Abstract Number: 025-0574

Controlled Order Release (OR) in Two-Level Multi-Stage Job Shops: An Assessment by Simulation

Name: Matthias Thürer
Institution: Federal University of São Carlos
Address: Department of Industrial Engineering
Rodovia Washington Luís, km 235
13565905 - São Carlos – Brasil
Email: mailboxmatthiasthurer@gmail.com

Name: Dr. Mark Stevenson
Institution: Lancaster University
Address: Department of Management Science
Lancaster University Management School
Lancaster University
LA1 4YX - U.K.
Email: m.stevenson@lancaster.ac.uk

Name: Dr. Cristovao Silva
Institution: CEMUC - University of Coimbra
Address: Mechanical Engineering Department
Pólo II Pinhal de Marrocos
3030 Coimbra – Portugal
Email: cristovao@gestao.dem.uc.pt

Name: Dr. Martin J. Land
Institution: University of Groningen
Address: Department of Operations
Faculty of Economics and Business
University of Groningen
9700 AV Groningen – The Netherlands
Email: m.j.land@rug.nl

Name: Dr. Moacir Godinho Filho
Institution: Federal University of São Carlos
Address: Department of Industrial Engineering
Rodovia Washington Luís, km 235
Caixa Postal 676 – Monjolinho
13565-905, São Carlos, SP, Brasil
Email: moacir@dep.ufscar.br

POMS 23rd Annual Conference
Chicago, Illinois, U.S.A.
April 20 to 23, 2012
Controlled Order Release (OR) in Two-Level Multi-Stage Job Shops: An Assessment by Simulation

Abstract
Most research which assesses the performance of Workload Control (WLC) order release methods assumes products have simple structures. But, in practice, products are often complex and consist of a number of sub-assemblies which flow through a ‘level 1’ job shop before converging on several final assembly operations in a ‘level 2’ assembly shop. Evaluating the performance of WLC order release methods in this context - referred to as a ”two-level multi-stage job shop” - is an important step towards improving the alignment between WLC theory and practice. We use simulation to assess the performance of four of the best-performing WLC order release methods in two-level multi-stage job shops. Results suggest that WLC order release has the potential to limit Work-In-Process (WIP) while reducing throughput times and the percentage of tardy jobs, thus improving overall performance. But it is important to consider: (i) when the assembly order should be considered for release to level 2 – as soon as the first sub-assembly in the routing of an order is complete at level 1, or only once all sub-assemblies are complete; and, (ii) where order release should be controlled – at level 1 only, or at both level 1 and 2. Results suggest that: (i) orders should be considered for release to level 2 as soon as the first sub-assembly in the routing of an order is complete; and, (ii) in general, controlling release at both level 1 and level 2 leads to a greater reduction in the percentage of tardy jobs.

1. Introduction
This paper uses simulation to assess the performance of Workload Control (WLC) order release methods in two-level multi-stage job shops, as commonly encountered in small to medium sized Make-to-Order (MTO) companies in practice (Portioli-Staudacher, 2000; Stevenson & Silva, 2008).
In this context, final products or assembly orders consist of multiple sub-assemblies. Sub-assemblies progress through the ‘level 1’ job shop independently before converging on the ‘level 2’ assembly shop, where a series of assembly operations have to be completed for the sub-assemblies to be converted into the final product or assembly order. The level 2 assembly shop is fed with sub-assemblies from the preceding level 1 job shop. Thus, Due Date (DD) adherence in multi-stage job shops is dependent on the timely progress through level 1 of all sub-assemblies that make up the final product or assembly order.

WLC is a Production Planning and Control (PPC) concept which simultaneously controls lead times, capacity and Work-In-Process (WIP) on the shop floor, integrating production and sales into a hierarchical system of workloads which buffers throughput against variance (Tatsiopoulos & Kingsman, 1983; Kingsman et al., 1989). One of its key control levels is order release which decouples the shop floor from higher planning levels (e.g. for managing customer enquiries). Jobs are not released immediately to the shop floor but held in a pre-shop pool of orders from which they are released onto the shop floor in time to meet DDs whilst keeping workload levels (i.e. Work-In-Process, or WIP) within limits or norms.

While a broad literature on order release methods in job shops exists (e.g. Land & Gaalman, 1998; Sabuncuoglu & Karapinar, 1999; Fredendall et al., 2010), it is generally assumed that products have simple structures and do not require (multiple) assembly operations to convert several sub-assemblies into a final product. In contrast, the order release literature which does consider assembly is scarce. The few existing studies (Portioli-Staudacher, 2000; Lu et al., 2010; Thürer et al., 2012) assume only one final assembly operation whereas, in practice, often a series of operations have to be undertaken to assemble the final product (e.g. Portioli-Staudacher, 2000; Silva et al. 2006). Thus, the literature does not provide answers to important questions regarding which method to apply when there is more than one final assembly operation. Moreover, research has not investigate when and where order release should be controlled in this context, i.e. whether to considered assembly orders
for release to level 2 as soon as the first sub-assembly in the routing of an order is complete at level 1, or only once all sub-assemblies are complete; and whether to control release at level 1 only, or at both level 1 and 2.

Consequently, researchers implementing WLC have found it difficult to apply existing theory (e.g. Silva et al. 2006; Hendry et al., 2008; Stevenson & Silva, 2008; Stevenson et al., 2011) and one of the most important contemporary WLC research problems concerns how WLC theory can be extended to assembly and multi-stage job shops (e.g. Hendry et al., 2008; Stevenson et al., 2011; Thürer et al., 2011). In response, this study:

- Determines which WLC order release method should be applied in multi-stage job shops in practice. This is based on comparing the performance of a set of four release methods, identified as best-performing from the literature, in two-level multi-stage job shops as commonly encountered in small to medium sized MTOs in practice.
- Investigates when assembly orders should be considered for release to level 2, i.e. as soon as the first sub-assembly in the routing of the assembly order is complete at level 1, or only when all sub-assemblies which make up an assembly order are complete.
- And, evaluates where order release should be applied, i.e. should order release only be controlled in the level 1 job shop? Or, in both the level 1 job shop and the level 2 assembly shop?

The remainder of this paper is organised as follows. Literature on order release in job shops and assembly shops is reviewed in Section 2 to identify the best-performing release methods from the literature to be included in this study. The simulation model is then outlined in Section 3 before the simulation results are presented in Section 4. Final conclusions, including managerial implications and future research directions, are provided in Section 5.
2. Literature Review

This review is structured as follows. First, Section 2.1 provides a brief review of WLC order release methods to identify a set of best-performing release methods from the literature to be included in our study. Then, the limited research on release methods in assembly shops is reviewed in Section 2.2.

2.1 WLC Order Release Methods

WLC order release methods control the amount of work on the shop floor. For this study, release methods are divided into those which release work from the pre-shop pool periodically and those which release work continuously. The two approaches are discussed in sections 2.1.1 and 2.1.2, respectively before a unique approach - which combines periodic with continuous release - is outlined in Section 2.1.3. In reviewing the order release methods, the four best-performing methods are identified for inclusion in this simulation study. For other approaches to classifying release methods, the reader is referred to Philipoom et al. (1993), Wisner (1995), Bergamaschi et al. (1997), Sabuncuoglu & Karapinar (1999) and Fredendall et al. (2010).

2.1.1 Periodic Release Methods

Periodic release methods make the decision to release orders at periodic time intervals, e.g. once a shift, day or week. The release procedure is similar among alternative methods (Land & Gaalman, 1998): (i) jobs in the pool are considered for release periodically according to a pool selection rule, such as Earliest Due Date (EDD) or Planned Release Date (PRD); (ii) the workload of a job is contributed to the load of the work centres in its routing; (iii) if this new load fits within the workload norm, the job is released and its load assigned; (iv) if one or more norms would be exceeded, the job must wait until at least the next release period and the workload is reset. This procedure is repeated until all jobs have been considered for release once.

The main difference between alternative periodic release methods is how the load is contributed. The main approaches are the probabilistic approach, which estimates the input to the direct load (or
queue) of a work centre over time and converts the indirect load contributed at release (the load upstream of a work centre) using a depreciation factor based on historical (probabilistic) data (see, e.g. Bechtle, 1988 and 1994); and, the ‘classical’ aggregate load approach (also known as the ‘atemporal’ approach), which does not consider the position of a work centre in the routing of a job: the direct and indirect load is simply aggregated (e.g. Bertrand & Wortmann, 1981; Hendry, 1989). An extension to the classical aggregate load approach was presented by Oosterman et al. (2000) and is known as the “corrected aggregate load approach”. This estimates the input to the direct load - like the probabilistic approach - but in a much simpler way. The workload which is contributed to a certain work centre is the corresponding workload divided by the position of a work centre in the routing of a job.

Of all the available periodic release methods presented in the literature, the corrected aggregate load and the probabilistic approaches have performed the best (see, e.g. Oosterman et al., 2000). Here, it is argued that the corrected aggregate load approach is the best solution for implementation in practice due to its simplicity. Therefore, this release method is selected to represent periodic release methods in this study.

2.1.2 Continuous Release Methods

Using a continuous release method means that the decision to release an order may be made at any moment in time, rather than periodically. Release is triggered by an event on the shop floor, such as: (i) the load of the bottleneck constraint falling below a predetermined level (see, e.g. Glassey & Resende, 1988; Enns & Prongue Costa, 2002); (ii) the load of any work centre falling below a predetermined level (e.g. Melnyk & Ragatz, 1989): or (iii) the load of the shop as a whole falling below a predetermined level (e.g. Melnyk & Ragatz, 1989; Qi et al., 2009). As soon as a release is triggered, jobs are chosen from the pre-shop pool according to a selection rule, e.g. EDD, PRD or Work-IN-Queue (WINQ – Melnyk & Ragatz, 1989) if the whole shop load is considered, until the workload exceeds the release trigger level. Melnyk & Ragatz (1989) compared the performance of a
continuous release method triggered by the work centre load with the performance of a shop load method and identified the former as the best; this result was later confirmed by Hendry & Wong (1994) and Sabuncuoglu & Karapinar (1999). As we focus on the job shop, where no stable bottleneck or gateway work centre can be identified, the bottleneck workload trigger is not applicable. Therefore, the work centre workload trigger is selected to represent continuous release methods in this study.

In addition, a second continuous release method is included in this study: the Superfluous Load Avoidance Release (SLAR) method for which outstanding performance results have been reported (see Land & Gaalman, 1998). SLAR releases work under two conditions: (i) a starving work centre; and (ii) no urgent jobs queuing in front of a work centre (but urgent jobs waiting in the pre-shop pool). In the first case, a job for which the first work centre in its routing is the starving work centre is selected from the pre-shop pool according to the PRD rule. In the second case, an urgent job for which the triggering work centre is the first work centre in its routing is released according to the Shortest Processing Time (SPT) rule. In contrast to the other release methods discussed above, SLAR bases its release decision on the urgency of jobs rather than on balancing the workload; it differentiates between urgent jobs (i.e. jobs for which the planned operation start time has passed) and non-urgent jobs (i.e. jobs for which the planned operation start time has not yet passed). The planned operation start time is given by the DD minus the sum of the remaining processing times and the remaining number of operations multiplied by a time-related slack factor $k$. As a result, the performance of SLAR depends largely on $k$.

2.1.3 Periodic and Continuous Release

An approach which combines periodic with continuous release, known as “LUMS OR” (the Lancaster University Management School Order Release rule), was presented by Hendry & Kingsman (1991). LUMS OR combines periodic release with a continuous starvation avoidance mechanism; thus, release is as for the other periodic release methods above (see Section 2.1.1) but if,
at any time, the workload in front of a work centre falls to zero (i.e. the work centre is starving), a job is actively pulled forward from the pre-shop pool. This means a job with the work centre which triggered the release as the first in its routing is released from the pool according to PRD and its load is contributed according to the approach applied by the periodic release method. A corrected version of LUMS OR, and hereafter referred to as “LUMS COR”, will be applied in this study. While the original LUMS OR method incorporated the ‘classical aggregate load approach’ to workload accounting over time, LUMS COR replaces this with the ‘corrected aggregate load approach’, given its superior performance (e.g. Oosterman et al., 2000; Thürer et al., 2011).

2.2 WLC Release and Assembly Shops
To the best of our knowledge, there are only three studies concerned with controlled order release in assembly shops: Portioli-Staudacher (2000), Lu et al. (2010) and Thürer et al. (2012). All three focused on the assembly job shop, i.e. only considered one final assembly operation rather than a series of assembly operations for the sub-assemblies to be converted into the final product.

Portioli-Staudacher (2000) applied periodic release using the probabilistic approach but did not compare results against immediate release or any other release method. Thus, the performance effect compared to alternative methods and simple dispatching remains unknown. Lu et al. (2010) also applied a periodic release method, this time the classical aggregate load approach. The authors showed that throughput time and the Mean Absolute Deviation of lateness (MAD) can be reduced compared to immediate release. However, as in Portioli-Staudacher (2000), performance was not compared to alternative release methods and Portioli-Staudacher’s (2000) results for the probabilistic approach cannot be directly compared with Lu et al.’s (2010) results for the classical aggregate load approach. Finally, Thürer et al. (2012) assessed the performance of LUMS COR showing that it outperforms immediate release; however, no comparison with periodic release, as applied by Portioli-Staudacher (2000) and Lu et al. (2010), or any other release method, was provided.
In conclusion, while a broad literature on order release in job shops exists, literature on order release in assembly shops is scarce. The available literature is restricted to assembly job shops where it is assumed that there is only one final assembly operation. There have been no attempts to compare the performance of alternative order release methods or to consider two-level multi-stage job shops. Thus, implementers in these contexts must select an order release method without any guidance. Moreover, when to consider orders for release to the level 2 assembly shop and where to control release (at the level 1 job shop or at both the level 1 job shop and level 2 assembly shop) has been neglected. As a result, researchers and practitioners have found it difficult to successfully implement existing theory in practice (e.g. Silva et al. 2006; Hendry et al., 2008; Stevenson & Silva, 2008; Stevenson et al., 2011).

3. Simulation

From the above, it follows that one of the most important outstanding WLC research questions is as follows:

*Which release method should be implemented in multi-stage job shops in practice?*

In response, the performance of four of the best-performing release methods identified from the literature is assessed in a two-level multi-stage job shop. In addition, and based on the results, when to consider an assembly order for release to the second level assembly shop (once all its sub-assemblies are complete or once the first sub-assembly in its routing is complete) and where to control release (i.e. at level 1 or at both level 1 and level 2) are addressed. Discrete event simulation is applied as it allows the gap between analytical research and practice to be bridged if the tractability of the problem is restricted, e.g. by feedback loops (Bertrand & Fransoo, 2002). The simulation model implemented is described in Section 3.1 before the release methods and subsequent
dispatching rule are detailed in Section 3.2 and Section 3.3, respectively. Finally, job characteristics are given in Section 3.4 and the experimental design is outlined in Section 3.5.

3.1 Overview of Shop Characteristics

A model of a two-level multi-stage job shop has been developed using the SimPy © module of Python ©. This model has been chosen as a generalisation of shop structures typically found in practice (e.g. Portioli-Staudacher, 2000). All of the sub-assemblies of an assembly order are produced in the level 1 job shop; finished sub-assemblies proceed to the subsequent level 2 assembly shop, where they are assembled into the final product through a series of predetermined assembly operations. The two levels are further described in what now follows below.

- **Level 1:** The level 1 job shop represents a pure job shop (Melnyk & Ragatz, 1989), i.e. the flow is undirected and completely random. The job shop contains six work centres, where each is a single and unique capacity resource. The number of operations per sub-assembly is equally distributed between one and six. All work centres have an equal probability of being visited and a particular work centre is required at most once in the routing of a sub-assembly.

- **Level 2:** The level 2 assembly shop represents a general flow shop (Oosterman et al. 2000). This shop is similar to the pure job shop but the routing vector, i.e. the sequence in which work centres are visited, is sorted. Thus, the flow is fully directed and assembly operations at upstream work centres always precede assembly operations at downstream work centres. The shop consists of six assembly work centres, where each is a single and unique capacity resource. All assembly work centres have an equal probability of being the assembly work centre where a certain set of sub-assemblies is assembled into an assembly order. A particular work centre is required at most once in the routing of an assembly order, meaning each assembly work centre only undertakes one assembly operation per assembly order. The number of sub-assemblies per assembly order is uniformly distributed between one and six.
The resulting job shop model is summarised in Figure 1. The probability that a job enters the shop floor, leaves the shop floor or moves to a certain work centre is indicated by the strength (or thickness) of an arrow. All sub-assemblies leaving the level 1 job shop move to the level 2 assembly shop.

To illustrate the multi-stage job shop, consider the following example. An assembly order is made up of three sub-assemblies; it therefore requires three assembly operations (two plus one – the first – which is considered preparatory work, i.e. setting up the assembly), which take place at three different work centres. The sequence in which assembly work centres are visited by the assembly order is predetermined and directed, i.e. there are typical upstream and typical downstream work centres. When the assembly order arrives at the multi-stage job shop, all three of its sub-assemblies are either released immediately (immediate release) to the level 1 job shop or are considered for release individually (if a release method is applied). Once sub-assemblies are completed, the assembly order is either released immediately to the level 2 assembly shop or considered for release individually according to a certain control policy (as will be described in Section 3.2.1 below). Precedence has to be considered whilst the assembly order progresses in the level 2 assembly shop.
Thus, a certain assembly operation cannot take place if the corresponding sub-assembly is not finished in the preceding level 1 job shop - the assembly order has to wait for the completion of the sub-assembly. The assembly order is complete and leaves the simulation once its final assembly operation is complete.

3.2 Controlled Order Release

As in previous studies, e.g. Land & Gaalman (1998) and Thürer et al. (2012), it is assumed that all orders are accepted, materials are available and all necessary information regarding routing sequence, processing time, etc is known. Sub-assemblies to the level 1 job shop and assembly order to the level 2 assembly shop are either released immediately (Immediate Release: IMM) or flow into a pre-shop pool to await release. Four different release methods are applied (see Table 1): corrected aggregate load approach (Periodic); WCPRD – Work Centre workload trigger Planned Release Date selection (Continuous); Superfluous Load Avoidance Release (SLAR); and, LUMS COR (Periodic and Continuous). These release methods were identified as the best-performing from previous research (see Section 2.1).

Table 1: Summary of Order Release Rules Applied in this Study

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name of Rule</th>
<th>Classification</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodic</td>
<td>Corrected Aggregate Load</td>
<td>Periodic Order Release Method</td>
<td>Releases jobs periodically up to the workload norm (upper bound).</td>
</tr>
<tr>
<td>LUMS COR</td>
<td>Lancaster University Management School (LUMS) Corrected Order Release</td>
<td>Periodic &amp; Continuous Order Release Method</td>
<td>Combines periodic with continuous release. Jobs are pulled onto the shop floor in-between periodic reviews if a work centre is starving.</td>
</tr>
<tr>
<td>SLAR</td>
<td>Superfluous Load Avoidance Release</td>
<td>Continuous Order Release Method</td>
<td>Releases jobs if a work centre is starving or there are no urgent jobs queuing in front of a work centre.</td>
</tr>
<tr>
<td>WCPRD</td>
<td>Work Centre workload trigger Planned Release Date</td>
<td>Continuous Order Release Method</td>
<td>Releases jobs if the direct load of any work centre falls below a predetermined level (lower bound).</td>
</tr>
</tbody>
</table>

Results for Periodic Release have been obtained by loosening the workload norm from 9 to 13 time units (5 steps); results for LUMS COR by loosening the workload norm from 8 to 12 time units.
results for WCPRD by loosening the workload trigger from 5 to 9 time units (5 steps); and, 
results for SLAR by loosening the slack factor $k$ from 2 to 6 time units (5 steps). The check period, 
i.e. the period between releases for Periodic Release and the periodic part of LUMS COR is set to 4 
time units. The Planned Release Date (PRD) for selecting jobs from the pool is given by the Planned 
operation Start Time (PST) of the first operation in the routing of a job (as will be defined in 
Subsection 3.3 on dispatching below). All parameters have been identified as best-performing in 
preliminarily simulation experiments.

3.2.1 Control Policies

There are two main factors which influence at which degree order release is implemented and control 
exercised in a two-level multi-stage job shop: (i) when control is exercised, i.e. whether assembly 
orders are considered for release as soon as the first sub-assembly in their routing is finished at level 
1 or only once all corresponding sub-assemblies are finished; and (ii) where control is exercised, i.e. 
whether release control is exercised at only one level (i.e. before level 1) or at both levels (i.e. before 
level 1 and 2). To reflect this, four different control strategies are applied in this study as follows (see 
also Table 2):

- **Control Policy I:** An assembly order enters the level 2 assembly shop immediately once the first 
  sub-assembly in its routing is finished. Level 1 and level 2 are not separated and release control is 
  only exercised at the level 1 job shop.

- **Control Strategy II:** An assembly order enters the level 2 assembly shop immediately once all of 
  its sub-assemblies are finished. Level 1 and 2 are separated and release control is only exercised at 
  the level 1 job shop.

- **Control Policy III:** This control policy is similar to Policy I, but release control is exercised at 
  both the level 1 job shop and the level 2 assembly shop. An assembly order is considered for 
  release to the level 2 assembly shop once the first sub-assembly in its routing is finished.
• **Control Policy IV**: This control policy is similar to Policy II, but release control is exercised at both the level 1 job shop and the level 2 assembly shop. An assembly order is considered for release to the level 2 assembly shop once all of its sub-assemblies are finished.

Table 2: Control Policies I-IV

<table>
<thead>
<tr>
<th>Consider Assembly Job for Release When:</th>
<th>Where: Release Control at Level 1 Job Shop</th>
<th>Level 1 Job Shop &amp; Level 2 Assembly Shop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st sub-assembly completed</td>
<td>Control Policy I</td>
<td>Control Policy III</td>
</tr>
<tr>
<td>All sub-assemblies completed</td>
<td>Control Policy II</td>
<td>Control Policy IV</td>
</tr>
</tbody>
</table>

Finally, if release control is exercised at both the level 1 job shop and level 2 assembly shop then the parameters for the release methods (i.e. workload norms, workload triggers or slack factors) are equal at both levels to maintain the number of necessary experiments at a reasonable level.

### 3.3 Dispatching Rules

The dispatching rule applied is PST; like the PRD selection rule, it is an integral part of the SLAR method and has interacted well with other WLC release methods in previous studies (e.g. Land & Gaalman, 1998). Moreover, it is similar to the Operation Completion Date (OCD) rule which was widely applied in previous research on assembly shops (e.g. Adam *et al.*, 1993; Smith *et al.*, 1993; Bertrand & van de Wakker, 2002). The job with the earliest PST, given by the DD minus the remaining total processing time and the number of remaining operations multiplied by a slack parameter $k$, is selected (see Equation (1)).

\[
PST = \text{Due Date} - (\text{Remaining Total Processing Time} + \text{Remaining Operations} \times k) \tag{1}
\]

First, the PSTs for the assembly operations are determined. Then, starting from the corresponding PST of each assembly operation (the sub-assembly DD), the PSTs of each sub-assembly are determined. For all experiments, except those concerned with SLAR, $k$ is set to 4 time units in the
level 1 job shop and 5 in the level 2 assembly shop. These values have been determined as best-performing in preliminary simulation experiments. For SLAR, as in Land & Gaalman (1998), $k$ for the PST dispatching rule is equal to the slack factor.

3.4 Job Characteristics and Due Date setting procedure

DDs are set exogenously by adding a random allowance factor to the job entry time. The random allowance factor follows a normal distribution with a mean of 110 time units and a standard deviation of 10 time units. This values have been set arbitrarily thus the percentage of tardy assembly orders is 20% under immediate release and no separation of level 1 and level 2.

Operation processing times (in the level 1 job shop) follow a truncated 2-Erlang distribution with a mean of 1 time unit and a maximum of 4 time units. To obtain the assembly times in the level 2 job shop, a constant of 2.5 time units is added to the truncated 2-Erlang distribution (with a mean of 1 time unit and a maximum of 4 time units). Thus, the utilisation levels in the level 1 job shop and the level 2 assembly shop are equal. The arrival rate of assembly orders follows an exponential distribution with a mean of 2.27 time units, thus the utilisation level is 90%. Tables 3 and 4 summarise the simulated shop and job characteristics, respectively.

**Table 3: Summary of Simulated Shop Characteristics**

<table>
<thead>
<tr>
<th>Shop Type</th>
<th>Multi-Stage Job Shop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routing Variability</td>
<td>Random routing, no re-entrant flows</td>
</tr>
<tr>
<td>No. of Work Centres (Job Shop)</td>
<td>6</td>
</tr>
<tr>
<td>No. of Work Centres (Assembly Shop)</td>
<td>6</td>
</tr>
<tr>
<td>Interchange-ability of Work Centres</td>
<td>No interchange-ability</td>
</tr>
<tr>
<td>Work Centre Capacities</td>
<td>All equal</td>
</tr>
</tbody>
</table>
Table 4: Summary of Simulated Job Characteristics

<table>
<thead>
<tr>
<th>No. of Work Orders per Assembly Job</th>
<th>Discrete Uniform[1, 6]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Operations per Sub-Assembly</td>
<td>Discrete Uniform[1, 6]</td>
</tr>
<tr>
<td>Operation Processing Times</td>
<td>Truncated 2–Erlang, $\mu = 1$, max = 4</td>
</tr>
<tr>
<td>Assembly Times</td>
<td>2.5 + a; a Truncated 2–Erlang, $\mu = 1$, max = 4</td>
</tr>
<tr>
<td>Inter-Arrival Times</td>
<td>Exp. Distribution, mean = 2.27</td>
</tr>
<tr>
<td>Set-up Times</td>
<td>Not Considered</td>
</tr>
<tr>
<td>Due Date Determination</td>
<td>Job entry time + a; a normal, mean = 110, $\sigma = 10$</td>
</tr>
<tr>
<td>Complexity of Product Structures</td>
<td>Simple dependent product structures</td>
</tr>
</tbody>
</table>

3.5 Experimental Design & Performance Measures

The experimental factors are summarised in Table 5. The experiments are full factorial for the four release methods (5 levels each), immediate release and the 4 control policies. Each cell is replicated 100 times. Results are collected over 10,000 time units. The warm-up period is set to 3,000 time units. These parameters allow us to obtain stable results whilst keeping the simulation run time at a reasonable level.

Table 5: Summary of Experimental Factors

<table>
<thead>
<tr>
<th>Control Policy (4 level)</th>
<th>Control Policy I - IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled Order Release (21 levels)</td>
<td>Periodic (Norm = 9, 10, 11, 12, 13); LUMS COR (Norm = 8, 9, 10, 11, 12); WCPRD (Trigger Level = 5, 6, 7, 8, 9); SLAR ($k = 2, 3, 4, 5, 6$); and, Immediate Release</td>
</tr>
</tbody>
</table>

The performance measures considered in this study are:

- **The assembly lead time**: The assembly order completion date (at level 2) minus the assembly order entry date (at level 1).
- **The percentage tardy**: The percentage of assembly orders which are tardy. The Mean Absolute Deviation of lateness (MAD), as applied by Lu et al. (2010) is not considered, as it would give no indication of the tardiness of orders: an order completed 10 time units early would have the same MAD as an order completed 10 time units late.
- **The level 1 throughput time (or sub-assembly manufacturing lead time):** The sub-assembly completion date minus the sub-assembly release date.

- **The level 2 throughput time (or assembly manufacturing lead time):** The assembly order completion date minus the date when the assembly order was released to the level 2 assembly shop.

Finally, the significance of the differences between the outcomes of individual experiments has been verified by paired t-tests which comply with the use of common random number streams to reduce variation across experiments. Whenever we discuss a difference in outcomes between two experiments, the significance can be proven by a paired t-test at a level of 97.5%.

### 4. Results

The presentation of results is structured as follows. The performance of the different release methods under the different control policies (I-IV) is first assessed in Section 4.1 before a final discussion of results to evaluate the release methods is presented in Section 4.2. Results are presented in the form of performance curves, where the assembly lead time (Section 4.1.1), the level 2 throughput time (Section 4.1.2) and the percentage tardy (Section 4.1.3) are each presented against the level 1 throughput time. The right-hand starting point of each curve represents the tightest workload norm level (for Periodic release and LUMS COR), workload trigger (for WCPRD) or slack factor (for SLAR).

#### 4.1.1 Performance of Release Methods: Assembly Lead Time

The performance curves obtained for the assembly lead time over the level 1 throughput time for control policies I & II are given in Figure 2a. The corresponding performance curves for control policies III & IV are given in Figure 2b.
The following can be observed from the results:

- **Which method to apply:** Figure 2a and Figure 2b illustrates that controlling order release significantly reduces the level 1 throughput time and thus the WIP in the level 1 job shop compared to immediate release. However, this is at the expense of an increase in the assembly lead time. Only LUMS COR maintains an assembly lead time equivalent to immediate release. LUMS COR and SLAR strike the best balance between the assembly lead time and level 1 throughput time.
throughput time performance. WCPRD performs the worst compared with the other three release methods and immediate release.

- When to consider assembly orders for release: From Figure 2a and Figure 2b it can be observed that Control Policy I outperforms Control Policy II and Control Policy III outperforms Control Policy IV. As expected, the assembly lead time increases if assembly orders are only considered for release to level 2 when all corresponding sub-assemblies have been completed (i.e. Control Policy II & IV).

- Where to exercise release control: When release is controlled at both levels, shorter assembly lead time results are obtained for Control Policy IV compared to Control Policy II (i.e. the assembly order is released to level 2 when all of its subassemblies are complete) if SLAR or LUMS COR is applied. For all other scenarios, the lead time remains equivalent or even deteriorates. The deterioration in performance is especially evident for WCPRD and Periodic Release. Lead time results deteriorate for both Control Policy III, when compared to Control Policy I, and for Control Policy IV, when compared to Control Policy II. However, this performance deterioration may be offset if a shift in WIP from the level 2 shop floor to the pre-shop pool in front of level 2 – which would increase inventory in front of the assembly shop – is perceived as beneficial by management. Thus, the decision at which level to control release is partly dependent on existing space restrictions within a company.

4.1.2 Performance of Release Methods: Level 2 Throughput Time

The shift in WIP from the level 2 shop floor to the pre-shop pool in front of level 2 can be observed from Figure 3a and 3b, where the results for the level 2 throughput time are presented for control policies I & II and control policies III & IV, respectively.
Figure 3: Level 2 Throughput Time over Level 1 Throughput Time: (a) Control Policy I and II; and (b) Control Policy III and IV

In more detail it can be observed:

- **Which method to apply:** As for the assembly lead time, LUMS COR and SLAR can be considered the best-performing methods because they strike the best balance between level 1 and level 2 throughput time under all experimental conditions.

- **When to consider assembly orders for release:** From Figure 2a and Figure 2b it can be observed that releasing an assembly order once the first sub-assembly in its routing has been completed in the level 1 job shop (Control Policy I & III) leads to a longer level 2 throughput time compared to releasing the assembly order once all of its sub-assemblies have been completed (Control Policy...
II & IV). The reason for this is the precedence (or the logical assembly sequence) which has to be followed when assembling an order. In other words, the difference between the performance curves for Control Policy I compared to Control Policy II and Control Policy III compared to Control Policy IV can be interpreted as the time which the assembly order spends waiting on the level 2 shop floor until its sub-assemblies are complete in the level 1 job shop. The strong difference in the pattern of the performance curves between Control Policy III and Control Policy IV for Periodic release and WCPRD (Figure 3b) indicates a high standard deviation of lateness of sub-assemblies caused by those two release methods. However, in general, the waiting time on the level 2 shop floor due to precedence is less than the time an assembly order would have to wait in front of the level 2 assembly shop for all of its sub-assemblies to be completed. This leads to better performance in terms of assembly lead time for Control Policy I & III compared to Control Policy II & IV, as is evident from Figure 2a and 2b, respectively (previously presented).

- **Where to exercise release control:** As may be expected, controlling release at the level 2 assembly shop (control policies III & IV) significantly reduces level 2 throughput time compared to control policies I & II where release to the level 2 assembly shop is not controlled.

**4.1.3 Performance of Release Methods: Percentage Tardy**

To conclude our presentation of results, results for the percentage tardy are provided in Figure 4a and 4b for control policies I & II and control policies III & IV, respectively. The performance curves show a similar pattern to the assembly lead time results. As DDs are set exogenously, the assembly lead time and percentage of tardy assembly orders are inter-related. The main difference is that, despite the fact that LUMS COR performed better in terms of lead time, it does not outperform SLAR in terms of percentage tardy. This indicates a better performance in terms of the standard deviation of lateness for SLAR compared to all of the alternative release methods tested in this study.
Figure 4: Percentage Tardy over Level 1 Throughput Time: (a) Control Policy I and II; and (b) Control Policy III and IV

4.2 Discussion: Comparison of Release Methods

To definitively compare the four release methods, and answer our research question, performance measures have been classified into two categories: Category 1 considers performance using ‘traditional’ measures, i.e. in terms of throughput time performance, WIP, and reductions in percentage tardy; and, Category 2 considers practical issues, including the simplicity of the method, how intuitive it is, and its ease of implementation. Performance in each category is described below before an overall assessment of the release methods is provided:
• **Category 1** - traditional performance measures: SLAR and LUMS COR perform best in terms of percentage tardy across all scenarios. Moreover, SLAR achieves the highest reduction in level 1 throughput time. SLAR can therefore be considered the method which strikes the best balance between reduced throughput time (i.e. reduced WIP) and percentage tardy. It is followed by LUMS COR. Periodic and WCPRD both perform poorly and do not outperform immediate release. The reason for this is the high standard deviation of lateness which results in long assembly lead times especially if release is controlled at both job shop levels (i.e. control policies III & IV). Therefore, for these two release methods, the best performance is achieved by only controlling the level 1 job shop.

• **Category 2** - practicality: LUMS COR, WCPRD and Periodic release are considered the easiest to implement. Moreover, these release methods use a distinct workload level – which is not specified for SLAR – that can be useful for maintaining clear dialog between different tiers of command in a company, e.g. between the shop floor supervisor & operators and between the supervisor & planning officer. SLAR may be considered the most difficult to implement as it is not as simple and intuitive as the other methods.

In contrast to Portioli-Staudacher (2000) and Lu *et al.* (2010), where periodic release was applied in assembly job shops (i.e. only one assembly operation), here it is argued that periodic release methods may be counterproductive and lead to performance deterioration compared to immediate release if applied in two-level multi-stage job shops (i.e. a set of assembly operations). The same is valid for WCPRD. LUMS COR and SLAR are considered the best overall options for implementation in this environment. Both achieve significant performance improvements compared to immediate release. Which method to choose for implementation in practice ultimately depends on whether the simplicity of LUMS COR over SLAR is perceived to offset the loss in level 1 throughput time reduction.
These conclusions can also be regarded as valid for the assembly job shop (as applied by Portioli-Staudacher (2000), Lu et al. (2010) and Thürer et al. (2012)) as results for this environment are indicated by Control Policy II – only controlling the level 1 job shop and releasing orders to the level two assembly shop once all of its work orders are complete – as the average lead time for all release methods can be considered equivalent at the level 2 assembly shop under immediate release.

5. Conclusion

This study has addressed one of the most important research questions raised in recent WLC literature: which release method should be applied in assembly shops (e.g. Hendry et al., 2008; Stevenson et al., 2011; Thürer et al., 2011). It focused on a two-level multi-stage job shop, as commonly encountered in small to medium sized MTOs in practice (e.g. Portioli-Staudacher, 2000; Stevenson & Silva. 2008). For the first time, the performance of a set of release methods has been compared to provide practitioners with guidance on which release method to apply in this context. The results of the experiments underline the potential of WLC and its release method to limit WIP and reduce the percentage of tardy jobs in this production environment and should give confidence for future implementations. More specifically, the SLAR and LUMS COR release methods have been identified as the best-performing solutions.

Moreover, the important questions regarding when to consider orders for release and where to control release have been addressed. In anticipation of future implementations of SLAR and LUMS COR, the following managerial implications are derived:

- **When to consider assembly orders for release?** Dividing (preceding) job shop and assembly shop does not lead to any performance improvement (i.e. Control Policy I & III compared to Control Policy II & IV). This implies that an assembly order should be released to the level 2 assembly shop as soon as the first sub-assembly in its routing is finished, rather than only when all the sub-assemblies that make up an assembly order are complete at level 1.
• Where to exercise release control?: Whether the release of order to both level 1 and level 2 should be controlled depends on whether shifting WIP from the pre-shop pool to sub-assembly inventory in front of the level 2 assembly shop is perceived as beneficial by the management of a given company. Thus, the decision depends to some degree on the space restrictions within a company. In general, controlling release to the level 2 assembly shop (i.e. Control Policy III & IV compared to Control Policy I & II) leads to a reduction in the percentage of tardy jobs.

Finally, the main limitations of this study are that: (i) it is assumed that the norm levels are equal in the level 1 job shop and level 2 assembly shop, despite the processing time differences; and, (ii) only simple backward infinite loading has been used to schedule operations. Future research should assess how the progress of sub-assemblies can be better coordinated, e.g. by considering the current load on the shop floor – thus using some kind of backward finite loading. This includes the determination of optimum workload norms or slack factors to ensure that the dispatching rule, scheduling mechanism and release method work together. Finally, this research addressed an important issue encountered in empirical research; future research should involve going back to the field to implement the findings of this study in practice.

Acknowledgments

Matthias Thürer would like to acknowledge FAPES – Fundação de Amparo a Pesquisa do Estado São Paulo – for the provided financial support (2011/10605-1).

References


