PRIORITY SCHEDULING RULES FOR A COMBINED PRODUCTION-INVENTORY SYSTEM: A SIMULATION ANALYSIS

William L. Berry, Assistant Professor
Krannert Graduate School of Industrial Administration
Purdue University
Lafayette, Indiana 47907

Abstract

One proposal for improving the combined performance of a shop and its associated inventory system is to incorporate inventory information into priority scheduling rules. This paper describes several ways that inventory information can be used for scheduling purposes. It also reports the results of simulation experiments that were conducted to evaluate the gain in performance resulting from the inclusion of inventory data in scheduling rules. The results of these experiments indicate that in designing scheduling rules, an increase in the amount and availability of inventory data is far less important in improving manufacturing performance than the application of a modest amount of inventory theory.

Introduction

The problem of planning, scheduling, and controlling the flow of work in job-shop-like facilities has received important attention in the Industrial Engineering literature. This work has partly been devoted to the investigation of priority scheduling rules -- rules which are used to dispatch work to machines at the time of actual production. In previous investigations these rules have been evaluated with respect to selected measures of shop performance, i.e. machine utilization, the amount of work-in-process inventory, the length of the manufacturing lead times, or the performance against due dates. Many shops, however, maintain an inventory of finished products and a significant part of their production is directed toward replenishing this inventory. In these cases an important criterion of the scheduling function is the combined performance of the shop and its associated inventory system.

One means of improving the effectiveness of the shop and its inventory system jointly is the incorporation of inventory information into the rule used for dispatching work to machines. This proposal involves feeding inventory data back to the shop for the purpose of improving the lead time performance of the individual products manufactured by the shop. Such information may include the use of: due date, on hand inventory, inventory shortage, and reorder point data. The purpose of this paper is to evaluate the gain in overall performance which results from the inclusion of inventory information in priority scheduling rules. We begin by describing several rules which consider inventory information and then report the results of simulation experiments that were conducted to evaluate the gain in performance resulting from the use of such information in making scheduling decisions.

Priority Scheduling Rules

The movement of orders through a shop is regulated by the manner in which scheduling decisions are made at individual machines at the time of actual production. These decisions involve determining the sequence in which orders are processed at each machine. When a priority scheduling rule is used to make sequencing decisions, all of the orders that are waiting to be processed at a machine are ranked according to a priority index assigned by the scheduling rule. Thus, the order having the highest scheduling priority is assigned to the machine when it becomes available to process another order.

The scheduling rule that is employed has an important influence on the operating performance of a shop. Consider, for example, the performance of two simple rules which have been widely investigated: 1) the Shortest Processing Time rule, and 2) the First in The System, First Served rule. When the Shortest Processing Time rule is used, the order requiring the least machine time receives the highest scheduling priority. This rule results in a short average manufacturing lead time, a low level of work-in-process inventory in a shop, and a high machine utilization. The First in The System, First Served rule is considered to be a more equitable procedure, since it schedules the order which has been in the shop the longest next. It results in a high predictability of manufacturing lead times by providing a low variance for this measure.[5,6]

While these measures describe the operating performance of a shop, they do not directly reflect the performance of a shop's inventory system. In extending the performance criteria to include both the shop and its inventory system, we are concerned with the costs of: placing orders, carrying both work-in-process and finished product inventory, and incurring inventory shortages. In the next three sections three scheduling rules will be described which work toward reducing these costs by using inventory information to improve the manufacturing lead time performance for the individual products produced by a shop. These rules are: 1) the Minimum Slack Time Per Remaining Operatio rule, 2) the Critical Ratio rule, and 3) the Two Class Shortest Processing Time rule. These rules are presented in Table 1 along with an example to illustrate the priority index calculation for each rule.

Slack Time Rule

In shops which manufacture made-to-order products, the Slack Time rule is directed toward improving the shop's performance against due dates. It keeps the work moving through a shop so that most orders are delivered on or near their promised delivery date, i.e. there are relatively few very early or very late deliveries.(1)[3,9]

When this rule is used to schedule replenishment orders, due dates can be established by using the same estimate of the manufacturing lead time used to set the reorder point for an individual product. Note, for example, that the 15 day lead time for order number 5 in Table 1 was used to set both the reorder point for this product and the due date for the current replenishment order. In this case only two days of the anticipated 15 day lead time remain, but eight days of shop time are needed to complete the order. Since the only other order that is behind schedule, order number 3, has a larger priority index (and thus a lower scheduling priority), order number 5 would be assigned to the machine next.

The scheduling problem thus becomes one of moving the replenishment orders through the shop so that they arrive in inventory at the end of the expected lead time, i.e. when the on hand inventory drops to the buffer stock level. By trying to get the actual manufacturing lead times to conform closely to the average lead time used to set the reorder point for each product, the Slack Time rule accomplishes two things. First, excessive inventory accumulations or shortages are avoided by timing the receipt of replenishment orders to correspond with the shop performance anticipated by the inventory policy. Second, and more importantly, the improved predictability of manufacturing lead times for individual products, i.e. a reduced lead time variance relative to other scheduling rules, translates into smaller buffer stock inventories and therefore lower inventory related costs.[4]

Critical Ratio Rule

The Critical Ratio rule makes explicit use of inventory information in machine scheduling decisions.[2,14] While the Slack Time rule considers inventory data only at the time that replenishment orders are placed, to establish the delivery date, the Critical Ratio rule relies on up-to-date inventory status information as orders progress through the shop. Specifically, the rule considers two types of information in sequencing orders: 1) the rate at which an item's inventory is being depleted, and 2) the rate at which an order is progressing toward completion. This information is combined to determine a critical ratio for each order waiting to be processed at a machine.

The example shown in Table 1 illustrates the rule's emphasis on speeding up the movement of orders through a shop which have relatively small on hand inventories, inventory shortages, or a substantial amount of work remaining to be completed. A comparison of order numbers 2 and 4 provides an example of this. Order number 4 has the smaller critical ratio since it has an inventory shortage of 10 units and more than half of its shop time remaining. Order number 2, on the other hand, has the same amount of work remaining to be completed, but it has a relatively large on hand inventory; its progress through the shop would therefore be slowed considerably by this scheduling rule. Thus, the Critical Ratio rule, like the Slack Time rule, is designed to accomplish reductions in finished product inventory levels and shortages, but by using more comprehensive and dynamic inventory information.

Two Class Shortest Processing Time Rule

The Two Class Shortest Processing Time rule was reported in a study of scheduling rules for single machine production-inventory systems by Baker.[1] Baker found this rule to be the most effective rule tested in a series of simulation experiments using the two bin inventory policy. It uses two types of information in sequencing orders: 1) operation processing times, and 2) inventory shortage data. The example shown in Table 1 illustrates the use of this rule. Here the two orders with inventory shortages, order numbers 3 and 4, receive a higher scheduling priority than the remaining orders. Of these two orders, the one requiring the least machine time, order number 4, would be processed next.

This two class version of the Shortest Processing Time rule (SPT) is designed to achieve the same shop performance objectives as the SPT rule mentioned previously, but with an improvement in inventory shortage costs.(2) That is,

---

(1) In previous job shop scheduling studies this rule has been found to be effective in minimizing the order lateness variance -- where order lateness is defined as the difference between the actual manufacturing lead time and the time allowed for manufacturing according to the order due date.[6]

(2) This rule is similar to the method described by McNaughton for sequencing a fixed set of orders at a single machine to minimize tardiness costs.[12] He sequences orders such that:

$$\frac{P_1}{A_1} \leq \frac{P_2}{A_2} \leq \cdot \cdot \cdot \leq \frac{P_n}{A_n}$$

where: P_n = operation processing time for the n^{th} order.

(Continued)

TABLE 1

Priority Scheduling Example

Order Number	Inventory Information					Shop Information					Priority Index		
	On Hand Inventory	Reorder Point	Safety Stock	Average Daily Usage	Total Manufacturing Lead Time	Accumulated Shop Time	Time Remaining Until Due Date	Current Operation Processing Time	Number of Operations Remaining	Shop Time Remaining	(1) Minimum Slack Time Rule	(2) Critical Ratio Rule	Two Class SPT Rule
1	150	160	10	10	15	1	14	5	2	8	+3	1.76	5
2	110	160	10	10	15	5	10	1	2	8	+1	1.29	3
3	-30	210	10	10	20	14	6	4	2	8	-1	-.36	2
4	-10	160	10	10	15	5	10	2	2	8	+1	-.12	1
5	30	160	10	10	15	13	2	3	2	8	-3	.35	4

Priority Scheduling Rules

Minimum Slack Time Per Remaining Operation: Slack Time is defined as the difference between the time remaining until the promised delivery date and the amount of shop time remaining to complete the order. Here shop time includes the machine processing time, the transport time between machine operations, and an allowance for the expected waiting time at each machine. The order with the smallest slack time per remaining operation value is processed next.

Critical Ratio: The critical ratio equals the on hand inventory divided by the reorder point multiplied by the total manufacturing lead time divided by the shop time remaining. A critical ratio is computed for each order and the order having the smallest critical ratio is processed next.

Two Class Shortest Processing Time: The orders at a machine are divided into two classes, those orders with inventory shortages and those for which there is inventory available. The orders with inventory shortages receive a higher scheduling priority, and of these orders, the one requiring the least machine time is processed next. When none of the orders have inventory shortages, the Shortest Processing Time rule is used.

$$(1)\quad \text{Priority Index} = \frac{\text{Time Remaining Until Delivery Date} - \text{Shop Time Remaining}}{\text{Number of Operations Remaining}}$$

$$(2)\quad \text{Priority Index} = \frac{\text{On Hand Inventory}}{\text{Reorder Point}} \cdot \frac{\text{Total Manufacturing Lead Time}}{\text{Shop Time Remaining}}.$$

(Note: The inventory data is measured in units and time is measured in days.)

it will result in a low average manufacturing lead time (aggregated across all products), a low work-in-process inventory level, and a high machine utilization. The reduction in inventory shortage costs is obtained by giving a higher scheduling priority to those orders with long operation processing times when inventory shortages exist for these products. Such orders are often neglected by the SPT rule, especially when the work load for a machine is heavy.[4] Thus, the use of inventory shortage information by the shop limits the length of time that orders with inventory shortages must wait in machine queues.

Simulation Experiments

The three scheduling rules described in the previous section vary substantially with respect to the type of inventory information considered. Since there may be additional costs associated with gathering and processing this information for scheduling purposes, we are interested in the gain in performance that can be attributed to its use.

We have conducted a series of simulation experiments to evaluate this trade-off between the use of more comprehensive information and the resulting gain in the combined system performance.(3) In these experiments we have compared the performance of the Shortest Processing Time (SPT) and the First in The System, First Served (FISFS) rules with that of the three rules presented above. Since the SPT and FISFS rules work toward reducing inventory related costs by improving the manufacturing lead time statistics without considering inventory data, they serve as a reference for determining the gain in performance which results from incorporating inventory data into scheduling rules.

Experimental Procedure

An important problem in conducting these experiments was that of determining the lot size and reorder point parameters to be used with each rule tested. This problem arises because each scheduling rule exhibits different manufacturing lead time statistics for the individual products. These variations, in turn, require the lot size and reorder point parameters to be selected for each rule so that the combined system costs are minimized. In comparing the combined system performance of two scheduling rules, it would clearly be unacceptable to base comparisons for one rule on results obtained with inventory policy parameters set with the performance statistics of another rule.

This problem is further complicated because of the number of inventory parameters to be determined. In this simulation, with an inventory of 15 products and two inventory parameters for each product, the task of determining a low cost set of parameters is indeed a formidable one.

To solve this problem we employed a procedure which guides the selection of inventory parameters for each rule individually by considering the cost, product demand, and inventory shortage factors for each product. The procedure is an iterative one described by Fetter and Dalleck for determining the inventory parameters jointly for a two bin policy.[7,11] By determining the lot size and reorder point parameters jointly, this procedure allows for cost trade-offs between the two parameters.

In using this procedure, a series of simulation experiments was run for each rule tested. Three measures were recorded for each of the 15 products during a given simulation run: 1) the average number of inventory shortages, 2) the lead time mean, and 3) the lead time variance. At the end of each simulation run the lot size and reorder point parameters for all of the products were recomputed for use in the succeeding run. When no further reductions in the combined system cost were observed, the sequence of runs was terminated. A sensitivity analysis was then conducted to determine if further reductions in the combined system cost could be obtained. In these runs the parameter values for all of the products were incremented by ± 10%. Although totally optimum inventory parameters may not be obtained with this procedure, the final parameter values appeared to reflect the minimum, or near minimum, cost level for each rule.

Experimental Results

The results shown in Table 2 indicate that there are important differences in the overall performance of the five rules. Note, for example, that the manufacturing lead times for the individual products obtained with the Critical Ratio rule are longer and far less predictable than for the Slack Time rule. This decline in lead time performance for the Critical Ratio rule caused the larger inventory system costs and explains much of the difference in performance between the two rules. In addition, the longer lead times for the Critical Ratio rule caused the shop to carry more work-in-process inventory, thereby contributing to larger overall costs.

The poor performance of the Critical Ratio rule is partially offset by the use of larger order quantities for the individual products than were used for the Slack Time rule. Larger orders, placed less frequently, helped to avoid some of the inventory shortages caused by this rule's unfavorable lead time performance. The lower ordering cost and the higher inventory carrying cost shown in Table 2 for this rule reflects this cost trade-off. This advantage can be

---

(2) A_n = tardiness cost per time unit for the n^{th} order.
In the case we considered above, the inventory shortage costs is similar to the tardiness cost.

(3) The details of the simulation model used in these experiments are provided in Appendix A.

Table 2

RESULTS BASED ON AN AVERAGE PRODUCT DEMAND OF 20 UNITS PER PRODUCT PER DAY

Scheduling Rule	Shortest Processing Time	First-In-System, First Served	Minimum Slack Time	Critical Ratio	Two Class: Shortest Processing Time
Shop Criteria:					
Machine Utilization	.846	.832	.831	.809	.831
Average Manufacturing Lead Time (in Days)*	5.8	6.3	6.3	7.8	5.1
Manufacturing Lead Time Standard Deviation*	1.21	.75	.50	1.98	1.08
Work-In-Process Inventory Cost**	$477	519	520	644	436
Inventory System Criteria:					
Mean Demand During Lead Time (in Units)*	115	126	126	156	103
Demand During Lead Time Standard Deviation*	27	19	15	42	24
Ordering Cost**	$376	348	347	287	346
Inventory Carrying Cost**	$650	539	427	740	523
Shortage Cost**	$188	59	91	138	132
Total Inventory Related Costs**	$1214	946	865	1165	1002
Combined System Criterion:					
Total Cost**	$1691	1465	1385	1809	1437

*Averaged Across All 15 Products
**Average Cost Per Day

Table 3

SAMPLE PRODUCT RESULTS WITH AN AVERAGE DAILY PRODUCT DEMAND OF 20 UNITS

Product Number	Manufacturing Lead Time (in days)	Manufacturing Lead Time Standard Deviation (in days)	Average Ordering Cost Per Day	Average Inventory Carrying Cost Per Day	Average Shortage Cost Per Day	Average Total Cost Per Day	Average Inventory Level	Average Number Of Shortages Per Day
Shortest Processing Time Rule:								
1	3.1	.63	$20	$ 16	$ 5	$ 41	63.5	.2
2	4.6	.84	28	36	6	70	53.9	.3
3	20.2	4.50	21	228	12	261	257.6	.6
Two Class - SPT Rule:								
1	3.6	.77	16	21	5	42	84.2	.2
2	4.8	.93	26	37	10	73	56.9	.5
3	6.9	1.38	30	60	12	102	67.6	.6
Minimum Slack Time Per Remaining Operation Rule:								
1	6.8	.44	16	19	5	40	75.5	.2
2	5.9	.72	26	35	9	70	53.2	.4
3	7.1	.51	30	36	11	77	40.8	.6
First In System, First Served Rule:								
1	6.3	.67	16	22	1	39	88.6	.1
2	5.9	.61	26	43	2	71	64.9	.1
3	7.1	.81	30	49	7	86	55.2	.4
Critical Ratio Rule:								
1	7.8	3.26	15	32	6	53	129.5	.3
2	7.5	1.67	21	61	4	86	93.0	.2
3	8.8	1.72	24	63	19	106	71.1	.9

attributed to the procedure we used to determine the inventory policy parameters jointly.

The Slack Time rule's effectiveness is largely because of its ability to control the manufacturing lead times to those anticipated by the inventory policy. This is reflected by the relatively small lead time standard deviation for this rule in Table 2. It should be noted, however, that the FISFS rule performed nearly as well as the Slack Time rule without considering inventory information. The 6% difference in total cost between these two rules is explained by the FISFS rule's 50% larger lead time standard deviation.

The two class version of the SPT rule also performed effectively but for quite a different reason. In comparison with the Slack Time rule, the shorter average manufacturing lead time of the two class SPT rule reduced the work-in-process inventory carrying cost to partially offset the additional inventory system cost. Moreover, the smaller average lead time helped to compensate for the extra inventory system cost caused by the rule's 75% larger lead time standard deviation.

Results with the Shortest Processing Time Rule

The increase in inventory system costs for those products having relatively long operation processing times largely explains the poor performance exhibited by the SPT rule. That these products take much longer to get through the shop is reflected by larger reorder points and increased buffer stock levels. The inventory system performance for three of the 15 products that we studied is presented in Table 3 to illustrate this point. These products were selected because they reflect the range of operation processing times in these experiments.(4)

The inventory system costs for products 1 and 2 are similar for all of the rules investigated, except for the Critical Ratio rule. The long operation processing times for product 3, however, caused its lead time mean and variance to increase substantially when the SPT rule was used. In this case, the need for additional buffer stock increased the inventory carrying cost considerably. It should be noted in Table 3, that the relatively simple two class SPT rule measurably improved the inventory system cost for product 3 when compared with the SPT rule results.

Conclusions

The results of these simulation experiments demonstrate that inventory information can be used in making scheduling decisions to improve the joint performance of a shop and its inventory system. In the case of the Slack Time rule, the use of inventory data only at the time replenishment orders are placed clearly improves the predictability of the manufacturing lead times for individual products. This improvement in the predictability of manufacturing lead times provided an important reduction in the inventory system costs.

The results obtained with the Critical Ratio rule, on the other hand, indicate that an increase in the timeliness of information for scheduling purposes does not necessarily lead to improved performance. The main problem with this rule is its loose control over manufacturing lead times. That is, it allows orders to wait in machine queues until they are needed to avoid inventory shortages. This lengthens the manufacturing lead times, increases the lead time variance, and results in a predictable increase in inventory system costs.

In contrast to the use of current inventory information in the Critical Ratio rule, one must consider the relatively good performance obtained with the FISFS rule which does not consider inventory information in scheduling production. This result has important practical implications for those involved in designing systems for planning and controlling manufacturing operations. One should carefully evaluate the cost of processing inventory information for use in making scheduling decisions. A relatively simple rule which is easily implemented, such as the FISFS rule, may be adequate -- especially when the cost of processing inventory information is significant.

Further research on the problem of priority scheduling in combined production-inventory systems may well focus on improving the two class version of the SPT rule suggested by Baker. In its present form, this rule does not consider differences between products regarding: 1) product demand rates, 2) inventory shortage cost, or 3) the number of outstanding inventory shortages. One way of considering these factors would be to apply the sequencing procedure described by McNaughton for single machine processing systems. By reducing the cost of inventory shortages, this version of the SPT rule may be an even more effective scheduling rule than the Slack Time rule.

APPENDIX A

Simulation Model

The simulation model used in these experiments consisted of a shop with 10 machines and an inventory control system with 15 products. The computer program was written in FORTRAN IV and required 45 to 60 seconds to execute on a CDC 6500, depending mainly on the number of orders in process.

At the start of each simulated day, the program performed the following steps: 1) the orders completed during the previous day were received into inventory, 2) daily demand was generated and the inventory status of each product was updated, 3) the inventory status of

---

(4) The operation processing times for these products are presented in Appendix B.

all products was reviewed and replenishment orders were placed using a two bin inventory policy developed for each product, and 4) the new replenishment orders were input directly into the shop. The program then scheduled the operations at individual machines during the 24 hour time interval between inventory reviews, using the scheduling rule specified. A number of features of the simulation model are significant:

1. The shop routing for each of the 15 products included 10 machine operations which remained fixed during the entire simulation run. The first and last operations were performed on machines 1 and 10, respectively, while the intervening operations were randomly selected from machine numbers 2 through 9. The machine time for each operation was randomly generated from an exponential distribution with a mean of 5.3 hours per 100 units. A one hour machine setup time was also incurred at each operation.

2. All 10 machines were operated 24 hours per day.

3. The following inventory data was used for each of the 15 products: 1) an average daily demand of 20 units (Poisson distributed), 2) an ordering cost of $100 per order to cover the machine setup cost, and 3) an inventory shortage cost of $20 per unit per day short. Inventory shortages were carried as backorders until they could be filled from replenishment orders. The shortage cost of $20 per unit was selected to provide a service level of approximately 90% in these experiments, i.e. a stockout probability of .10.

4. The finished product inventory carrying cost varied among the products from $.25 to $.88 per unit per day. This value was obtained by summing the per unit operation processing times for each product. These same values were used to determine the work-in-process inventory carrying cost for partially completed orders in the shop.

Experimental Procedure

An identical sequence of daily product demands was employed in each of the experiments. The start up phase included: 1) preloading the shop with 25 partially completed orders, 2) randomly generating the on hand inventory balance for each product, and 3) performing an initial run of 20 days. Data collection began on the 21st day and continued for 280 days. Approximately 1000 replenishment orders were completed by the shop during this period.

APPENDIX B

Representative Products

Operation Number	Operation Processing Time Per 100 Units Product 1	Product 2	Product 3
1	7.76	2.47	11.04
2	0.78	6.70	15.58
3	5.14	6.19	5.43
4	0.19	0.53	0.71
5	2.26	4.44	18.57
6	3.29	1.83	5.64
7	0.15	2.34	4.23
8	0.27	24.01	1.56
9	4.01	14.01	19.89
10	0.71	3.29	5.66
	24.56	55.84	88.31

REFERENCES

1. Kenneth R. Baker, "Inventory Control and Priority Dispatching In A Multi-Product Production-Inventory Model," Department of Industrial Engineering, University of Michigan, Technical Report No. 69-28, December, 1969.

2. Richard R. Berger, "Critical Ratio Simplified," Production and Inventory Management, Vol. 11, No. 2, 1970, pp. 15-20.

3. Robert G. Brown, Decision Rules for Inventory Management, Holt, Rinehart, and Winston, 1967, pp. 307-318.

4. Robert G. Brown, "Simulations to Explore Alternate Sequencing Rules," Naval Research Quarterly, June 1968, Vol. 15, No. 2, pp. 281-286.

5. Elwood S. Buffa, Production-Inventory Systems: Planning and Control, Irwin, 1968, Chapters 11 and 12.

6. Richard Conway, William L. Maxwell, and Louis Miller, Theory of Scheduling, Addison-Wesley Publishing Company, 1967.

7. Robert B. Fetter and Winston C. Dalleck, Decisions Models for Inventory Management, Irwin, 1961.

8. Donald Gross and A. Soriano, "The Effect of Reducing Leadtime on Inventory Levels -- Simulation Analysis," Management Science, Vol. 16, No. 2, October 1969.

9. William K. Holstein, "Production Planning and Control: Integrated," Harvard Business Review, May-June 1969, pp. 121-140.

10. The Production Information and Control System, IBM Data Processing Manual, 1968, pp. 69-81.

11. Claude McMillan and Richard F. Gonzalez, *Systems Analysis: A Computer Approach to Decision Models*, Irwin, 1968, Revised Ed., Chapters 7 and 8.

12. R. McNaughton, "Scheduling with Deadlines and Loss Functions," *Management Science*, October, 1959.

13. Ashok R. Pai and Keith L. McRoberts, "Simulation Research in Interchangeable Part Manufacturing," *Management Science*, Vol. 17, No. 12, August, 1971.

14. Arnold O. Putman, E. Robert Barlow, and Gabriel N. Stilian, *Unified Operations Management*, McGraw-Hill Book Company, 1963, pp. 183-185.