia and, thus, the deterministic problem has to be extended to a stochastic version, the SELSP.

Two major differences can be seen between production plans for a deterministic environment on the one hand and a stochastic environment on the other hand. Firstly, a rigid cyclic production plan will not suffice anymore in a stochastic environment, since one has to be responsive to the dynamic changes in this environment. This means that dynamics have to be included in the production plan. Secondly, in a stochastic environment the inventories for the individual products play a more important role than in the deterministic case, as indicated by Sox et al. [41]. Inventories now do not only reduce the number of set-ups in a cycle, but they also serve as hedge against stock-outs and scheduling conflicts due to the variation in demand, production or set-up times. The inventory levels needed to guaranty a certain service level increase therefore drastically in a stochastic environment. Irrespective of these differences, Bourland [8] correctly notes that the insights and results obtained from the deterministic studies remain useful, since a deterministic production plan is frequently used as a basis for the solution of the stochastic problem.

In contrast to the deterministic problem, until the end of the seventies the SELSP received almost no attention in the literature. Most likely, this lack of attention was not caused by an absence of practical interest, but by the intrinsic analytical complexity of the problem. To illustrate this feeling we cite here the following conclusion drawn by Vergin and Lee [43] in 1978 concerning the state of the research on the SELSP at that time:

*The literature is almost completely void of not only the development of analytical models, but even of discussion of the problem. A thorough review of the production scheduling and inventory management journals and books would almost suggest that the scheduling problem does not exist. Yet the multiple product single machine*
The first papers that recognized the need for analytical models for the SELSP were published by Vergin and Lee [43] and by Graves [25]. Although both papers make the assumption, in contrast to the SELSP, that the set-up times equal zero, these papers can be seen as the beginning of the research on the SELSP. More particular, in the paper of Vergin and Lee [43] a set of simple dynamic heuristics has been compared by simulation for the problem with deterministic production and set-up times. They conclude that the, at that time widely used, deterministic solution cannot effectively be used in a stochastic environment. Furthermore, it turns out that no heuristic exists that is better than its competitors in all situations. In Graves [25], a periodic review policy is studied. He first formulates the single-product problem with an infinite horizon as a discrete-time Markov decision problem over a two-dimensional state space consisting of the stock level and the machine status. This single-product problem is solved by the policy iteration method on a truncated state space. Since the computation time to solve the multi-product problem increases dramatically as the number of products increases, Graves [25] resorts to a heuristic in the multi-product case that relies on the subsequent analysis of a large number of single-product problems. Therefore, the notion of composite products is introduced, which are nothing more than aggregations of individual products. The developed heuristic may at each point in time idle the machine, produce the product with the least inventory or produce the product that is currently set-up.

The present paper provides an overview of the research on the SELSP, which started two decades ago with the above-mentioned papers of Vergin and Lee [43] and Graves [25]. Because of the early existence of the research on the SELSP, several interesting research questions are still unanswered. With this paper we hope to start a discussion on the question of what are important problems concerning the SELSP both from a theoretical and a practical point of view.
3 Critical elements of the production plan

The present section describes the critical elements of a production plan, i.e., the production sequence (Subsection 3.1) and the lot sizing policy (Subsection 3.2). Based on these elements we propose a classification of the production strategies for the SELSP in Subsection 3.3.

A production plan for the machine must describe for each possible state of the system whether to continue production of the current product, whether to switch to another product or whether to idle the machine, as described by Bourland and Yano [9]. Federgruen and Katalan [16] correctly argue that a policy making the aforementioned decisions at each point in time in a dynamic fashion based on a complete state description of the system will lead to the optimal costs. However, in general the bottleneck is part of a larger multi-stage production process. Therefore, we should not only be interested in the costs at the bottleneck, but also in the coordination of the plan with upstream and downstream stages of this process. Scheduling decisions at the bottleneck do not only affect the performance of the bottleneck itself, but they also strongly influence the performance of the other stages, as indicated by Bourland [8] and Bourland and Yano [9]. When this coordination between the stages is improved, the estimates of total lead time may get better and, therefore, the company may be able to set more realistic due dates, as described by Anupindi and Tayur [7] and Erkip et al. [12].

In order to simplify the coordination with downstream and upstream stages, these stages may impose all kinds of restrictions on the production plan of the bottleneck. In general, these restrictions concern the following two major elements of the production plan:

1. Production sequence;

2. Lot sizing policy.

The above two decisions are of course not the only decisions which ought to be made. The production manager has, for instance, to decide on the idle times between the production runs and the safety stocks as well. However, the sequencing and the lot sizing decisions can be seen as the most critical ones.
3.1 Production sequence

The first major restriction we identify is whether a *fixed production sequence* is used or not. A fixed production sequence means that there exists a pre-defined *order* and *frequency* for the production of the individual products. For example, if products A, B and C have to be produced a possible production sequence would be A-B-A-C. A sequence in which each product is produced exactly once in each cycle is called a *pure rotation sequence*, e.g., C-A-B. By fixing the production sequence downstream and upstream stages exactly know the order in which products enter and have to leave their own machines, respectively. According to Bourland [8], this leads to some kind of stability. Moreover, fixed production sequence strategies are easy to implement on the floor.

For production strategies using a fixed production sequence an additional classification can be made, i.e., whether or not a pre-defined *fixed cycle length* is used. In the sequel we adapt the definition of Bourland [8] for this concept, i.e., the cycle length is equal to the time between two successive completions of the production sequence. Fixing the cycle length has the advantage that one exactly knows when the production plan starts over again.

Based on the above criteria concerning the production sequence and the cycle length, the production strategies for the SELSP can be divided into the following three categories:

- *Dynamic production sequence*;
- *Fixed production sequence in combination with a dynamic cycle length*;
- *Fixed production sequence in combination with a fixed cycle length*.

When the production manager makes the decision to use a strategy from the latter two categories, he mainly trusts lot sizing strategies as tool to respond to the stochastic environment of the plant. It goes without saying that a strategy from the first category leads to a cost reduction in comparison with a strategy from the latter two categories. A trade-off should be made whether or not this cost reduction is outweighed by the less efficient coordination in the production line or with the environment of the plant.
3.2 Lot sizing policy

The second critical element of a production plan is the deployed lot sizing policy. Roughly speaking, we can distinguish two general classes of such policies:

1. **Global lot sizing policies**: lot sizing decisions may depend on the complete state of the system, i.e., stock levels of all the individual products and the state of the machine;

2. **Local lot sizing policies**: lot sizing decisions only depend on the stock level of the product currently set-up.

A global strategy has the advantage, in contrast to a local policy, that it is able to react to changes in the complete state of the system at each point in time. When a local strategy is deployed, typically a *fixed-quantity* policy or a *base-stock* policy is used. In the former a fixed quantity of a product is produced when production is commenced, whereas in the latter the machine will continue production until a pre-defined target inventory level has been reached. The latter type of policy is also known as *produce-up-to* policies. A major drawback of the base-stock policy is that one single product, for which a high demand arrives in a certain period of time, may dominate the machine for a while. The impact of this phenomena on the other products are stock outs, highly variable cycle lengths and high costs. In the literature a number of variations on the standard base-stock policy, which in the sequel of the paper is referred to as the *exhaustive* base-stock policy, have been introduced to circumvent this drawback, e.g.,

- **Gated base-stock policies**: when the machine starts production of a product, it will continue production until a production batch has been completed the size of which equals the difference between the base-stock level and the starting inventory level;

- **Quantity-limited base-stock policy**: when the machine starts production of a product, it will continue production until either the base-stock level has been reached or a maximum number of products has been produced;

- **Time-limited base-stock policy**: when the machine starts production of a product, it will continue production until either the base-stock level has been reached or a maximum
An additional advantage of bounding the production runs is the facilitation of preventive maintenance and cleaning of the machine.

### 3.3 Classification of the literature

In Subsections 3.1 and 3.2 we have introduced two critical elements of a production plan: the sequencing and the lot sizing policy, respectively. By using these elements a classification of the strategies for the SELSP can be made, which is depicted in Figure 2. Based on this classification, we review in the next section the existing literature on the SELSP.

### 4 Strategy classes for the SELSP

In this section we discuss the production strategies proposed for the SELSP. In doing so, we follow the classification introduced in the preceding section.

#### 4.1 Dynamic production sequence

The papers concerning the two types of dynamic production sequence policies, i.e., using a global or a local lot sizing policy, are described in the remainder of the present subsection.
4.1.1 Global lot sizing policy

Sox and Muckstadt [40] propose a finite-horizon discrete-time stochastic optimization model under the assumption of the availability of overtime and deterministic production and set-up times. They propose a method for finding optimal or near-optimal solutions applicable for small problems by using a Lagrangian decomposition algorithm. Sox and Muckstadt [40] assume that a set-up for a product is incurred even if the same product was produced in the preceding period. They argue that this assumption can easily be relaxed at the expense of increased computation times.

Qiu and Loulou [35] present a multi-product model with limited stock space for each product under the assumption of backlogging, Poisson demand and deterministic production and set-up times. They model the problem as a continuous-time semi-Markov decision problem with an infinite horizon, where the state space consists of the individual stock levels and the status of the machine. By using the successive approximations technique a policy can be derived on a truncated finite state space, which is then extended to a near optimal policy for the original model. The general structure of such a near optimal policy for a two-product example is shown in Figure 3. That is, a policy creates in general four regions. When the individual stock levels are in regions I and II, products one and two have to be produced, respectively, irrespective of the status of the machine. In region IV nothing is produced, whereas in region III we should continue the production of the current product. Obviously, the switching curves between these regions, and thus the optimal policy, depend on the stock levels of both products. This means that a local lot sizing policy is, in general, not optimal. Finally, Qiu and Loulou [35] conclude that for problems consisting of more than two products, their solution procedure cannot be both efficient and accurate due to the notorious curse of dimensionality. They suggest using composite products in these cases as in Graves [25]. Though they state that preliminary results indicate that this aggregation of products is promising, no results have appeared in the open literature yet.

Finally, we want to mention work of Karmarkar and Yoo [28], in which the SELSP is also studied under the assumption of backlogging, deterministic production and set-up times. They present a formulation of a discrete-time dynamic stochastic programming problem over
Figure 3: Structure of a near optimal policy of Qiu and Loulou [35].
Region I: produce product 1; Region II: produce product 2; Region III: produce product currently set-up; Region IV: produce nothing.

a finite horizon under the assumption of time-varying stochastic demand. Several Lagrangian relaxations of the problem are introduced, which can provide lower and upper bounds for the original problem. The results on small-scale problems are, however, not very encouraging.

4.1.2 Local lot sizing policy

Production strategies that fall within this class are often of the so-called independent stochastic control type (see, e.g., Sox et al. [41]). This means that the lot sizing decisions are made locally according to standard single product inventory control strategies such as \((s, Q)\) or \((s, S)\) policies, whereas the sequencing decisions are dynamically resolved by using priority rules. In particular, Zipkin [46] studies a continuous-time model under the assumption of backlogging, Poisson demand processes and generally distributed set-up and production times. The individual lot sizing policies are of the \((s, Q)\) type, while the batches for the various products are produced in a first come first served (FCFS) order. Zipkin [47] makes the additional assumption that the production time of a batch is (nearly) independent of the size of this batch. He derives optimal batch sizes and reorder points with respect to total costs.

Altiok and Shiue [3, 4, 5] study a model for the SELSP, in which an \((s, S)\) policy is implemented. The sequencing decision is made based on either one of the following two priority
rules. The first one is a standard priority rule, which states that when the stock level of the product currently set-up reaches its base-stock level, the machine starts production of the highest priority product with stock level below its reorder point. The second priority rule is a cyclical one, which serves the products with stock below their reorder points in a cyclical manner. In the case of backlogging, Altiok and Shiue [3] present an approximate analysis for the case of three products, while in Altiok and Shiue [5] this paper is extended to the N-product case. More specifically, an approximate iterative procedure is developed, which assumes independence among the stock levels of the products. In Altiok and Shiue [4], an exact analysis is performed for the lost sales case under the additional assumptions of phase-type distributed production times and exponential distributed set-up times.

4.2 Fixed production sequence in combination with a dynamic cycle length

In the context of fixed sequence strategies using a dynamic cycle length, we can distinguish between strategies using a global lot sizing policy or a local one. These two types of strategies are discussed in more detail in the remainder of the present subsection.

4.2.1 Global lot sizing policy

Markowitz et al. [33] propose a model, in which the demands are allowed to follow general renewal processes. At each point in time, the production manager has the following options: produce the product currently set-up, idle the machine or switch to the next product in the production sequence. Markowitz et al. [33] study the cases of set-up times and costs separately, but the combination can be analyzed without much additional effort. Motivated by well-known heavy-traffic limit theorems, a time-scale decomposition is made. By doing so, the SELSP can be approximated by a diffusion control problem. This problem can explicitly be solved for the set-up cost case, whereas one has to resort to an algorithmic procedure in the set-up time case. The paper is completed with a numerical evaluation of the resulting policies, in which, among others, a discrete-event simulation is used. These policies can be characterized via three regions of the total workload as shown in Table 1, which is copied from Markowitz et al. [33]. The total workload equals the sum of the production times for
products in stock minus the sum of production times for products backlogged, i.e., the actual stock position of the system. In Markowitz and Wein [34] the same kind of heavy traffic analysis is applied to all kinds of related stochastic multi-product single-machine scheduling problems.

Bourland and Yano [9] use a two-level hierarchical policy for the SELSP under the assumption of deterministic production and set-up times. Their strategy assumes that for each individual product a reorder point is given. Besides idle time and safety stocks, overtime may be used to respond to the stochastic demand. Since the backlogging costs are higher than the overtime costs, no demand is backlogged. At the upper level, the planning level, a cyclic schedule is obtained without neglecting the stochasticity of the demand. At this level one decides on the cycle length, stock levels and idle time allocations given the reorder points. Moreover, this planning level sets targets for the lower level, the control level. At this lower level, a control rule is defined that tries to follow the target schedule. The control policy does not alter the production sequence, but it may move the production starts forward or backward in time and, thus, the actual cycle length may differ from the target length. Roughly speaking, their control policy works as follows: at the moment the next product in the sequence hits its reorder level before the end of the production interval of the current product, the remainder of the production quantity of the current product is made on overtime and the production run of the next product commences immediately (see Figure 4).

<table>
<thead>
<tr>
<th>Workload</th>
<th>Cycle length in set-up time problem</th>
<th>Cycle length in set-up cost problem</th>
<th>Cheapest Product</th>
<th>Other Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much larger than zero</td>
<td>Set small to avoid holding costs, time is cheap</td>
<td>Set large to avoid set-ups at expense of holding costs</td>
<td>Excess buffer is created</td>
<td>Balanced between excess and backlog</td>
</tr>
<tr>
<td>In the neighborhood of zero</td>
<td>Dynamic trade-off between wasted time and holding costs, time is important</td>
<td>Dynamic trade-off between set-up costs and holding costs</td>
<td>Balanced between excess and backlog</td>
<td>Balanced between excess and backlog</td>
</tr>
<tr>
<td>Much smaller than zero</td>
<td>Set large to reduce backorders, time is critical</td>
<td>Set small to minimize backlogging of all products</td>
<td>Severe backlogging is allowed</td>
<td>Balenothead</td>
</tr>
</tbody>
</table>
The production quantity is then determined by a so-called *match-up* lot sizing policy, which is defined as follows in Bourland [8]: a match-up policy schedules production of a product in such a way that the stock level at the *planned* completion time - and not necessarily on the *actual* completion time - of the production run will be equal to the base-stock level. In Bourland [8], it is shown by means of a simulation study that such a match-up policy follows the target cycles more effectively compared to a standard (exhaustive) base-stock policy. By using simulations Bourland and Yano [9] conclude that in comparison to safety stocks and overtime, idle times are a very expensive tool against demand uncertainty.

Gallego [20] proposes a three-level production strategy. At the first level, Gallego [20] constructs the production sequence, the production quantities and the idle times based on a deterministic procedure, but he does not add safety stocks yet. At the second level, Gallego [20] derives a policy which recovers the target schedule at minimal excess over average costs after a single disruption. A disruption may for instance be a machine failure, lack of raw materials, variations in demand or power shortages. The recovery of the target schedule is realized by adjusting the production quantities without altering the production sequence. The size of these adjustments for a product depends in general on the stock levels of all the individual products. In Gallego [21] sufficient conditions are given for a base-stock recovery policy to be optimal. In case of a base-stock policy one only needs to monitor the stock level of the product currently set-up. Several authors (see, e.g., Anupindi and Tayur [7] and Sox *et*
al. [41]) see in these conditions a first step in the derivation of optimality conditions for the SELSP. At the third level, Gallego [20] adds safety stocks in order to efficiently use the control policy, that was shown to be optimal after a single disruption, in a stochastic environment.

Leachman and Gascon [31] develop a so-called dynamic cycle lengths heuristic in a discrete-time model under the assumption of non-stationary demand and deterministic production and set-up times. The first step in their heuristic is the calculation of target cycle lengths in each review period via a deterministic approach by using moving averages of the demand forecasts. The second step is the determination of the operational cycle lengths, which are the minimal reductions of the target cycle lengths for which there is an adequate probability that these modified cycles can be maintained. This reduction is achieved by proportionally reducing the production quantities of all products in a cycle, while maintaining the fixed production sequence. The last step is the possible insertion of an idle period in a cycle, when all products have sufficient stock. Leachman et al. [32] improve the heuristics by increasing the lengths of the operational cycles, i.e., the operational cycles closer to the target cycles, which result in lower costs and improved customer service. Furthermore, the decision rule concerning the insertion of an idle period is refined as well. In a later paper (Gascon et al. [22]), an extensive simulation study on the performance of the dynamic cycle lengths heuristic is undertaken. It is concluded that the performance of the heuristic is satisfactory as long as the load is not extremely high. Moreover, it turns out that the dynamic cycle length heuristic outperforms simple heuristics such as the EMQ rule.

Fransoo et al. [19] study a model for the SELSP under assumptions comparable to those of Leachman and Gascon [31] and Leachman et al. [32]. In particular, it is assumed that demand that cannot be fulfilled from stock is lost. Fransoo et al. [19] show numerically that the performance of the dynamic cycle lengths heuristic proposed by Leachman and Gascon [31] significantly decreases if the load increases. Therefore, an alternative heuristic is developed, that is able to keep the cycle lengths stable. By means of a simulation study, it is shown that this stable cycle length heuristic outperforms the dynamic cycle lengths heuristic when the total demand rate is close to or exceeds the production rate. Fransoo [18] and Fransoo et al. [19] give the following qualitative reason for this result. If one expects that a product will
run out of stock in the forthcoming cycle, the dynamic cycle lengths heuristic will decrease the cycle length. Hence, the relative set-up frequency increases and less capacity is available for production. Consequently, it becomes even more difficult to fulfill future demand.

4.2.2 Local lot sizing policy

Anupindi and Tayur [7] present a fixed production sequence policy under the assumption of backlogging and deterministic production and set-up times. They model a very general demand process, i.e., a non-Markovian compound demand process in which demand arrives for sets of products, and assume state-dependent set-ups. They allow a large number of lot sizing strategies which are all variations of base-stock policies. A simulation-based procedure is developed to obtain the optimal base-stock levels for various performance measures. These performance measures contain not only the traditional product-focused measures, such as costs and service levels based on individual products, but also order-focused measures, like the order response time. Their paper is completed with numerical results for theoretical instants as well as for an industrial application. From these results, it clearly emerges that the popular product-focused performance measures based on costs cannot be used as substitute for order-focused measures.

Wagner and Smits [44] suggest a two-level continuous time model, where at the upper level an optimal fixed cycle schedule with respect to the expected set-up and holding costs is derived. At the lower level a periodic base-stock policy, a so-called \((R, S)\)-policy, is used. The review periods are fixed and determined by the solution at the planning level, while the optimal base-stock levels are obtained by an algorithm developed by Smits et al. [38]. The planning and control levels are then instantaneously optimized by an integrative approach, which uses a Local Search optimization technique.

Vaughan [42] studies a model for the SELSP under the assumption of correlated demand, backlogging, deterministic production and set-up times. His policy is characterized by an exhaustive base-stock policy and a target cycle length. This means that if a cycle is ended within the target length, the machine is idled. If not, the next cycle will commence immediately. Vaughan [42] concludes that demand correlation both increases the variance of the
cycle length and causes correlation between the demand per period and the cycle length. Both effects lead to an higher variance of the total demand during a cycle and, thus, larger safety stock levels are needed compared to the uncorrelated demand case.

The next strategy we describe is introduced by Federgruen and Katalan [14, 15, 16]. Besides the basic assumptions for the SELSP as described in Section 2.1, the following additional assumptions are made in their model: unfilled demand is backlogged, the demand for a product follows a Poisson process and products are produced by an exhaustive or a gated base-stock policy. The production manager is allowed to insert a fixed idle time prior to the set-up for a product in order to reduce the set-up frequencies and, hence, the average set-up costs. Federgruen and Katalan [14, 15, 16] show that the total average costs only depend on the total idle time inserted in a cycle and not on the complete vector of idle times. Hence, a strategy is completely specified by the vector of the base-stock levels and the total amount of idle time in a cycle. A dual polling model is defined by identifying each product with a queue and the demand process of a product with the arrival process at the corresponding queue (see Figure 5). In an earlier paper (Federgruen and Katalan [13]), an efficient algorithm is developed to compute the complete steady-state queue size distributions of this dual polling problem for a given total idle time. The optimal total idle time can be obtained by a numerical procedure. Finally, since the queue size distributions do not depend on the base-stock levels, these base-stock levels can be computed by solving standard newsboy problems. In Federgruen and Katalan [16], their research is completed by the construction of the optimal production sequence.

Although the papers of Federgruen and Katalan [14, 15, 16] are without doubt seminal papers in the field of the SELSP, the introduced strategy has two major drawbacks. First of all, exhaustive, and even gated, base-stock policies do not seem to be the optimal lot sizing policies for the SELSP, since for both policies one high-demand product can dominate the machine resulting in very high cost for the other products. Secondly, the machine incurs a set-up for a certain product even when there is no shortfall for this product, which is of course suboptimal, as mentioned by Sox et al. [41]. Recently, in the literature three extensions of the papers by Federgruen and Katalan have appeared: one developed by Grasman et al. [24],
Figure 5: Dual polling model for 8 products.

one introduced by Krieg and Kuhn [29, 30] and one proposed by Winands et al. [45].

Grasman et al. [24] extend the exhaustive base-stock model of Federgruen and Katalan by adding random yields for the cases of backlogging, lost sales and expediting. In case of backlogging they derive similar newsboy equations in order to obtain optimal base-stock levels. In case of lost sales or expediting they have to resort to a heuristic for computing the (approximate) optimal base-stock levels. Krieg and Kuhn [29, 30] introduce continuous-time models for single-stage multi-product Kanban systems, which are completely identical to the SELSP with lost sales. It is assumed that demands arrive according to mutually independent Poisson processes and that the production and set-up times are exponentially distributed. In Krieg and Kuhn [29], a system is analyzed with state-independent set-ups, whereas in Krieg and Kuhn [30] state-dependent set-ups are modeled, i.e., no set-up for a product is incurred when there is no shortfall. Production quantities are in both models determined by an exhaustive base-stock policy. They decompose the multi-product Kanban system into multiple single-product single-server vacation models. The individual subsystems can be evaluated numerically by an approximate continuous-time Markov chain. Winands et al. [45] study a two-product system with state-dependent set-ups under the assumption of backlogging, in which the first product is produced according to the exhaustive base-stock
policy, whereas the production of the second product follows a quantity-limited policy. They evaluate the system and derive the optimal base-stock levels for a given value of this quantity limit. This paper constitutes a start for the analysis of the quantity-limited base-stock policy with state-dependent set-ups in a system with more than two products.

Finally, it is noteworthy that the strategies of Federgruen and Katalan [14, 15, 16] and the above three extensions are some of the very few policies allowing for an analytical evaluation and optimization.

4.3 Fixed production sequence in combination with a fixed cycle length

In the following subsections we describe the two types of fixed production sequence strategies using a fixed cycle length. That is, production plans using a global lot sizing policy on the one hand and local lot sizing strategies on the other hand.

4.3.1 Global lot sizing policy

In the Master’s thesis of Giezenaar [23], a case study of a chemical plant in the Netherlands is presented, for which a fixed production sequence strategy in combination with a fixed cycle length has been developed under the assumption of deterministic production and set-up times. At the beginning of a cycle the production quantities are determined according to base-stock policies. When scheduling conflicts arise caused by the fixed cycle length, the production quantities are rationed in such a way that the production runs fit in this cycle length.
4.3.2 Local lot sizing policy

Erkip et al. [12] introduce a discrete-time model under the assumption of backlogging, in which the production and set-up times are deterministic. They propose a fixed cycle strategy, where fixed production times are allocated to products and where a base-stock policy is used (see Figure 6). This means that not only the sequence and the total cycle length are fixed, but also the available capacity for each individual product is pre-defined. When the fixed amount of production time has expired, the product is not produced until the next cycle. Furthermore, when the product is on base-stock level before the end of the production interval, one does not switch to the next product in the sequence and thus the machine is idled. Their strategy is modeled as a quasi-birth-death process, which can be solved numerically by the matrix-analytic method. Dellaert [10] mentions some drawbacks of such a fixed cycle strategy. The most important one is that no pooling effect is obtained by the fixed pre-allocation of production capacities to products. For a more detailed description of fixed cycle strategies in the context of make-to-order production situations, see Chapter 3 of Dellaert [10].

5 Discussion and future research

Table 2 summarizes the classification of the literature discussed in Section 4. From this overview some open research areas in the SELSP become evident. In particular, it leads to the following possible topics for further research.

Comparison of the different types of strategies. The first observation that emerges from the preceding section is the fact that most papers focus on the evaluation or optimization of a single class of policies without distinct comparison with other strategies. Instead, research should pay more attention on the comparison between the different policies. By this we mean not only a comparison within a particular class of policies, e.g., fixed production sequence strategies with a local lot sizing policy, but also between the introduced classes.

We admit that this is a rather complicated task for a couple of reasons. First of all, each individual paper uses its own test set to evaluate the performance of the policies. In order
Lot sizing strategy

<table>
<thead>
<tr>
<th>Production sequence</th>
<th>Global</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic</td>
<td>Karmarkar and Yoo [28]</td>
<td>Altiok and Shiue [3], [4], [5]</td>
</tr>
<tr>
<td></td>
<td>Qiu and Loulou [35]</td>
<td>Zipkin [46]</td>
</tr>
<tr>
<td></td>
<td>Sox and Muckstadt [40]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fransoo et al. [19]</td>
<td>Federgruen and Katalan [14], [15], [16]</td>
</tr>
<tr>
<td></td>
<td>Gallego [20], [21]</td>
<td>Grasman et al. [24]</td>
</tr>
<tr>
<td></td>
<td>Gascon et al. [22], [31], [32]</td>
<td>Krieg and Kuhn [29], [30]</td>
</tr>
<tr>
<td></td>
<td>Markowitz et al. [33], [34]</td>
<td>Smits et al. [38], [44]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vaughan [42]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Winands et al. [45]</td>
</tr>
<tr>
<td>Fixed + fixed cycle length</td>
<td>Giezenaar [23]</td>
<td>Erkip et al. [12]</td>
</tr>
</tbody>
</table>

Table 2: An overview of the classification.

to make a fair comparison among the proposed strategies, one needs a generally accepted and widely used test set. Secondly, it is not very likely that a certain policy will outperform the competing strategies in each production environment. Rather, we expect that in some situations a certain policy is superior, but, in other situations, a competitor is more effective. In other words, the performance of a particular strategy is contingent upon the production environment in which it is used. Nevertheless, it is important to identify in which situation a certain policy is effective. Thirdly, some policies can not be implemented in each environment, since this environment may impose all kinds of restrictions on the production plan, e.g., a fixed production sequence or a fixed cycle length. Lastly, strategies may perform well with respect to the total costs, but we have to realize that a cost function can never capture all the relevant information.

Irrespective of the aforementioned difficulties, there is without doubt need for a large-scale (simulation) study comparing the performance of a large number of production plans for the SELSP. This study should contain evaluations of the most important strategies proposed in the literature and a comparison between them as well as some simple heuristics used as benchmarks. Such a study would get us a step nearer to the answering of basal, yet very important, questions like what is the loss of optimality incurred by the fixation of a production sequence and in which settings should a cycle length be stable or when should it be highly
variable? Finally, we emphasize that we certainly do not claim that there is a complete lack of attention in the literature for the comparison between strategies and the answering of the aforementioned questions. However, we believe that the results are still too scattered and should be more structured.

Empirical study. As argued in the preceding paragraph, the performance of a particular production schedule is highly contingent on the production environment in which it is deployed. So far, research has paid insufficient attention to the identification of the characteristics of the environments in which the SELSP is encountered. This observation shows the need for a large-scale empirical study that investigates the main characteristics of these production environments such as the plant characteristics, product characteristics, production process characteristics and demand characteristics.

Dynamic production sequence. By using a dynamic production sequence in combination with a global lot sizing policy the production plan with optimal costs can, in principle, be obtained. However, due to the curse of dimensionality only optimal policies for at most three products have been obtained in the literature. Therefore, we strongly advocate more research aiming at the development of intuitively appealing (near-optimal) dynamic production strategies for the SELSP, which can be applied in situations with a large number of different products.

Fixed production sequence in combination with a dynamic cycle length. As can be seen from Table 2, most of the research concerning the SELSP concentrates on fixed production sequence strategies in combination with a dynamic cycle length. The choice for the analysis of a fixed production sequence is mainly motivated by practice, because a fixed sequence facilitates coordination within the production line. Nevertheless, we have to recall that the assumption of a fixed sequence is often made for the tractability of the analysis as well.

The construction of an optimal fixed production sequence policy using a dynamic cycle length should of course contain the following two steps:
1. Construction of the optimal production sequence, i.e., the order and frequency for the production of the individual products;

2. Determination of the optimal system parameters, e.g., base-stock levels, given this production sequence.

However, in many papers dealing with fixed production sequence strategies the first step is lacking, and the focus is put on the second step. One could argue that this is in line with practice, in which often the production manager indeed has no rights to decide on the production sequence. That is, this sequence is either imposed by the downstream and upstream stages of the production line or by strongly sequence-dependent set-ups. Nonetheless, the lack of the first step in the construction of a fixed production sequence policy seems to be a serious deficiency, since optimizing a system under a non-efficient production sequence is like finding the right answer for the wrong problem. This observation is supported by Bourland and Yano [9] and Federgruen and Katalan [16], who argue and show that the production frequency and order for a product has a serious impact on the variance of the total demand in a cycle and, therefore, on the required safety stock level. We support their conclusion that the production sequence plays a key part in the SELSP and that certainly more research is needed in the construction of efficient production sequences.

Finally, lot sizing decisions in fixed production sequence policies are often made according to exhaustive or gated base-stock policies. As described before, these types of lot sizing strategies possess several drawbacks, such as highly variable cycle lengths and, thus, high safety stock levels. The analysis of more refined policies like the quantity-limited or time-limited base-stock policy provides challenges for further research.

**Fixed production sequence in combination with a fixed cycle length.** Table 2 shows that research on strategies deploying a fixed production sequence in combination with a fixed cycle length is still in a very early stage. Future research on these strategies would, therefore, be desirable.

In particular, as in the preceding paragraph, the optimization of policies deploying a fixed production sequence policy in combination with a fixed cycle length ought to include the
following steps:

1. Construction of the optimal production sequence and of the optimal cycle length;

2. Determination of the optimal system parameters, e.g., base-stock levels, given this production sequence and cycle length.

So far, the literature has shed insufficient light on the first step of the above procedure. Frequently, an optimal deterministic solution is implemented instead, although several authors have concluded before that there is a significant discrepancy between the optimal deterministic sequence and cycle lengths on the one hand and the optimal stochastic counterparts on the other hand (see, e.g., Bourland and Yano [9], Federgruen and Katalan [16] or Sox et al. [41]). Therefore, one should take stochasticity into account when constructing the sequence and cycle length as indicated by Bourland [8] and not straightforwardly implement deterministic solutions. The importance of the cycle length follows from the fact that this length determines not only to a large extent the levels of both the cycle stocks and safety stocks, but also the set-up frequency.

**Impact distributions.** An interesting research topic is the study of the impact of the various probability distributions, i.e., the demand, set-up and production processes, on the performance of the machine and on the structure of the optimal production plan. So far, hardly any results have been obtained on this issue. For instance, how does the variation of the production process affect the performance of the bottleneck? It may be that the cost reduction by optimizing the production plan is negligible in comparison with the reduction that could be obtained by getting more control on the production process. Comparable questions can be asked for the set-up and the demand process.

Especially, for the latter it is often assumed in analytical studies that it follows a Poisson distribution, though this seems not be a realistic assumption from a practical point of view. Therefore, it would be interesting to extend the analytical results obtained under a Poisson assumption to more general demand processes. Furthermore, almost all research on the SELSP assumes that the demand processes for the individual products are mutually independent.
To the best of our knowledge, the only paper addressing demand correlation is by Vaughan [42], in which a simulation study is performed on the impact of a correlated demand on the production schedule. An interesting research topic would be to analyze the effect of demand correlation on the optimal production plan in more detail by means of an analytical study.

**Buffering restrictions.** The usual assumptions in the SELSP literature are that buffer space is infinite and that products are non-perishable. There are some papers that relax the first assumption and that study production environments with limited stock space for each product (see, e.g., Qiu and Loulou [35]). However, no paper studies the SELSP in combination with perishable goods. The limited shelf life of not only the finished products but also of the raw materials asks for more refined production plans than the ones employed in the standard SELSP. For example, the safety stock levels of the perishable products and the total cycle length should be bounded. For a literature overview of combined make-to-order and make-to-stock production situations for food processing companies, in which products are highly perishable, see Soman et al. [39]. In conclusion, we believe that both extensions, limited stock space and perishable goods, are worth studying.

**Local lot sizing strategy incorporating global information.** In the present paper we have distinguished between local lot sizing strategies and the global variants. We now want to discuss a fixed production sequence strategy deploying a local lot sizing strategy whilst at the same time incorporating (partial) information of the rest of the system. This strategy originates from the *fiber distributed data interface* (FDDI) protocol used in optical local area networks (see, e.g., Altman and Liu [6]). We assume that a base-stock policy is implemented. Furthermore, each individual product possesses its own timer. When production of a tagged product ends, the corresponding timer is started. At the next production start of this product, the maximal production time is determined by the difference between the pre-defined fixed cycle length and the value on this timer. This means that when the machine starts production of the tagged product, it will continue production until either the base-stock level has been reached or the timer has ended, whichever occurs first. By doing so, the cycle lengths are bounded, and, thus, the total production capacity as well. Now, mix flexibility is mainly used
to react to the stochastic environment. As seen before in Fransoo [18] and Fransoo et al. [19], these characteristics of a policy in general lead to low costs and high service levels, especially in the case of a high demand level. Although the lot sizing decisions are thus made only on local information, i.e., the stock level and timer of the product currently set-up, the timer contains some global information of the rest of the products. Analysis of the performance of this policy in the context of the SELSP would be highly interesting.

Acknowledgement

The authors have strongly benefited from many discussions with Onno Boxma and Ton de Kok. Furthermore, we like to thank Will Bertrand, Nico Dellaert, Nesim Erkip, Simme Douwe Flapper and Jan Fransoo for their valuable comments. Doug Down is acknowledged for pointing our attention to the strategy originating from the FDDI protocol as described in Section 5 of the present paper.

References


