The Effects of Controlled Release and Congestion in a Flow Shop With Prohibited Early Shipments

Warren W. Fisher[1], Samir Barman[2]

[1] Stephen F. Austin State University, College of Business (wfisher@sfasu.edu)
[2] University of Oklahoma, Price College of Business (sbarman@ou.edu)

Abstract

Experiments were conducted with a simulation model of a flow shop with prohibited early shipments under a wide variety of conditions including congestion, which was modeled separately as "blocking" and as "material handling cost (MHC)." Congestion was found to significantly degrade most standard measures of shop performance, with stronger effects at high utilization and with blocking. Contrary to intuition, controlled release was found to be of little benefit under blocking. Further, controlled release under both blocking and MHC would appear to have a negative effect on flow time, percent of jobs tardy, work in process (WIP) cost, and late cost. Early cost and MHC were shown to benefit from controlled release under some circumstances, so total cost can benefit from controlled release depending on system configuration and cost parameters. With one such combination, a moderate level of release control was found to perform better on total cost than either immediate release or a stronger level of control.

1. Introduction

In job shops and flow shops, it is common practice for arriving jobs to be held in the shop's planning phase for a period of time before "releasing" them to production. Sometimes this is done as a matter of convenience, i.e., jobs arriving during the day may be reviewed at the start of business the next day. In other cases, release may be intentionally delayed in an effort to improve shop performance.

One apparent reason to delay release is to avoid the negative effects of shop congestion. If jobs are (temporarily) released faster than the shop can process them, work-in-process (WIP) inventory cost might increase. If jobs are released that are not due in the near future, they may absorb resources that are needed by jobs that are more immediately due, thereby causing those jobs to be completed late. Further, releasing a large number of jobs to production may actually reduce the system's effective production capacity in at least two ways: (1) material movement may be hampered if there is a limited capacity system for moving jobs between work centers, such as in the system reported by Sabuncuoglu and Karapinar [34, 35]; or (2) if there is limited WIP or queue space for jobs, work centers may be "blocked" from starting a new job when the downstream work center has no space for the job just finished. An extreme but common subset of situation (2) is the case where there is no intermediate buffer space at all [14]. A variation of situation (2) is where the jobs can be moved to a temporary storage location when there isn't room for them at the upcoming work center. In an automotive paint and body shop, for example, cars would normally go from body work to the paint shop. If there is insufficient room in the paint shop, they could be moved elsewhere temporarily. An extra cost for the movement would of course be incurred, but production would never be blocked. Such an extra cost is denoted "material handling cost" (MHC) for this paper.

Still another reason for controlled release is that releasing jobs earlier than necessary may cause them to be completed significantly before their due dates. If early shipments are prohibited (e.g., if the customer is using just-in-time processing and doesn't need or want the items delivered early), jobs completed early will accumulate finished goods inventory cost unnecessarily.

Unfortunately, most studies of controlled release fail to show a consistent advantage over releasing jobs immediately, even with prohibited early shipments. One likely reason is that delaying release can increase the number of tardy jobs. In terms of cost impact, total cost (when lateness cost is considered) can be worse under controlled release even though finished goods inventory cost is better. This effect is shown by Barman and Fisher [6], who found that controlled release improved total cost only at low utilization (which would lend itself to early completion under immediate release) and low late-to-early penalty cost. Conversely, when the shop is busy, the authors observed that immediate release along with an appropriate sequencing rule seems to perform well. A possible reason is that, in a busy shop, there are enough jobs in each work center queue so that an effective sequencing rule can pick out the jobs that need to be processed most immediately. One can also see how flow time could be negatively affected by controlled release. If release is controlled to the degree that work centers run out of work, average flow time will obviously be worse than if the work centers can be kept busy. However, it should be possible to mitigate this effect with the proper choice of release rule.

The failure of controlled release to consistently improve overall performance has been termed a "paradox" by Sabuncuoglu and Karapinar [35]. They go on to observe that the studies showing poor or mixed performance by controlled release fail to explicitly model the congestion that occurs in a shop where jobs are released to the floor faster than they can be processed. Typically, the models used in analyzing controlled release allow an essentially unlimited number of jobs in the queue of each machine. When the system is busy in that environment, early completion is rarely a problem, and (as noted above) having a large number of jobs in each queue permits a due date-based sequencing rule to pick out the ones that need to be processed immediately. The authors show that (at least in the system they modeled) the presence of congestion appears to make controlled release more attractive.

This research explores the effects of controlled release and congestion on the performance of a flow shop with prohibited early shipments. In part, this continues and expands on previous works [34, 35] to show the degree to which controlled release can improve system performance in the presence of congestion. The paper is organized as follows. The next section provides a review of related literature, after which a description of the computer simulation model used in this study is presented. Subsequent sections describe the experimental design, results, and conclusions.

2. Related Literature

Since the publication of the seminal works by Plossl and Wight [29], and Wight [38], numerous studies have addressed the role of the order review and release (ORR) function on the shop floor control. Wisner [39] provided an excellent up-to-date review of the studies and findings pertaining to the ORR research. Sabuncuoglu and Karapinar [34] presented a comprehensive classification of the ORR methods proposed in the literature and a summary of major findings. In addition, Land and Gaalman [22], and Van Ooijen [37] discussed in detail the principles of workload control and various load-based order release rules. An extensive overview of the components and activities of ORR, along with its role on the shop floor control, is presented in [7, 24, 31].

Generally, order release policies can be separated into two categories, infinite loading and finite loading methods, each of which could operate under a forward- or backward-looking algorithm [39]. While there exist many studies investigating the effectiveness of these policies, the results appear to be mixed [12, 18, 27]. For example, Melnyk and Ragatz [24] claimed that ORR policies are more critical than dispatching rules because they affect in-process inventory, job lead time and job priority. Irastorza and Deane [17] observed mixed results on the performances of forward finite and forward infinite release strategies. It is also reported that the use of ORR may reduce the job queue time and work-in-process inventory, but it does not necessarily improve overall lead times [25]. Therefore, an effective strategy could be releasing the jobs immediately to the floor instead of filtering them through an ORR mechanism [26]. Cigolini et al. [12] concluded that the performances of ORR strategies are highly dependent on factors such as due date tightness, shop utilization level, dispatching rule, etc. Sabuncuoglu and Karapinar [34] illustrated, through the use of their DLR (Due date and Load based Release) algorithm, the advantage of incorporating both load and due date information simultaneously to improve system performance. Also reported elsewhere [8] is that release policies do reduce variance of job lateness, but the extent of improvement depends on how the jobs are sequenced using their operation due dates.

Ragatz and Mabert [31] showed that simple release rules, such as Backward Infinite Loading (BIL) and Modified Infinite Loading (MIL), could outperform other complex rules when used with non-due date oriented dispatching rules. Ahmed and Fisher [2] reported that the performance of release policies is contingent upon the choice of shop utilization level, due date policies, and the dispatching rules used. Kim et al. [21] concluded that both output flow control and bottleneck flow control are better mechanisms than the dynamic flow control approach for releasing work to the production system.

Results on the performance of ORR policies that extend beyond simple strategies to include more complex capacity-sensitive and work-in-process (WIP) related techniques are also mixed [28, 33]. Philipoom et al. [28] introduced a capacity-sensitive release method called path based bottleneck (PBB) and showed that it performed better than simple strategies only under tight due date conditions. However, under loose due date situations, a traditional release-date based ORR policy, such as modified infinite loading (MIL) rule, might be a better choice. Roderick et al. [33] provided evidence to suggest that a WIP based release strategy performs better than a shop bottleneck based strategy under a variety of shop conditions for throughput, due date, and WIP related performance criteria. Zozom et al. [40] presented heuristic algorithms to determine the release times of new jobs to minimize the maximum job tardiness using the information on current shop floor conditions. Other studies [15, 26-27, 36] reported that the shop performance could be enhanced if ORR policies are preceded by another level of control that attempts to smooth the potential workload.

While most prior research used due date and throughput related shop measures, a relatively small subset considered cost-based performance criteria [2, 6, 28, 31] in evaluating the effectiveness of different ORR policies. With regard to the studies pertaining to job release mechanisms, Ahmed and Fisher [2], and Ragatz and Mabert [31] included average total cost per period as the primary shop performance criterion. Barman and Fisher [6] chose to compare work-in-process, earliness, lateness, and total costs while investigating the effects of combining simple priority rules and controlled release of jobs in situations where early shipments are prohibited. In addition, Philipoom et al. [28]

considered inventory holding and customer service costs, along with several traditional shop based measures, in order to compare the performance of three order review and release policies.

As noted above, prohibited early release and/or shop congestion are characteristics that would seem to favor controlled release. Many of the studies on controlled release do indeed include the assumption of prohibited early shipments, including [2, 6, 11, 19, 31]. The impact of congestion, however, has received less than sufficient research attention. As reported in [3], "blocking" occurs in flowshops with limited or no buffer capacity between processing stages, such as those in paint shops, aluminum, steel, food-processing, and petrochemicals processing industries. While there have been a number of papers devoted to the blocking case [1, 3, 14, 20, 23], those are typically devoted to very specific situations, such as cases with no buffer space at all between work stations and cases devoted to optimizing a single performance measure over a static set of jobs. For example, Bagchi et al. [3] developed heuristic algorithms to estimate the minimum makespan time in a two-stage blocking flowshop problem with material handling constraint. McCormick et al. [23] presented heuristics to minimize the cycle time in a flowshop with blocking and fixed capacity buffers between machines. The mathematical connections between a blocking flowshop and a no-wait flowshop, in which no waiting is allowed between the processing stages, were explored in [1]. Karabati and Kouvelis [20] analyzed the blocking flowshop problem with an integer programming model. A thorough treatment of algorithms, computational complexities, and results in the no-wait and no-buffer-blocking cases are provided in a survey by Hall and Sriskandarajah [14].

Sabuncuoglu and Karapinar [34, 35] address the problem of congestion with a simulation model of a job shop in which the system has a finite capacity at each work center and a transporter system to move jobs between centers. If a job arrives at a work center that is already at capacity, the transporter must move the job to an auxiliary storage area, thereby consuming limited transporter capacity. In [35], allowing an intermediate number of jobs on the shop floor was found to produce shorter flow times than allowing either a high or low number. This is consistent with intuition: if the release process is too strict, machines may run out of work, but if too many jobs are released, the resulting congestion makes the shop inefficient. The authors also report that due date-related measures appear to be improved by controlled release when congestion is explicitly modeled. In addition, "continuous" release rules (those in which checks for release are conducted continuously) outperformed "periodic" release rules (those with checks only at specific time intervals) on both mean flow time and due date-related performance measures.

In summary, evidence suggests that controlled release is more likely to result in improved system performance if early shipments are prohibited and if congestion directly affects the ability of jobs to flow through the system, both of which are obviously the case in many real-world job shops. Unfortunately, there are few studies in the literature that use both of these experimental conditions, and none of those focus on cost results. This paper expands the body of scheduling research by filling that void. Experiments are conducted separately with blocking and material handling cost.

3. Computer Model

3.1 Overall Operation

This study is based on a simulation model written in SLAM II [30] that is similar to models used in previous studies [4, 5, 6]. The model represents a flow shop with three work centers (designated WC1, WC2, and WC3), each of which has two machines. Jobs arrive by a Poisson process with a mean time between arrivals of one hour. Processing time is determined upon arrival and follows a truncated exponential distribution with parameters set to produce a specified utilization level. As done in many prior studies including [4, 5, 6, 9, 32, 34, 35], the due date of each job is determined from the job's total processing time using the total work content (TWK) method so as to provide a specified level of "tightness" (see section 5.1 below for details).

Arriving jobs are placed in a "pre-shop" queue, from which they may be released to production immediately or with a delay as described below. Jobs in the pre-shop queue are checked for release every time there is an event that could result in the need to release a job. Relevant events include every arrival, every time that a job is completed in work center one (WC1), and at certain other times depending on the rule in effect. This procedure effectively constitutes "continuous" release checking, which was found by [35] to produce better performance on mean flow time and due date-related performance measures compared to "periodic" checking. The model is capable of changing the sequence in which jobs in the pre-shop queue are checked, but all experiments reported in this paper involved checking them first-come-first-served.

All jobs entering the shop have the same routing and require three operations, one in each of WC1, WC2, and WC3 in that order. Jobs at any work center wait in a single queue if both machines are busy. As a machine becomes free, a job is selected from the queue using the priority rule in effect. Following the operation, the job is routed immediately to the next work center except when affected by congestion as described in the next section. Since early shipments are prohibited, items completed at WC3 before the due date must be held in finished goods inventory until the due date.

3.2 Congestion

The shop may have limited work in process (WIP) storage space, with the result that shop performance may be directly affected by congestion. Space available and used are measured as a number of jobs in each queue, allowing simulation of a shop where all jobs are similar in space requirements. The space used is called “jobs in queue” (JIQ). If the available space is exceeded, work must either stop (causing “blocking”) or be temporarily moved out of the way (incurring “material handling cost”) according to the simulation option in effect. Each of the three work centers has the same amount of space, denoted maximum jobs in queue (MAXJIQ).

“Blocking” reflects a shop in which production must be stopped when available WIP storage space is exceeded. At WC1, the possibility of blocking affects job release. That is, if releasing a particular job would put WC1 above MAXJIQ, the job is placed in a special holding area rather than being released. Those jobs have priority over jobs in the pre-shop queue when WC1 conditions eventually allow jobs to enter. At WC2 and WC3, blocking may affect production at the previous WC. If WCi (i=2 or 3) is currently at MAXJIQ with both servers busy, a job completed at WC(i-1) is not permitted to move to WCi, thereby preventing the WC(i-1) machine from starting another job. Blocking at WC3 may affect the entire shop: if WC3 is at capacity, jobs may stack up at WC2, eventually moving it to capacity as well; WC2 can then have the same effect on WC1, eventually preventing any jobs from being released. There is no blocking in the WC3 completion area. Note that WIP inventory is increased upon completion of a production step whether or not blocking occurs, since completion of the step implies an expenditure of time and money associated with processing the job.

As an alternative to blocking, the shop may have the ability to move excess WIP from the shop floor to a temporary storage location. Obviously, doing so would incur an additional cost, which is denoted “material handling cost” (MHC). MHC can occur in any work center but is easily prevented in WC1 with a release rule based on jobs in WC1. There are two parameters for MHC: a threshold (MAXJIQ) and a cost per job moved (MHC). If the arrival of a job at a WC causes the number of jobs to exceed the allowed maximum, an MHC event occurs and the shop is charged for the cost of moving the material. The penalty is assessed no more than once per job per work center, thereby reflecting the total cost of moving the job to the temporary storage location and later back to the shop floor. A job that invokes MHC continues to be counted in the center’s queue until it moves into processing. Material handling other than that associated with MHC is not modeled.

3.3 Cost Computations

Shop costs other than MHC are similar in structure to those used in [2, 6, 28, 31] and consist of work in process (WIP) cost, finished goods inventory cost (early penalty), and late delivery penalty. All costs are computed on a per week (40 hours) basis. Work in process (WIP) cost is accumulated with the completion of each production step and is proportional to processing time. Specifically, there is no WIP cost for a job until it has completed processing on the first work center. It then begins accumulating WIP cost at the rate of $.10 per hour of completed processing time (on this job) per week. The rate for a job rises to reflect the new amount of completed processing time at the end of the second work center. At the end of the third work center, a determination is made of the number of weeks that the job is earlier or later than its due date. Early jobs accumulate finished goods inventory cost (early penalty cost) from completion to due date at the rate of $.10 per hour of total processing time on this job per week. Jobs completed after the due date are assessed a late penalty at a rate proportional to the early penalty rate. The model reports cost results for late-to-early penalty ratios of 1:1 ("low"), 20:1 ("moderate"), and 50:1 ("high"). These result in late penalty costs of $.10, $2.00, and $5.00 per hour of completed processing time per week late, respectively. These ratios are similar to those used in prior research [2, 6, 28, 31].

3.4 Performance Measures

Consistent with [2, 6, 28, 31, 34, 35], the performance measures for this research include:

- --average flow time (job arrival to delivery).
- --percentage of jobs tardy.
- --average WIP cost per week
- --average earliness penalty per week
- --average late penalty per week at all three ratios
- --average MHC per week
- --average total cost per week at all three ratios, i.e., the sum of average WIP cost per week, average early penalty per week, average late penalty per week, and average MHC per week.

Both cost measures and flow time are considered to be especially important, reflecting the situation where the shop must be able to complete work quickly and at low cost in order to be competitive.

3.5 Procedures Implemented

As noted above, the model incorporates either blocking or MHC, with the degree of restriction of each controlled by MAXJIQ. A variety of release and sequencing rules are implemented. Let:

A_i = arrival time (hour number), job i
R_i = release time (hour number), job i
D_i = due date (hour number), job i
P_i = total processing time in hours, job i
Q_i = number of other jobs in queue (waiting) at the work centers job i has left to visit, including any jobs that have been blocked from entry into WC1.
$k_1, k_2, \ldots k_n$ = planning factors (determined experimentally)

Release rules implemented include the following. Note that no job can ever be released if WC1 is blocked as described above:

Immediate Release (IMM): $R_i = A_i$. That is, arriving jobs are released immediately to WC1 as long as WC1 isn't blocked. Although this is the simplest rule, past research has often found it to work well under certain experimental conditions. For this and the other rules below, if WC1 is blocked at time R_i, the job is released at the soonest time after R_i when WC1 becomes unblocked. If more than one job is waiting for release when WC1 becomes unblocked, the jobs are released first-come-first served.

Modified Infinite Loading (MIL): $R_i = D_i - k_1 * P_i - k_2 * Q_i$. If $R_i \leq$ "now" at the time of examination, the job is released; otherwise the job is not released and is checked again no later than at time R_i (at which time it will be released unless WC1 is blocked). Some implementations of this rule [2, 31] multiply k_1 by the number of operations in the job, thereby giving earlier release dates (all other things equal) to jobs with more operations. Since all jobs have three operations in this shop, it is more logical to multiply the parameter by processing time so that longer jobs have earlier release dates.

Maximum Work in Shop (MXWS): a job is released only if the job's total work time plus the total remaining work time of jobs currently in the shop is less than or equal to a set level (k_3). Work remaining in shop is defined to be the sum of the uncompleted work of all jobs in the shop EXCEPT those in the pre-shop file and those that have completed processing in WC3 (jobs "released" but blocked at WC1 are counted as being in the shop). The amount is increased by the full processing time at all three work centers upon job release, and it is decreased by the processing time of the step completed when a job is completed at a WC. Jobs are checked in the prescribed sequence and are released so long as maximum work is not violated. Obviously, this rule provides an extremely simple approach to controlling congestion in the shop by limiting the amount of pending work on the shop floor.

MXWS1: like MXWS but, if a particular job would not be released because of work in the shop, the job may be released anyway if an operator would otherwise be idle at WC1. This rule attempts to improve on MXWS by preventing WC1 from running out of work. As shown by Van Ooijen [37], use of release rules which give priority to orders that require resources which currently are underloaded (and delay orders requiring resources that are overloaded) can improve delivery performance.

Maximum Jobs in Shop (MXJS): a job is released only if the total number of jobs in the shop (including this job) will be less than or equal to a set number (k_4). Total jobs in shop is defined to be the number of jobs released to the shop (including those blocked at WC1) that have not yet completed processing in WC3. The number is increased immediately upon job release, and it is decreased when a job is completed at WC3 whether or not the due date has been reached. Jobs are checked in the prescribed sequence and are released so long as the maximum number is not violated. This rule is analogous to MXWS but simpler, since processing time is not used. It is included for testing because both blocking and MHC are based on a number of jobs rather than processing times.

MXJS1: like MXJS but, if a particular job would not be released because of the number of jobs in the shop, the job may be released anyway if an operator would otherwise be idle at WC1. This rule attempts to improve on MXJS by preventing WC1 from running out of work.

Maximum Jobs in WC1 (MXJC1): similar to MXJS but based only on WC1. A job is released only if the total number of jobs in WC1 (including this job) will be less than or equal to a set number (k_5). Total jobs in WC1 is defined to be the number of jobs released to the shop that have not yet moved to WC2 or further. The number is increased immediately upon job release, and it is decreased when a job moves from WC1 to WC2 (which might involve delay after processing if WC1 is blocked due to the status of WC2). Jobs are checked in the prescribed sequence and are released so long as the maximum number is not violated.

Simplified Due Date and Load Based Release (SDLR): this rule uses information from both the job's due date and shop status. It is based on the DLR rule, which was reported to perform well by [34]. Release is determined as follows:

1. R_i is calculated as for the MIL rule for each job in the pre-shop queue.
2. Jobs for which $R_i \leq (\text{now} + k_6)$ are designated "priority" jobs and are considered further for possible release. That is, jobs are considered only if they need to be released soon in order to meet the due date. The parameter k_6 is sometimes called a time "fence."

3. The priority jobs are checked in ascending order of R_i and are released if the release action does not exceed maximum allowable jobs or work in the shop as limited by one of rules MXWS, MXWS1, MXJS, MXJS1, or MXJC1.

The following priority rules are available for sequencing decisions at the three work centers. Although the model is capable of using different rules in different centers, all centers used the same rule for this research. These rules are also available for use in the release process in that each can be selected for use as the sequence in which jobs in the pre-shop queue are checked for release:

First-come-first-served (FCFS): select jobs for processing (or release) in their order of arrival to the work station (or shop).

Shortest Processing Time (SPT): from all jobs waiting to be processed at any given work center, the one with the shortest processing time at that work center is selected next. When used for release, the rule selects first the job with the shortest total processing time.

Earliest Due Date (EDD): the job that has the most imminent due date is selected next for processing (or release).

Modified Shortest Processing Time (SI^X): a truncated SPT rule, it classifies the jobs waiting to be processed according to the slack (time till due minus processing time remaining) of each job. Jobs with a zero or negative slack are placed into a priority queue, while the rest are placed into a normal queue. The priority queue is processed first followed by the normal queue; however, the SPT rule is used in both queues to determine the job processing (or release) order.

Of the above rules, SPT, EDD, and SI^X represent rules found to work well in previous research [4, 5, 6]. The FCFS rule is included as a classical benchmark.

4. Experimental Design

Experiments were conducted in four phases. In phase one, the model was subjected to thorough testing to ensure that it operates properly. Tests included:

--verifying that the model produces queuing statistics consistent with classical M/M/2 results when its parameters are set appropriately (IMM release, no blocking, Poisson arrivals, exponential service, and FCFS sequencing).

--tracing entity flow to ensure correct release, sequencing, blocking, and processing behavior.

--verifying accuracy of cost calculations and other performance measures.

--verifying that the simulation run block size was large enough to produce approximately independent blocks. A block size of 20,000 simulated hours (with the first block of each set removed to eliminate any possible initialization bias) passed the test for independence by Fishman [13].

--verifying the absence of initialization bias in the remaining batches after discarding the first batch of each set.

Phase two consisted of experiments identifying system behavior under various experimental conditions. The results are presented below and were used to set values of the parameters and experimental conditions for the main experiments, which comprised phase three. That phase consisted of a full factorial analysis of the effects of six factors: release rule, sequencing rule, due date tightness, utilization, congestion level, and congestion effect (MHC or blocking). Phase four consisted of sensitivity experiments with selected model parameters and experimental conditions.

5. Results

5.1 System Behavior and Parameter Selection

The following experiments were conducted to identify how the simulated system behaved under various experimental conditions. The results were used to establish values of the parameters and experimental conditions for the main experiments as indicated:

a. a "base case" configuration was established as: IMM release (except for experiments setting release parameters), SI^X sequencing, loose due dates, low utilization, and no congestion. IMM was chosen because its simplicity gives broad appeal, and it has been shown to perform as well or better than more complex rules in many studies including [6]. SI^X was chosen because it performed well in previous research with a similar model [6] as well as in preliminary experiments with the current model. In particular, SI^X produces flow time nearly as good as SPT and is better than the other sequencing rules on both lateness and medium-ratio total cost.
b. TWK parameter values were set under the base case for 10% tardy as "loose" due dates and 30% tardy for "tight", separately for low and high utilization. The percentages were selected to be consistent with [34, 35]. TWK=3.8 for loose, low utilization; 7.2 for loose, high utilization; 2.7 for tight, low utilization; and 5 for tight, high utilization.

c. low (82%) and high (92%) utilization were set up using the base case via parameters for a truncated exponential service distribution. These levels were selected as consistent with most previous related research, including [4, 5, 6, 28, 31, 34, 35]. The minimums and maximums serve to prevent the unrealistically small and large times that would occur with pure use of the exponential distribution and are consistent with the other research cited above. In addition, truncating job length (especially the maximum length) serves to control output variance, which otherwise would be quite high for certain rule combinations. For low utilization, mean processing time=1.64 hours per work center with a minimum of .25 and maximum of 7. For high utilization, mean processing time=1.84 hours per work center with a minimum of .3 and maximum of 8.

Experiments demonstrating the effects of congestion were also conducted and are presented in Table 1. Various levels of congestion were created by varying MAXJIQ, separately under MHC and blocking, each of which were tested under low and high utilization. Each row represents the mean from three batches of 20,000 simulated hours each except that the percent of jobs affected came from just the first batch. Various values of MHC were tested in an effort to find a number that would yield similar effects on total cost, medium ratio, with both MHC and blocking under MAXJIQ values suitable for "loose" and "tight" congestion. As a result, MAXJIQ values were selected as indicated in Table 2. When combined with MHC cost per event of $.20, the indicated MAXJIQ values caused total cost at the medium ratio to be roughly 5.5% higher under loose congestion and MHC than with no congestion (MAXJIQ unlimited). That percentage increase holds for both low and high utilization. Under tight congestion and MHC, total medium-ratio cost is roughly 12% higher than with no congestion, again for both high and low utilization. Under low utilization and blocking, the total medium-ratio cost increases are roughly 2% and 13% for loose and tight congestion respectively, or about the same order of magnitude as with MHC. Under high utilization, costs increase by considerably larger percentages under blocking than they do under MHC. The MHC value of $.20 per occurrence reflects a few minutes for a minimum wage worker to move a small job a short distance. Sensitivity experiments were conducted with a larger MHC value.

Several other results from Table 1 are noteworthy:

a. At higher levels of congestion (lower MAXJIQ values) under blocking, actual work center utilization increases well above the target values. This is because a worker cannot pass on a completed job to the next work center when that center is at space capacity. Note that the effects are extremely strong in WC1 at high utilization.
b. Under MHC, congestion has no effect on flow time, percent of jobs tardy, or cost measures other than MHC. MHC per week rises, of course, with more congestion. Under blocking, however, flow time, percent of jobs tardy, late cost, and total cost-medium ratio all rise with more congestion, and the rise is quite dramatic under high utilization.
c. Early cost, as might be expected, moves inversely with late cost under blocking.
d. Surprisingly, WIP cost tends to decrease with higher congestion levels under blocking. It would appear that the effect of blocking to prevent entry into WC1 more than offsets the increased WIP resulting from jobs that are blocked between work centers.

Once the MAXJIQ values for levels of congestion were set, experiments were conducted under the base case with loose congestion to find the parameters for each release rule that resulted in the best performance on the total cost-medium ratio performance measure. Results across release rules were compared, and the following rules were eliminated from further consideration: MAXJS, MAXJS1, MAXWS, MAXWS1, and SDLR. That left the IMM, MIL, and MXJC1 rules in consideration. MXJC1 actually performed better than IMM or MIL, but the latter two were retained because they are classical benchmarks. Additional experiments were then conducted to identify separate parameters for high and low utilization for each of tight, loose, and unlimited (i.e., no congestion effect) congestion. Selected parameters for the MIL and MXJC1 rules are presented in Table 2. While the indicated parameter values are certainly not claimed to be optimal, it is believed that they represent good performance.

5.2 Main Experiments

The main experiments were designed to obtain a more complete picture of the degree to which shop performance with and without congestion is affected by controlled release under various experimental conditions. To do so, a full factorial analysis of six factors was conducted as follows (parameter settings for the various options are provided above):

--three release rules: IMM, MIL, and MXJC1.
--two sequencing rules: SPT, and SI^X. Plans to include the FCFS and EDD sequencing rules were dropped when it was discovered that the model became unstable under some experimental conditions with those rules.
--two levels of due date tightness: loose and tight.
--two levels of average machine utilization: low and high.
--three levels of congestion: tight, loose, and unlimited, representing small, medium, and unlimited work center queue space (MAXJIQ) respectively.

--two congestion effects: blocking and MHC.

Table 1 Congestion Effects on Shop Performance

Max JIQ	M/B*	Util. Level	Actual Utilization WC1	WC2	WC3	Flow	Pct. Tardy	MHC*	WIP*	Early Cost*	Late Cost-M*	Total Cost-M*	Percent Affected+
50	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.00%
25	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.00%
20	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.02%
15	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.06	1.49	5.86	1.78	9.19	0.22%
10	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.27	1.49	5.86	1.78	9.40	1.16%
8	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.48	1.49	5.86	1.78	9.61	2.20%
7	M	L	81.33%	81.50%	81.23%	9.53	8.93%	0.72	1.49	5.86	1.78	9.85	3.41%
6	M	L	81.33%	81.50%	81.23%	9.53	8.93%	1.14	1.49	5.86	1.78	10.27	5.58%
5	M	L	81.33%	81.50%	81.23%	9.53	8.93%	1.90	1.49	5.86	1.78	11.03	9.26%
4	M	L	81.33%	81.50%	81.23%	9.53	8.93%	3.27	1.49	5.86	1.78	12.40	15.56%
3	M	L	81.33%	81.50%	81.23%	9.53	8.93%	5.47	1.49	5.86	1.78	14.60	25.21%
50	M	H	91.50%	91.67%	91.37%	17.84	12.50%	0.05	3.19	15.40	14.58	33.22	0.00%
25	M	H	91.50%	91.67%	91.37%	17.84	12.50%	0.50	3.19	15.40	14.58	33.67	0.11%
20	M	H	91.50%	91.67%	91.37%	17.84	12.50%	0.68	3.19	15.40	14.58	33.86	0.64%
15	M	H	91.50%	91.67%	91.37%	17.84	12.50%	0.99	3.19	15.40	14.58	34.16	2.17%
10	M	H	91.50%	91.67%	91.37%	17.84	12.50%	1.94	3.19	15.40	14.58	35.11	6.55%
8	M	H	91.50%	91.67%	91.37%	17.84	12.50%	3.04	3.19	15.40	14.58	36.22	11.35%
7	M	H	91.50%	91.67%	91.37%	17.84	12.50%	3.94	3.19	15.40	14.58	37.11	15.02%
6	M	H	91.50%	91.67%	91.37%	17.84	12.50%	5.18	3.19	15.40	14.58	38.35	20.37%
5	M	H	91.50%	91.67%	91.37%	17.84	12.50%	6.91	3.19	15.40	14.58	40.08	28.05%
4	M	H	91.50%	91.67%	91.37%	17.84	12.50%	9.17	3.19	15.40	14.58	42.35	38.24%
3	M	H	91.50%	91.67%	91.37%	17.84	12.50%	11.92	3.19	15.40	14.58	45.09	50.24%
50	B	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.00%
25	B	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.00%
20	B	L	81.33%	81.50%	81.23%	9.53	8.93%	0.00	1.49	5.86	1.78	9.13	0.02%
15	B	L	81.47%	81.50%	81.23%	9.54	8.97%	0.00	1.49	5.86	1.78	9.13	0.24%
10	B	L	81.70%	81.67%	81.23%	9.55	9.07%	0.00	1.47	5.85	1.83	9.15	1.74%
8	B	L	81.93%	81.83%	81.23%	9.62	9.43%	0.00	1.46	5.83	2.06	9.35	4.28%
7	B	L	82.20%	82.03%	81.23%	9.73	9.93%	0.00	1.44	5.79	2.44	9.67	7.76%
6	B	L	82.70%	82.30%	81.23%	9.91	10.90%	0.00	1.42	5.73	3.13	10.28	11.57%
5	B	L	83.83%	82.93%	81.17%	10.38	13.00%	0.00	1.41	5.57	4.79	11.77	19.14%
4	B	L	86.07%	84.20%	81.20%	11.64	17.93%	0.00	1.39	5.21	10.37	16.97	32.02%
3	B	L	90.60%	86.53%	81.20%	16.47	34.30%	0.00	1.36	4.13	36.62	42.12	58.91%
50	B	H	91.53%	91.67%	91.37%	17.84	12.50%	0.00	3.17	15.40	14.62	33.19	0.00%
25	B	H	92.10%	91.73%	91.33%	18.43	14.03%	0.00	2.87	15.12	17.09	35.09	8.26%
20	B	H	92.20%	91.80%	91.37%	18.59	14.20%	0.00	2.79	15.07	18.08	35.94	12.61%
15	B	H	92.40%	91.93%	91.37%	19.10	14.57%	0.00	2.70	14.96	21.02	38.68	17.20%
10	B	H	93.63%	92.50%	91.33%	22.67	20.83%	0.00	2.61	13.60	35.38	51.58	24.47%
8	B	H	94.90%	93.03%	91.33%	28.53	28.93%	0.00	2.52	12.02	69.82	84.36	40.17%
7	B	H	95.93%	93.53%	91.40%	34.14	36.23%	0.00	2.45	10.55	102.50	115.50	52.29%
6	B	H	97.33%	94.13%	91.43%	49.78	54.40%	0.00	2.37	7.54	215.34	225.25	67.11%
5	B	H	100.00%	95.57%	91.43%	276.29	100.00%	0.00	2.29	0.00	2561.07	2563.36	100.00%

* M = with MHC congestion at $.20; B = with blocking. Costs are per week.

\+ percent of jobs experiencing MHC or blocking as applicable at WC1.

The design produced a total of (3 x 2 x 2 x 2 x 3 x 2) = 144 cells in the experiment. Multiple observations of each cell were collected using the batch means method in which one simulation run is divided into equal subgroups, treating each subgroup as a single independent observation (see [13] for a thorough discussion of the methodology). A total of 20 observations were collected for each cell, consisting of 10 batches from each of two independent runs (not counting the first batch of each run, which was discarded). The batch means procedure is common in job shop scheduling research and has been used in several papers related to the research presented here, including [4, 5, 6, 25, 31]. The results were analyzed with ANOVA.

Table 2 Congestion and Release Rule Parameters

Utilization*	Congestion Tightness*	MAXJIQ*	MXJC1+	MIL K1/K2+
L	U	999	13	3.4, 3.5
L	L	8	11	3.5, 3.5
L	T	6	9	3.4, 3.5
H	U	999	12	6.5, 3.0
H	L	10	13	6.5, 3.0
H	T	7	10	6.3, 3.0

*Utilization is L=low or H=high; congestion tightness is U=unlimited or no congestion, L=loose, or T=tight. MAXJIQ is the maximum jobs allowed in any queue before congestion has an effect and is set for the indicated combination of utilization and congestion tightness.
+MXJC1 is the parameter for the Maximum Jobs in WC1 rule, and MIL K1/K2 are the parameters for the Modified Infinite Loading rule. See text.

Table 3 presents the ANOVA results for all performance indicators except late and total cost at the low and high ratios, which had significance values similar to the respective medium ratio measures. Somewhat surprisingly, the release rule was not found to significantly affect performance on percent of jobs tardy, MHC, late cost, or total cost. The release rule is marginally significant for flow time, with IMM significantly better than MIL and slightly (not significantly) better than MXJC1. The release rule does significantly affect WIP and early cost, with IMM best on WIP and MIL best on early cost. Since the interactions between release rule and the other factors including congestion and congestion effect were mostly not significant, the experiments fail to show that controlled release is advantageous, even in the presence of congestion.

Table 3 ANOVA (Probability Results are Random)

Factor(s)	Flow	Percent Tardy	MHC	WIP	Early	Late-M	Tot-M
Rel	0.0269	0.2281	0.0761	0.0003	<.0001	0.1408	0.1496
Seq	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DDtight	0.487	<.0001	0.0002	<.0001	<.0001	<.0001	<.0001
Util	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Congest	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
CEffect	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Rel*Seq	0.3822	0.8014	0.6242	0.4406	0.9399	0.3321	0.3266
Rel*DDtight	0.0237	0.0859	0.039	0.0018	<.0001	0.2602	0.2916
Rel*Util	0.1381	0.3362	0.2798	0.0157	0.0736	0.142	0.1374
Rel*Congest	0.5469	0.947	0.3136	0.7551	0.7674	0.5479	0.5357
Rel*CEffect	0.6054	0.9166	0.0761	0.4768	0.9728	0.4289	0.4234
Seq*DDtight	0.2858	<.0001	<.0001	<.0001	0.926	0.2905	0.2512
Seq*Util	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Seq*Congest	0.0017	0.001	<.0001	<.0001	0.0189	0.0074	0.0063
Seq*CEffect	<.0001	<.0001	<.0001	<.0001	<.0001	0.0008	0.0023
DDtight*Util	0.5365	<.0001	0.612	0.2055	<.0001	<.0001	<.0001
DDtight*Congest	0.872	0.2039	0.0025	0.2493	<.0001	0.0379	0.023
DDtight*CEffect	0.6954	0.1043	0.0002	0.0616	<.0001	0.0275	0.0204
Util*Congest	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Util*CEffect	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Congest*CEffect	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Combining ANOVA significance and Duncan's Multiple Range Tests produced the following additional observations:

--flow time is lowest under SPT, low utilization, looser congestion levels, and MHC. Due date tightness had no effect.

--percent tardy is lowest under SPT, loose due dates, low utilization, looser congestion levels, and MHC.

--MHC is lowest under SPT, loose due dates, low utilization, and looser congestion (there is no MHC under blocking).

--WIP is lowest under SPT, loose due dates, low utilization, tighter congestion, and blocking.

--early cost is lowest under SI^X, tight due dates, low utilization, tighter congestion, and blocking.

--late cost is lowest under SI^X, loose due dates, low utilization, looser congestion, and MIL.

--total cost is lowest under SI^X, loose due dates, low utilization, looser congestion, and MIL under all three lateness cost ratios <u>except</u> that total cost-low is better under <u>tight</u> due dates and <u>blocking</u>.

--the degree of the above effects varied with the particular factor combination as indicated by the significant interactions.

--some three-way interactions that included the release rule and congestion factors were tested but not found to be significant.

5.3 Sensitivity Experiments

To better understand the system's behavior, experiments were run comparing IMM and MXJC1 under FCFS, loose due dates, low utilization, tight congestion, and blocking (parameters all per the factor definitions in the main experiments), which would seem at first glance to be amenable to controlled release. The MXJC1 parameter was set to 7 (lower than the 9 used in the main experiments) to force even more control of job release. The comparisons were run with common random numbers, i.e., each experiment was started from the same seeds. Surprisingly, performance on all measures was <u>identical</u>, with the only difference being that MXJC1 caused jobs to wait in the pre-shop queue instead of in the WC1 "release blocked" holding area. The comparison was repeated with SI^X sequencing. In that case, there was <u>little</u> difference in release rule performance, except that MXJC1 produced more lateness and consequently higher total cost (all ratios) than IMM. MXJC1 would likely perform closer to IMM with a larger parameter. Interestingly, the "wait" (i.e., time blocked) at WC1 was much less with MXJC1 than IMM, but the wait at WC2 and WC3 was almost identical for the two different release rules.

The above test was repeated with MHC instead of blocking and with MHC changed from $.20 to $1.00 per event. As before, there was little difference in performance between IMM and MXJC1 <u>except</u> that MHC per week was significantly less under MXJC1. Closer inspection revealed that almost all of the savings came from WC1. That is, MXJC1 significantly reduced MHC at WC1, but had almost no effect on WC2 and WC3. Also, <u>total</u> cost was <u>lower</u> under MXJC1 at all ratios.

Additional experiments were then conducted to display the effects of degree of release control on system performance; results are in Table 4. Using SI^X sequencing, loose due dates, low utilization, tight congestion, and MHC at $1.00 per event, 50 observations were taken with each of IMM and MXJC1 at the indicated parameter values. Experiments were conducted as five independent runs of 10 batches each (11 with the first discarded), and common random numbers were used in that each five-run set was started with the same seeds. Table 4 shows that, as release goes from strong control to no control, flow time, percent of jobs tardy, WIP, and late cost all steadily <u>decline</u>. Over the same range, early cost steadily <u>increases</u>. MHC, however, is lower with a <u>moderate</u> level of control. The result is that <u>total</u> cost for the moderate lateness ratio is also lower with a <u>moderate</u> level of control. Total cost for low and high lateness ratios behave similarly. Figure 1 illustrates 95% confidence intervals on total medium cost. Note that the confidence intervals for MXJC1 at 6 and 8 do <u>not</u> overlap those from the other JC1 values or IMM, thereby providing strong support for the benefit of moderate release control, at least on total medium cost.

Table 4 MXJC1 vs. IMM Release Rule Performance at MHC=$1.00

Release	Flow	% Tardy	MHC	WIP	Early	Late-Med	Tot-Med
JC1/4	10.369	14.232%	4.099	1.546	5.649	3.309	14.602
JC1/6	9.828	10.512%	3.771	1.522	5.806	2.459	13.557
JC1/8	9.700	9.712%	3.718	1.518	5.838	2.259	13.332
JC1/10	9.662	9.432%	6.149	1.517	5.848	2.200	15.712
JC1/12	9.655	9.426%	6.177	1.516	5.848	2.190	15.733
IMM	9.644	9.384%	6.158	1.515	5.851	2.157	15.681

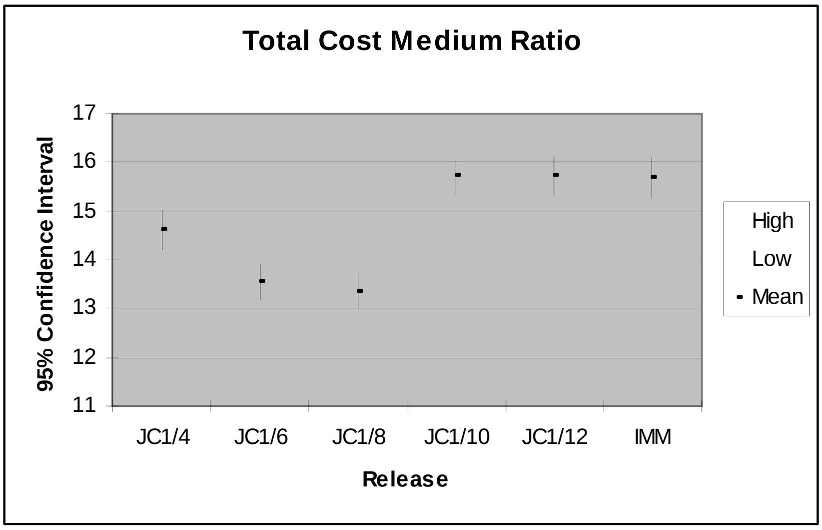

Fig. 1 Medium Total Cost Confidence Interval by Release Rule (see text)

Table 5 reports the results of experiments similar to those in Table 4 but with MIL instead of MXJC1. Similarly to the MXJC1 results, as MIL release goes from strong control to no control, flow time, percent of jobs tardy, WIP, and late cost all steadily decline. Surprisingly, MHC also declines steadily! That is, strong release control produces the highest MHC and vice versa! As before, early cost steadily increases with less control. The result is that total cost for the moderate lateness ratio is highest under the strongest release control and vice versa. Total cost for low and high lateness ratios behave similarly.

Table 5 MIL vs. IMM Release Rule Performance at MHC=$1.00

Release*	Flow	% Tardy	MHC	WIP	Early	Late-Med	Tot-Med
MIL/2.5	12.681	21.188%	14.353	1.757	3.969	5.873	25.954
MIL/2.8	11.742	16.394%	11.301	1.686	4.516	4.281	21.784
MIL/3.1	10.931	13.120%	8.962	1.618	5.018	3.240	18.837
MIL/3.4	10.268	10.912%	7.305	1.564	5.447	2.596	16.914
MIL/3.7	9.797	9.744%	6.409	1.529	5.752	2.257	15.947
IMM	9.644	9.384%	6.158	1.515	5.851	2.157	15.681

*MIL parameter is K1; all K2=3.5

Note: lower values of K1 and K2 cause longer release delays, all other things equal.

6. Conclusions

Experiments were conducted with a simulation model of a flow shop with prohibited early shipments under a wide variety of conditions including various congestion levels and two congestion effects (MHC and blocking). Results indicate that increased congestion degrades most measures of system performance, with stronger effects at high utilization and with blocking.

Contrary to intuition, controlled release is found to be of little if any benefit under blocking. This appears to be because the release rule mainly affects entry to WC1, but there is no cost difference between waiting in the pre-shop file and waiting while blocked from entry into WC1. After the jobs get to WC1, the release rule has no effect on what they do in WC2 and WC3.

Under MHC, a moderate level of release control with the MXJC1 rule produces less MHC per week than strong control or no control. If MHC per event is large enough (such as the $1.00 level used in the sensitivity experiments), total cost appears to also benefit from a moderate level of release control. Under the MIL release rule, however, only the early cost measure benefits from increased control. Unless early cost were to be much higher relative to other costs than it was in these experiments, MIL would appear to offer little net benefit for any level of release control. Further, controlled release with any rule would appear to have a negative effect on flow time, percent of jobs tardy, WIP, and late cost.

Future research needed includes work similar to the main experiments but with MHC per event as a factor. The experimental parameters also require adjustment so that a broader array of sequencing rules can be tested without the model becoming unstable. With a higher value of MHC per event, other release rules besides MXJC1 might perform better. Another experimental variable worth considering is the pre-shop check sequence, i.e., the order in which jobs are checked for release. It would also be interesting to see whether controlled release were more attractive if due dates had a random component rather than being based completely on processing time.

Finally, it seems likely that controlled release will never have the impact in a flow shop that it would in a job shop. In a job shop, jobs would go from release to all work centers, so releasing too many would affect them all. In the flow shop simulated herein, release primarily impacts WC1. Therefore, there is a need for research similar to that reported here but in a job shop environment.

References

[1] Abadi, I. N. K., N. G. Hall and C. Siskandarajah, "Minimizing Cycle Time in a Blocking Flowshop," *Operations Research*, Vol. 48, pp 177-180, 2000.

[2] Ahmed, I. and W. W. Fisher, "Due Date Assignment, Job Order Release, and Sequencing Interaction in Job Shop Scheduling," *Decision Sciences*, Vol. 23, No. 3, pp 633-647, 1992.

[3] Bagchi, T. P., S. Naranpanawe and C. Siskandarajah, "A Blocking Flowshop Scheduling Problem with Material Handling Constraint," *International Journal of Operations and Quantitative Management*, Vol. 9, No. 1, pp 1-20, 2003.

[4] Barman, S., "Simple Priority Rule Combinations: An Approach to Improve Both Flow Time and Tardiness," *International Journal of Production Research*, Vol. 35, No. 10, pp 2857-2870, 1997.

[5] Barman, S., "The Impact of Priority Rule Combinations on Lateness and Tardiness," *IIE Transactions on Scheduling and Logistics*, Vol. 30, No. 5, pp 495-504, 1998.

[6] Barman, S. and W.W. Fisher, "Combined Simple Priority Rules and Controlled Release with Prohibited Early Shipments," *International Journal of Operations and Quantitative Management,* Vol. 10, No. 3, pp 215-234, 2004.

[7] Bergamaschi, D., R. Cigolini, M. Perona and A Portioli, "Order Review and Release Strategies in a Job Shop Environment: A Review and a Classification," *International Journal of Production Research*, Vol. 35, No. 2, pp 399-420, 1997.

[8] Bertrand, J. W. M., "The Use of Workload Information to Control Job Lateness in Controlled and Uncontrolled Release Production Systems," *Journal of Operations Management*, Vol. 3, No. 2, pp 67-78, 1983.

[9] Blackstone Jr., J. H., D. T. Phillips and G. L. Hogg, "A State of the Art Survey of Dispatching Rules for Manufacturing Job Shop Operations," *International Journal of Production Research,* Vol. 20, No. 1, pp 27-45, 1982.

[10] Breithaupt, J. W., M. J. Land and P. Nyhuis, "The workload Control Concept: Theory and Practical Extensions of Load Oriented Order Release, *"Production Planning and Control,"* Vol. 13, pp 625-638, 2002.

[11] Christy, D. P. and J. J. Kanet, "Manufacturing Systems with Forbidden Early Shipment: Implications for Choice of Scheduling Rules," *International Journal of Production Research*, Vol. 28, No. 1, pp 91-100, 1990.

[12] Cigolini, R., M. Perona and A. Portioli, "Comparison of Order Review and Release Techniques in a Dynamic and Uncertain Job Shop Environment," *International Journal of Production Research*, Vol. 36, No. 1, pp 2931-2951, 1998.

[13] Fishman, G. S., Principles of Discrete Event Simulation, John Wiley, New York, 1978.

[14] Hall, N. G. and C. Siskandarajah, "A Survey of Machine Scheduling Problems with Blocking and No-wait in Process," *Operations Research*, Vol. 44, pp 510-525, 1996.

[15] Hendry, L. C. and S. K. Wong, "Alternative Order Release Mechanisms: A Comparison by Simulation," *International Journal of Production Research*, Vol. 32, No. 12, pp 2827-2842, 1994.

[16] Henrich, P., M. Land, G. Gaalman and D-J. van Der Zee, "Reducing Feedback Requirements of Workload Control," *International Journal of Production Research*, Vol. 42, No. 24, pp 5235-5252, 2004.

[17] Irastorza, J. C. and R. H. Deanne, "A Loading and Balancing Methodology for Job Shop Control," *AIIE Transactions*, Vol. 6, pp 302-307, 1974.

[18] Kanet, J. J., "Load-limited Order Release in Job Shop Scheduling Systems," *Journal of Operations Management*, Vol. 7, No. 3, pp 44-58, 1988.

[19] Kanet, J. J. and D. P. Christy, "Manufacturing Systems with Forbidden Early Shipment: Implications for Setting Manufacturing Lead Times," *International Journal of Production Research*, Vol. 27, No. 5, pp 783-792, 1989.

[20] Karabati, S. and P. Kouvelis, “The Interface of Buffer Design and Cycle Scheduling Decisions in Deterministic Flow Lines,” *Annals of Operations Research*, Vol. 50, pp 295-317, 1994.

[21] Kim, S., K. R. Davis and J. F. Cox, “An Investigation of Output Flow Control, Bottleneck Flow Control and Dynamic Flow Control Mechanisms in Various Simple Line Scenarios,” *Production Planning and Control*, Vol. 14, No. 1, pp 15-32, 2003.

[22] Land, M. J. and J. C. Gaalman, “The Performance of Workload Control Concepts in Job Shops: Improving the Release Method,” *International Journal of Production Economics*, Vol. 56-57, pp 347-364, 1998.

[23] McCormick, S. T., M. L. Pinedo, S. Shenker and B. Wolf, “Sequencing in an Assembly Line with Blocking to Minimize Cycle Time,” *Operations Research*, Vol. 37, pp 925-935, 1989.

[24] Melnyk, S. A. and G. L. Ragatz, “Order Review/Release and Its Impact on the Shop Floor,” *Production and Inventory Management Journal*, Vol. 29, No. 2, pp 13-17, 1988.

[25] Melnyk, S. A. and G. L. Ragatz, “Order Review/Release: Research Issues and Perspectives,” *International Journal of Production Research*, Vol. 27, No. 7, pp 1081-1096, 1989.

[26] Melnyk, S. A., G. L. Ragatz and L. Fredendal, “Load Smoothing by the Planning and Order Review/Release Systems: a Simulation Experiment,” *Journal of Operations Management*, Vol. 10, No. 4, pp 512-523, 1991.

[27] Melnyk, S. A., K. C. Tan, D. R Denzler and L. Fredendal, “Evaluating Variance Control, Order Review/Release and Dispatching: a Regression Analysis,” *International Journal of Production Research*, Vol. 32, No. 5, pp 1045-1061, 1994.

[28] Philipoom, P. R., M. K. Malhotra and J. B. Jensen, “An Evaluation of Capacity Sensitive Order Review and Release Procedures in Job Shops,” *Decision Sciences*, Vol. 24, No. 6, pp 1109-1133, 1993.

[29] Plossl, G. W. and O. W. Wight, “Capacity Planning and Control,” *Production and Inventory Management*, Vol. 14, No. 3, pp 31-67, 1973.

[30] Pritsker, A. A. B., Introduction to Simulation and SLAM II, A Halsted Press Book, John Wiley & Sons, New York; Systems Publishing Corporation, West Lafayette, Indiana, 1986.

[31] Ragatz, G. L. and V. A. Mabert, “An Evaluation of Order Release Mechanisms in a Job-Shop Environment,” *Decision Sciences,* Vol. 19, No. 1, pp 167-189, 1988.

[32] Ramasesh, R., “Dynamic Job Shop Scheduling: a Survey of Simulation Research,” *Omega*, Vol. 18, No. 1, pp 43-57, 1990.

[33] Roderick, L. M., D. T. Phillips and G. L. Hogg, “A Comparison of Order Release Strategies in Production Control Systems,” *International Journal of Production Research*, Vol. 30, No. 3, pp 611-626, 1992.

[34] Sabuncuoglu, I. and H. Y. Karapinar, “A Load-Based and Due-Date Oriented Approach to Order Review/Release in Job Shops,” *Decision Sciences*, Vol. 31, No. 2, pp 413-447, 2000.

[35] Sabuncuoglu, I. and H. Y. Karapinar, “Analysis of Order Review/release Problems in Production Systems,” *International Journal of Production Economics,* Vol. 62, pp 259-279, 1999.

[36] Salegna, G. J. and P. S. Park, “Workload Smoothing in a Bottleneck Job Shop,” *International Journal of Operations and Production Management,* Vol. 16, No. 1, pp 91-110, 1996.

[37] Van Ooijen, H. P. G., “Delivery Performance Improvement by Controlled Work-order Release and Work-center Load Balancing, “*International Journal of Production Economics,* Vol. 56-57, pp 661-675, 1998.

[38] Wight, O. W., “Input/Output Control: a Real Life Handle On Lead Time,” *Production and Inventory Management*, Vol. 11, No. 3, pp 9-31, 1970.

[39] Wisner, J. D., “A Review of the Order Release Policy Research,” *International Journal of Operations and Production Management,* Vol. 15, No. 6, pp 25-40, 1995.

[40] Zozom Jr., A., T. J. Hodgson, R. E. King, A. J. Weintraub and D. Cormier, Integrated Job Release and Shop-floor Scheduling to Minimize WIP and Due-dates," *International Journal of Production Research*, Vol. 41, No. 1, pp 31-45, 2003.