Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

BEHAVIOR OF AN ORDER RELEASE MECHANISM IN A MAKE-TO-ORDER MANUFACTURING SYSTEM WITH SELECTED ORDER ACCEPTANCE

Amitava Nandi

Optical Components Division
Nortel Networks
185 Corkstown Road
Nepean, Ontario, K2H 8V4, CANADA

Paul Rogers

Dept. of Mechanical & Manufacturing Engineering
2500 University Drive N.W.
University of Calgary
Calgary, Alberta T2N 1N4, CANADA

ABSTRACT

The value of holding orders in a pre-shop pool, prior to their release to the factory floor, is a somewhat controversial topic. This is especially true for make-to-order manufacturing systems, where, if capacity is fixed and exogenous due dates are inflexible, having orders wait in a pre-shop pool may cause the overall due date performance of the system to deteriorate. In such circumstances, selective rejection of orders offers an alternative approach to dealing with surges in demand whilst maintaining acceptable due date performance. This paper reports on the behavior of such a make-to-order manufacturing system under a control policy involving both an order release component and an order acceptance/rejection component.

1 INTRODUCTION

In modern manufacturing, the need to maintain low inventory, short lead-times and high due date performance has gained in importance when compared with the conflicting desire to achieve high resource utilization. Today, competing in a global market demands not just high quality but also dependable deliveries. As Miller and Roth (1988) have noted, the ability to distinguish one manufacturer from another based on high product quality has been replaced by time and service-related capabilities. However, the ability to reliably meet due dates, particularly in situations where allowable flow times are short is coupled with the ability to keep work in progress (WIP) levels under control. In a seminal work, Little (1961) showed the theoretical relationship between WIP, throughput and manufacturing lead-time, highlighting that there is a critical level of WIP that should not be exceeded in a manufacturing system if lead-time guarantees are to be achievable. In a similar vein, early work by Wight (1970) advocated that work should not be added to the shop at a rate that exceeds the rate at which the work can be completed if low and predictable flow times are to result. Thus, early and fundamental work has clearly shown that managing and controlling WIP inventories requires well-defined Order Review and Release (ORR) strategies.

Over the last 30 years, ORR has become a recognized topic of research, with many authors and practitioners identifying that ORR can provide a number of benefits to the operation of manufacturing systems. However, some researchers have criticized and challenged the benefits of ORR and concluded that although judicious release of orders to the shop floor certainly reduces the average overall *manufacturing* lead-time of a shop, it may increase the overall *customer* lead-time, thus yielding poorer delivery performance. This is especially true when demand spikes occur, since accepting all customer orders and making them wait in an order release pool (ORP) until the shop floor load is small enough can result in a very large ORP with certain orders spending considerable time in the pool before they are released. As Melnyk et al. (1994) have commented, the order release mechanism alone is incapable of reducing the system (or customer) flow time, which makes it necessary to have a good planning system to feed the orders into the order release pool selectively.

It is these fundamental insights into the nature of flow in resource-constrained systems under varying demand that motivated the work reported in this paper exploring the potential benefits of judiciously rejecting certain orders at times of high congestion. The objective of this paper is to examine the behavior of such a two-stage input control system where orders are judiciously accepted and where the accepted orders might still be delayed before releasing them to the shop floor at an opportunistic time. This paper is structured as follows. The next section highlights certain related research in the general area of ORR. After this, the hypothetical system, used in this research as a test bed, is described in detail, including material on the acceptance/rejection rule and the order release rule that have been implemented. Selected experimental results are then provided, which are followed by a concluding section that also includes some material indicating future research.

2 RELATED RESEARCH

Research in ORR can be traced back to the 1960s with a substantial volume of research appearing since then. Researchers explored ORR, in terms of holding orders and releasing them opportunistically, and reported mixed results. A number of researchers have reported on the positive benefits of ORR in terms of improved delivery performance while others have reported to the contrary.

Ackerman (1963) observed higher reliability in on-time completion in the case of immediate release, when compared to a Backward Infinite Loading rule in combination with due date oriented dispatching policies. Irastorza and Deane (1974) found that in a job shop environment, a Finite Forward Loading (FFL) rule outperformed immediate release when paired with the shortest processing time dispatching rule in terms of average tardiness and lateness variance. Shimoyashiro et al. (1984) found FFL to outperform immediate release in terms of job lateness. Ragatz and Mabert (1988) tested five release mechanisms with four dispatching rules in a job shop system. The results showed that controlling the release of work to the shop floor can substantially improve the performance of the system in terms of total shop cost (where this consists of late delivery cost and holding cost for both work in progress and finished-goods inventory) and also deviation from due dates. Bobrowski and Park (1989) studied the effects of several release mechanisms on a dual resource constrained job shop in conjunction with two dynamic, due date oriented dispatching rules and two levels of due date tightness. This research indicated that the release mechanisms produce significant performance improvements over an immediate release rule. Ashby and Uzsoy (1995) reported on the development of a number of scheduling policies integrating order release, group scheduling, and order sequencing for a group technology cell in the presence of sequence-dependent setup times and dynamic job arrivals. Their results showed that the new scheduling policies, which consider setup times as well as due dates in both order release and job sequencing decisions, substantially improve due date performance.

In conflict with the above positive results, many researchers have reported experimental results showing that using an ORP and ORR mechanisms (as opposed to immediate release) can lead to worse performance. The most notable work along this theme includes Melnyk and Ragatz (1989), Baker (1984), Bertrand (1983a, 1983b), and Kanet (1988). According to these researchers, although the introduction of order release mechanisms reduces the manufacturing lead-time, it might not reduce customer order lead-time or system flow time. Utilizing ORR and an ORP simply shifts waiting time from the shop to the order release pool. The fact that the overall system flow time cannot be reduced by the ORR mechanism alone has also been supported by Melnyk and Ragatz (1988) and by Melnyk et al. (1994).

Melnyk et al. (1994) and Fredendall and Melnyk (1995) concluded that the performance of an ORR system is strongly dependent on the presence of variance control at both the planning system and the shop floor levels. If the order release pool is exposed to high variability of external workload, ORR cannot release all of the work in a timely fashion, and waiting time in the order release pool increases. Similarly when the variance of time on the shop floor is high, ORR becomes overwhelmed, and cannot make "good" order release decisions.

There is another, smaller, school of research in ORR, where just the arrival process is controlled by deciding if an incoming order from the customer will be accepted or rejected. In the research on queuing theory, Lippman (1975), Lippman and Ross (1971), Miller (1969), and Scott (1969, 1970) are examples where this kind of "extreme" control on accepting or rejecting customers was exercised. On the experimental side, Wester et al. (1992), ten Kate (1994), and Philipoom and Fry (1992) also tested their systems with this kind of extreme control.

In the light of the above discussion, it is interesting to investigate the behavior of a two-stage input control system, where the external orders are judiciously accepted and the accepted orders are opportunistically released to the shop floor.

3 RESEARCH METHODOLOGY

In the research reported here, a hypothetical manufacturing system has been explored as a test bed to evaluate various control policies under various operating conditions. This section describes this manufacturing system, and details the order acceptance/rejection rule, the order release rule, and the relevant performance measures of the system.

3.1 Description of the System

The manufacturing system considered here, is a make-to-order manufacturing system where each order arriving from the external customers is for one unit of a specified product (there being a broad product range). Order inter-arrival times are sampled from a Gamma distribution, permitting both the average arrival rate and the arrival process variability to be easily and independently varied (so that the acceptance/rejection rule can be tested under different conditions). The acceptance of a candidate order is guided by an order *acceptance/rejection rule* (this rule is described in detail in subsection 3.2.2). It is assumed that all customer orders are of equal urgency and that the due date for each order is set at the time of order acceptance to equal the order's arrival time plus a constant flow allowance (denoted by the acronym FTA in the following text).

Accepted orders are temporarily held in a pre-shop pool called the Order Release Pool (ORP) before releasing them to the shop. The orders in the ORP are released to the shop floor according to the order release rule as described in subsection 3.2.3.

The manufacturing facility processing the orders comprises 10 machines arranged in 4 workstations with the even numbered workstations having 3 machines and the odd numbered workstations having 2 machines. A job can visit each workstation no more than once and, further, is constrained to visit any workstations visited in increasing numerical sequence. This restriction results in there being 143 different routes through the facility, each being associated with a unique job type (note that the total number of operations will vary between 1 and 4). Each of the 10 machines has a known mean processing time, with actual times being sampled from a Gamma distribution (again, permitting acceptance/rejection rules to be easily evaluated under different conditions in terms of average processing rate and service process variability). The jobs in each machine queue are prioritized according to the simple first-in-system-first-served dispatching rule wherein the job that has entered into the system (not the local queue) the earliest is selected.

This system can be tested under various environmental conditions, which are set by choosing levels for a number of experimental factors including: (i) overall demand level; (ii) demand level variability; (iii) process time variability; and (iv) due date tightness.

3.2 Description of the Two-Stage Control Policy

While the manufacturing system is in operation, a two-stage control policy makes decisions regarding order acceptance and order release as necessary. When a customer attempts to place an order, with the manufacturer deciding whether or not to accept the order, the accept/reject decision is taken depending on an *acceptance/rejection rule*. In general, a subsequent decision is also necessary to decide which of the accepted orders will be released from the ORP and what is the most appropriate release time for the order. This decision is taken on the basis of an *order release rule*.

It should be noted that this paper is focusing on only a subset of the entire model that was developed. The full model is highly modular and includes a number of alternative rules for each of the two types of decision that must be made (i.e., alternative accept/reject rules and alternative order release rules). The focus of this paper is on the performance of a particular accept/reject rule in concert with a particular order release rule. Further information on the full range of control policies implemented in the model can be found in Nandi (2000).

3.2.1 Quantities Involved in Different Rules

The control rules under investigation in this paper make use of dynamically-updated information on the state of the manufacturing system being controlled. Certain key components of this state information are particularly relevant to the operation of the rules, and these are defined as follows:

- *TotAccL*: The total amount of estimated remaining work content of all orders that have been accepted.
- *AccLOM(i)*: This is defined for each machine *i* as the portion of *TotAccL* which must be performed at that machine.
- *TotRelL*: This is defined as the total amount of estimated remaining work content of all of the *released* jobs in the system at this moment.
- *RelLOM(i)*: This is defined, for each machine *i*, as the portion of *TotRelL* which must be performed at machine *i*.

All of the above quantities are updated, as required, whenever: (i) a new job is accepted; (ii) a job is released from the ORP to the shop floor; (iii) a machine begins/ends processing an operation of a job. E.g., when a new job is accepted, the *AccLOM(i)* values for all machines visited by the job increase by the expected processing time of the job at those machines (while the *TotAccL* increases by the total expected processing time, over all operations). Also, whenever an operation starts or ends at a machine, all four quantities above are decremented by half of the job's expected operation time.

3.2.2 Acceptance/Rejection Rule

The acceptance/rejection rule that has been explored in the work reported in this paper can be stated as follows:

> Accept the order if the accepted load on the busiest machine on the candidate order's route is less than a specified maximum value.

For a job of type j, arriving at time t, if $AccLOM(i)_t < RL$, for all $i \subset q_j$, then the order is accepted, otherwise it is rejected, where:

$AccLOM(i)_t$	=	The value of *AccLOM(i)* at the time of the arrival of the order, t.
q_j	=	The set of all machines on the route of an order of type j.
RL	=	A constant (and the sole parameter of this particular acceptance/rejection rule).

As the machine with the heaviest workload would tend to delay the completion of an order more so than less loaded machines, controlling the input of orders based on this critical machine may make more sense than looking at the total load on the entire shop when making accept/reject decisions. As should be obvious from the above description, increasing this rule's control parameter, *RL*, should result in less orders being rejected, but may also lead to increased congestion and hence increased order tardiness.

3.2.3 Order Release Rule

The order release rule that has been explored in the work reported in this paper can be stated as follows:

> Release the order if the released load on the busiest machine on the order's route is less than a specified maximum value.

If $RelLOM(i)_t < CL$, for each $i \subset q_j$, an order is released from the order release pool (otherwise it continues to be held in the order release pool), where:

$RelLOM(i)_t$	=	The value of $RelLOM(i)$ at time t.
q_j	=	The set of all machines on the route of the order (of type j) in question.
CL	=	A constant (and the sole parameter of this particular order release rule).

Although orders can be released at any time, and in any quantity, if permitted by this order release rule, the way in which the state information is dynamically updated means that the order release decision needs to be revisited only each and every instant at which an operation at a machine starts or ends (see 3.2.1). Further, when the order release decision is revisited, each job in the ORP is considered for release in a specific priority sequence determined by the well-known *minimum slack per operation* priority rule. If two orders have the same slack per operation and both are eligible to be released, the tie is broken on the basis of the earlier entry time into the system.

There is also a special arrangement to release an order from the ORP forcibly, if it is not released "normally" by the active order release rule within a certain interval after its arrival. This duration is individually determined for each order on its arrival to the ORP in the following manner. At the moment a new order enters the ORP, the average waiting time in any shop queue experienced by a similar order (similar with respect to the number of steps involved) is noted and is multiplied by the number of steps involved in the new order. If this product is less than the flow time allowance of the new order, the said duration is set equal to this product. Otherwise the said duration is set to zero, i.e., the new order is released from the ORP immediately.

This rule considers detailed information on the state of the system at the time of the decision. Specifically, it considers the maximum estimated released load at any machine on an order's route at the time of checking the possibility of the order's release and it tries to keep congestion under control by keeping the load at each individual machine below a defined maximum limit. As should be obvious from the above description, increasing this rule's control parameter, *CL*, should result in orders spending less time in the release pool, but the impact of varying *CL* on total flow time cannot in general be predicted (since shop floor flow times might increase by more than any decrease in ORP times).

3.3 Performance Measures of the System

The overall goal of workload control systems like the two-stage one being explored here is to improve customer satisfaction. With this in mind, there are two components of customer *dis*satisfaction that are considered as relevant in the context of the specific system being studied here. Specifically, customers will be increasingly dissatisfied as either of the following quantities increases:

- *%R*: the percentage of arriving orders that are rejected by the acceptance/rejection rule.
- *%T*: the percentage of arriving orders that are accepted but which do not complete on time.

Thus, one appropriate principal performance measure for this system can be defined as the sum of these two quantities (*%R+%T*) since it can serve as an indication of the total proportion of customer orders that are not dealt with "satisfactorily". This measure is also useful from the manufacturer's perspective in that *%R* might represent lost revenue while *%T* might represent tardiness penalty costs. It is thus desirable for the system to aim to reduce the sum of the two kinds of loss. An additional performance measure that is also monitored to assess due date performance in a different manner is *average order tardiness*.

Five additional performance measures, four of which represent different views or components of flow time and inventory levels, are also reported on in this paper. These provide additional insight into the behavior of the manufacturing system under different conditions and under different values of rule control parameters. The additional five measures can be defined as follows:

- *OFT*: is the average overall flow time of orders from the time of arrival until the time of completion of the last operation.
- *OMLT*: is the average overall manufacturing lead time of orders from the time of release to the shop floor until the time of completion of the last operation.
- *OWTORP*: is the average overall waiting time in the ORP from the time of arrival until the time of release to the shop floor. Of course, *OFT* = *OMLT* + *OWTORP*.
- *TotWIP*: is the average total accepted (but not yet complete) workload in the system, in terms of the total hours of remaining processing time of the accepted jobs. This measure is the time average

of the dynamically-updated quantity *TotAccL* defined in 3.2.1.

- *IRL*: is the average total released (but not yet complete) workload in the system, in terms of the total hours of remaining processing time of the released jobs. This measure is the time average of the dynamically-updated quantity *TotRelL* defined in 3.2.1.
- *Util*: is the average machine utilization for the system. Clearly, this will be related to the earlier measure *%R* since the more jobs are rejected, the lower the machine utilization must be.

4 EXPERIMENTAL RESULTS AND ANALYSIS

This section reports on a set of experiments conducted on the test bed described in section 3. The objective of this set of experiments, was to study the behavior of the two-stage input control mechanism, in terms of a number of important performance measures, when the control limits involved in the acceptance/rejection rule and the order release rule are varied. The performance measures included are *%R+%T*, *%R*, *%T*, *Tardiness*, *OFT*, *OMLT*, *OWTORP*, *TotWIP*, *IRL*, and *Util*.

In the specific experiments run, *RL* was varied through {10, 15, 20, 25, 30, 35} hours and for each value of *RL*, *CL* was varied through {5, 10, 15, 20, 25, 30, 35} hours, so that there are 42 (6 levels x 7 levels) different experiments in total. For each of these experiments, the overall demand level was kept at 85% (in terms of average machine utilization if *all* jobs were accepted) and both the demand level variability and the process time variability were set to 10% (expressed in terms of the coefficient of variation of the respective random quantity, i.e., interarrival times and service time). The demand level is controlled by setting the mean of the inter-arrival time distribution (knowing the mean of the process time distribution and the product mix). Further, the flow allowance parameter noted earlier, *FTA*, was set to equal 30 hours. Note that this parameter effectively determines the level of *due date tightness* of the system (with the chosen value yielding approximately 37.1% tardy orders, if all orders were accepted and were released to the shop floor immediately, under the already specified demand level and variability levels).

In carrying out each experiment, the system has been simulated for 5 replications, during each of which statistics were collected for 72000 hours after a warm-up period of 11520 hours (these values were chosen to as to yield confidence interval half-widths for the primary performance measure of interest, *%R+%T*, whose size is within 0.1% of the measure's mean values).

The detailed results can be seen in Table 1 on the next page of the paper. The subsections below attempt to highlight the key qualitative and quantitative aspects of the behavior of the system under the two-stage control policy.

4.1 Effect on Average Released Load (*IRL*) and Average Accepted Load (*TotWIP*)

One important aspect of performance concerns the inventory levels in the manufacturing system, both in terms of the load on the shop floor (*IRL*) and the load in the system as a whole (*TotWIP*). This subsection includes observations concerning how these two quantities are impacted by the choice of workload control policy parameters.

Firstly, we can make the following observations on how performance varies as the order release rule's control parameter, *CL*, is varied for a constant value of the acceptance/rejection rule's control parameter, *RL*:

- For a particular value of *RL*, the average incomplete released load (*IRL*) increases as *CL* increases. This is as we would expect, since allowing more released workload on each machine on the shop floor (i.e., a higher value of *CL*) can only increase the total average shop floor workload.
- Similarly, for each value of *RL*, the average accepted load (*TotWIP*) decreases as *CL* increases. In general, we cannot say that this is expected (since this would be prejudging the impact of order release, and the research literature documents that fact that delayed release is sometimes of benefit and sometimes not). It is probable that the results obtained here *(TotW*IP decreasing with increasing *CL* for all *RL* values) are due to the low overall demand level and that if demand were higher, there might be some situations where increasing CL might increase *TotWIP*.
- Further, for any *RL*, if *CL* is set high enough (specifically, if $CL \geq RL$) the released load will equal the accepted load ($IRL = TotWIP$) since a high enough release limit for the order release rule is equivalent to a policy of immediate release of accepted orders.

Next, we can make the following observations on how performance varies as the value of the acceptance/rejection rule's control parameter, *RL*, is varied for a constant value of the order release rule's control parameter, *CL*:

- For a particular *CL*, as *RL* increases, average values of both *IRL* and *TotWIP* increase. This is as might be expected since increasing *RL* (for constant *CL*) is equivalent to allowing more orders to be accepted which can only increase congestion in the system as a whole.

Table 1: Performance Measures for Different Combinations of the Control Policy Parameters (*RL* and *CL*)

RL (hr)	*CL (hr)*	*%R+%T*	*%R*	*%T*	*Tardiness (hr)*	*OFT (hr)*	*OMLT (hr)*	*OWTORP (hr)*	*TotWIP (hr)*	*IRL (hr)*	*Util (%)*
10	0	23.74	23.74	0.00	0.227	13.11	11.85	1.26	54.09	46.46	63.05
	5	23.12	23.12	0.00	0.000	12.86	11.91	0.94	53.14	47.04	63.64
	10	21.30	21.30	0.00	0.000	12.42	12.42	0.00	50.70	50.70	65.28
15	0	15.72	15.68	0.05	1.115	15.77	13.82	1.95	75.25	62.08	70.44
	5	15.42	15.38	0.04	1.201	15.53	13.84	1.68	74.02	62.22	70.72
	10	14.19	14.15	0.04	0.968	14.97	14.25	0.72	70.65	65.27	71.84
	15	13.32	13.29	0.03	1.256	14.66	14.66	0.00	68.19	68.19	72.62
20	0	11.57	8.99	2.57	1.664	19.42	16.50	2.92	103.98	82.64	76.62
	5	11.32	8.86	2.47	1.689	19.20	16.51	2.70	102.73	82.53	76.73
	10	10.46	8.24	2.23	1.693	18.67	16.79	1.87	99.15	84.43	77.32
	15	9.50	7.52	1.98	1.698	18.20	17.23	0.97	95.51	87.63	77.98
	20	8.73	6.90	1.82	1.684	17.84	17.84	0.00	91.54	91.54	78.54
25	0	15.36	6.49	8.86	2.416	21.61	18.11	3.50	120.52	94.19	78.93
	5	14.90	6.37	8.53	2.401	21.39	18.10	3.28	119.18	93.93	79.04
	10	13.73	5.98	7.75	2.422	20.83	18.30	2.53	115.28	95.09	79.40
	15	12.33	5.46	6.87	2.385	20.25	18.64	1.61	110.79	97.55	79.88
	20	10.72	4.92	5.80	2.428	19.66	19.34	0.32	104.74	101.99	80.39
	25	10.44	4.75	5.70	2.418	19.55	19.55	0.00	103.31	103.31	80.55
30	0	32.22	3.71	28.51	3.754	24.91	20.52	4.39	144.98	110.89	81.49
	5	31.45	3.65	27.80	3.734	24.69	20.51	4.18	143.57	110.54	81.56
	10	29.61	3.44	26.16	3.709	24.12	20.63	3.49	139.40	110.94	81.76
	15	27.02	3.12	23.90	3.651	23.38	20.83	2.55	133.68	112.31	82.06
	20	23.73	2.71	21.03	3.591	22.60	21.44	1.16	126.11	116.21	82.44
	25	22.59	2.55	20.04	3.567	22.31	21.79	0.52	122.80	118.33	82.59
	30	21.54	2.43	19.11	3.526	22.05	22.05	0.00	119.79	119.79	82.70
35	0	39.76	2.70	37.06	4.966	26.51	21.70	4.81	156.55	118.72	82.45
	5	39.22	2.66	36.56	4.972	26.32	21.70	4.62	155.28	118.41	82.49
	10	36.94	2.50	34.44	4.876	25.67	21.73	3.93	150.36	118.09	82.63
	15	34.19	2.28	31.91	4.836	24.92	21.91	3.01	144.59	119.28	82.84
	20	30.11	1.99	28.11	4.783	23.96	22.40	1.55	135.72	122.44	83.10
	25	28.38	1.83	26.55	4.689	23.57	22.72	0.85	131.75	124.42	83.26
	30	27.02	1.71	25.31	4.651	23.22	23.11	0.11	127.50	126.52	83.37
	35	26.62	1.69	24.93	4.656	23.16	23.16	0.00	126.80	126.80	83.38

4.2 Effect on System Flow Time (*OFT*), Shop Floor Flow Time (*OMLT*), and ORP Time (*OWTORP*)

Another important aspect of performance concerns the various flow times (in passing through the ORP, the shop floor, and the system as a whole). This subsection includes observations concerning how the three flow time quantities are impacted by the choice of workload control policy parameters. There will be some similarities here with the results for the "inventory" performance measures but the effects will not be identical since the throughput rate of the system is also changing as the control policy parameters are varied (i.e., the λ in Little's Law is not constant for this system; if it were, then flow times would simply be linearly related to inventory levels).

Firstly, we can make the following observations on how performance varies as the order release rule's control parameter, *CL*, is varied for a constant value of the acceptance/rejection rule's control parameter, *RL*:

- For a particular value of *RL*, the average time in the order release pool (*OWTORP*) decreases as *CL* increases. This is as we would expect since allowing work onto the shop floor earlier should decrease the time spent in the ORP.
- On the other hand, for each value of *RL*, the time spent on the shop floor (*OMLT*) increases as *CL* increases. This too is as expected since allowing work onto the shop floor earlier can only increase congestion which can only increase shop floor flow time.

- For all the experimental scenarios reported in this paper, the total flow time (*OFT*) decreased as *CL* was increased, for all values of *RL*. For different values of the overall demand level, this might not, in general, be the case.

Next, we can make the following observations on how performance varies as the value of the acceptance/rejection rule's control parameter, *RL*, is varied for a constant value of the order release rule's control parameter, *CL*:

- At a particular *CL*, if *RL* decreases, all three of the flow time performance measures (*OFT*, *OMLT* and *OWTORP*) decrease. This is as expected since allowing more jobs to be accepted can only increase congestion within the system, thus both ORP and shop flow times can only increase (and since total flow time is the sum of these two components, it must increase too).

4.3 Effect on Percent Rejected (*%R*), Percent Tardy (*%T*) and Total Loss Percentage (*%R+%T*)

The primary performance measures of interest in this study are the two components of "loss" (due to rejected jobs and tardy jobs) and their sum which makes up the overall measure we are seeking to minimize via our choice of optimal values for the control policy parameters. This subsection includes observations concerning how the three losses are impacted by the choice of workload control policy parameters.

Firstly, we can make the following observations on how performance varies as the order release rule's control parameter, *CL*, is varied for a constant value of the acceptance/rejection rule's control parameter, *RL*:

- For a particular value of *RL*, both components of loss (and therefore their sum too) decrease as *CL* increases. This is as might be expected since releasing work to the shop floor earlier should allow it to be completed, on average, faster. Although this is true for all experiments reported here, we do not expect this to be generally true since there may be other cases, e.g., at higher demand levels, where earlier release might worsen performance.

Next, we can make the following observations on how performance varies as the value of the acceptance/rejection rule's control parameter, *RL*, is varied for a constant value of the order release rule's control parameter, *CL*:

- For a particular value of *CL*, the percentage of jobs rejected (*%R*) decreases as *RL* increases. This is as expected since allowing more work into the system can only reduce the amount that does not enter. Were *RL* increased to a high enough value, this would be equivalent to a "*full acceptance*" scenario wherein no jobs at all are rejected.
- Conversely, although we are free to increase *RL* without limit, thereby reducing our "*rejection losses*", doing so increases congestion in the system and leads to deteriorating due date performance an increasing percentage tardy (*%T*). As might be expected, the numeric results show that for all levels of *CL*, *%T* increases as *RL* increases.
- For all the experiments reported here, since *%R* decreases at a decreasing rate and *%T* increases at an increasing rate as *RL* increases, for each value of *CL* there is an optimal value of *RL* that yields the least overall loss (*%R+%T*).

4.4 Effect on Utilization and Tardiness

Since average utilization is closely (and inversely) related to the percentage of jobs that are rejected, how *Util* varies with *RL* and *CL* can be inferred from the results in the previous subsection:

- For a particular value of *RL*, average machine utilization (*Util*) increases as *CL* increases since releasing work earlier allows it to be completed faster, on average, which allows additional jobs to be accepted. Note that this relationship may not be present in general since at higher overall demand levels, earlier release could conceivably lead to greater congestion and lower flow times.
- More significantly, for a particular value of *CL*, *Util* increases as RL increases, since allowing more jobs to be accepted directly impacts the machine utilization.

The average order tardiness (*Tardiness*) appears to be much more strongly influenced by the control parameter of the acceptance/rejection rule than it is by that of the order release rule. As the results in Table 1 show, there is very little variation of tardiness with *CL* (for constant *RL*) for the set of experiments conducted whereas for constant *CL*, *Tardiness* increases significantly with increasing *RL*. The latter phenomenon is as might be expected since allowing more jobs to be accepted can only increase congestion which can only worsen due date performance.

5 CONCLUSIONS

An interesting conclusion from the limited set of experiments reported upon here is that for the specific scenario simulated, it appears to be better to release orders to the shop floor as soon as they are accepted by the system if the objective is to minimize the sum of rejection losses and tardiness losses (in terms of *%R+%T*). While this agrees with some of the earlier literature reviewed above (see section 2) it does not mean that immediate release will *always* be the best (i.e., there may very well be other combinations of experimental factors for which non-immediate release can outperform immediate release). Further experimentation is required to see what, if any, combinations of experimental factors (i.e., demand, arrival and process variability, due date tightness) lead to some form of delayed release yielding the best overall performance.

From the set of experiments carried out, the best values for the control parameters (for this set of experimental factors only) appears to be *RL*=20 and *CL*=20. The results clearly show that under these circumstances, giving the system the ability to reject jobs (via the implemented acceptance/rejection rule) can significantly outperform a system that must accept all arriving jobs. Further experimentation is ongoing along two directions: (i) exploration of how the optimal choice of control parameters depends on the experimental factors; (ii) development, implementation, and testing of alternative acceptance/rejection rules to identify rules that might outperform the one whose performance is reported upon in this paper.

ACKNOWLEDGMENTS

Thanks are due to the Natural Sciences and Engineering Research Council (NSERC) of Canada for partially funding some of the research reported in this paper.

REFERENCES

Ackerman, S.S. 1963. Even flow, a scheduling method for reducing lateness in the job shop. *Management Technology* 3 (1): 20-32.

Ashby, J.R., and R. Uzsoy. 1995. Scheduling and order release in a single-stage production system. *Journal of Manufacturing Systems* 14 (4): 290-306.

Baker, K.R. 1984. Sequencing rules and due-date assignments in a job shop. *Management Science* 30: 1093-1104.

Bertrand, J.W.M. 1983a. The use of workload information to control job lateness in controlled and uncontrolled release production systems. *Journal of Operations Management* 3 (2): 79-92.

Bertrand, J.W.M. 1983b. The effect of workload dependent due-dates on job shop performance. *Management Science* 29 (7): 799-816.

Bobrowski, P.M., and P.S. Park. 1989. Work release strategies in a dual resource constrained job shop. *Omega* 17 (2): 177-188.

Fredendall, L.D., and S.A. Melnyk. 1995. Assessing the impact of reducing demand variance through improved planning on the performance of a dual resource constrained job shop. *International Journal of Production Research* 33 (6): 1521-1534.

Irastorza, J.C., and R.H. Deane. 1974. A loading and balancing methodology for job shop control. *AIIE Transactions* 6 (4): 302-307.

Kanet, J.J. 1988. Load-limited order release in job shop scheduling systems. *Journal of Operations Management* 7 (3): 44-58.

Lippman, S.A. 1975. Applying a new device in the optimization of exponential queuing systems. *Operations Research* 23: 687-710.

Lippman, S.A. and S. Ross. 1971. The streetwalker's dilemma: a job shop model. *SIAM Journal of Applied Mathematics* 20: 336-344.

Little, J.D.C. 1961. A proof of the queuing formula $L = \lambda W$. *Operations Research* 9: 383-387.

Melnyk, S.A., and G.L. Ragatz. 1988. Order review/release and its impact on the shop floor. *Production and Inventory Management Journal* 29 (2): 13-17.

Melnyk, S.A., and G.L. Ragatz. 1989. Order review/release: research issues and perspectives. *International Journal of Production Research* 27 (7): 1081-1096.

Melnyk, S.A., K.C. Tan, D.R. Denzler, and L. Fredendall. 1994. Evaluating variance control, order review/release and dispatching: a regression analysis. *International Journal of Production Research* 32 (32): 1045-1061.

Miller, B.I. 1969. A queueing reward system with several customer classes. *Management Sciences* 16:234-245.

Miller, J.G., and A.V. Roth. 1988. *Manufacturing Strategies: executive summary of the 1988 North American manufacturing futures survey*. Boston, MA: Boston University.

Nandi, A. 2000. *Input control strategies for make-to-order manufacturing systems via order acceptance/rejection*. Ph.D. Thesis, Department of Mechanical & Manufacturing Engineering, University of Calgary.

Philipoom, P.R., and T.D. Fry. 1992. Capacity-based order review/release strategies to improve manufacturing performance. *International Journal of Production Research* 30 (11): 2559-2572.

Ragatz, G.L., and V.A. Mabert. 1988. An evaluation of order release mechanisms in a job shop environment. *Decision Sciences* 19 (2): 167-189.

Scott, M. 1969. A queueing process with some discrimination. *Management Sciences* 16: 227-233.

Scott, M. 1970. Queueing with control on the arrival of certain types of customers. *CORS Journal* 8: 75-86.

Shimoyashiro, S., K. Isoda, and H. Awane. 1984. Input scheduling and load balance for job shop. *International Journal of Production Research* 22 (4): 597-605.

ten Kate, H.A. 1994. Towards a better understanding of order acceptance. *International Journal of Production Economics* 37: 139-152.

Wester, F.A.W., J. Wijngaard, and W.H.M. Zijm. 1992. Order acceptance strategies in a production-to-order environments with setup times and due dates. *International Journal of Production Research* 30 (6): 1313-1326.

Wight, O. 1970. Input/output control a real handle on lead time. *Production and Inventory Management* 11 (3): 9-31.

AUTHOR BIOGRAPHIES

AMITAVA NANDI is a Process Engineer at Nortel Networks in Nepean, Ontario. His interests include system design and analysis, productivity improvement, manufacturing systems modeling, statistical process control, and multivariate data analysis. He holds a Ph.D. from the University of Calgary (in Canada), and an M.Sc. from Brunel University (in England). He can be contacted by email at <amitava@nortelnetworks.com>.

PAUL ROGERS is an Associate Professor in the Department of Mechanical and Manufacturing Engineering at the University of Calgary. His research and teaching interests include production planning and control systems, manufacturing strategy, distributed and dynamic scheduling in electronics manufacturing, object-oriented modeling for intelligent manufacturing, simulation of intelligent agent systems, and models for the design and analysis of manufacturing systems. He is a Professional Engineer registered with the Association of Professional Engineers, Geologists, and Geophysicists of Alberta (*APEGGA*), a member of *IIE*, *APICS*, and *INFORMS*, and serves on the Editorial Board of the *International Journal of Computer Integrated Manufacturing*. He holds Ph.D. and M.Eng. degrees from Cambridge University in England. He can be contacted by email at <rogers@ucalgary.ca>.