HEURISTICS FOR DYNAMIC SCHEDULING OF MULTI-CLASS BASE-STOCK CONTROLLED SYSTEMS

Bora Kat
Department of Industrial Engineering, Middle East Technical University, 06531 Ankara, TURKEY, Phone: +90 (312) 210 4793, E-mail: bora@ie.metu.edu.tr
Zeynep Müge Avşar
Department of Industrial Engineering, Middle East Technical University, 06531 Ankara, TURKEY, Phone: +90 (312) 210 4681, E-mail: muge@ie.metu.edu.tr

Abstract

Dynamic scheduling of an exponential single-server facility processing different types of items one by one is studied for the case of Poisson demand arrivals. Inventories of the items are managed by base-stock policies and backordering is allowed. Structure of the optimal scheduling policy is investigated numerically with respect to a weighted average of the fill rates. Performance of the optimal policy is compared to those of two well-known policies, Longest Queue and First-Come-First-Served, and alternative policies are generated by heuristics in order to approximate the optimal policy.

Keywords: Multi-class, Queueing systems, Dynamic scheduling, Base-stock, Fill rate.

1 Introduction

This study is to investigate structure of the optimal scheduling policy for multi-class base-stock controlled systems where demand that can not be satisfied upon its arrival is backordered and preemption is allowed. It is possible to consider the system outlined above as a manufacturing system or a repair shop. In case of a repair shop, demands represent failures of repairable items. When an item fails, it joins its own queue in the repair shop. One has to decide how many spares should be kept in the repair shop. Many spares bring extra cost, fewer spares, on the other hand, would cause frequent interruptions of the activities at the sites where the repairable items are in use as in the case of many military applications. The performance measure considered in this study is steady-state fill rate, defined as the long-run fraction of the time demand is satisfied immediately upon its arrival. Instead of working with individual fill rates of the items, a weighted average of them is introduced as the aggregate fill rate to represent the overall system performance.

Only the most related studies in the literature are reviewed next. Zheng and Zipkin [3] consider the same system as the one investigated in this study for two types of items with identical demand and service rates, and holding and backorder costs. Assuming that the longest queue (LQ) policy is applied preemptively, the authors analytically show that the performance of LQ policy is always better than the performance of first-come-first-served (FCFS) policy with respect to average cost criterion over infinite horizon when both holding and backorder costs are convex. Ha [1] studies two-item single-server make-to-stock production system in order to investigate structure of the optimal policy under the expected discounted holding and backorder costs over infinite horizon allowing parameters (Poisson demand rates, exponential service rates and cost figures) to be different for the items. He proves that a time-stationary priority rule (a hedging point policy) is optimal when there are backorders of both items. van Houtum et al. [2] approximate performance of the symmetric LQ system where symmetry means identical (Poisson) demand rates and (exponential) processing times for the items. The work by van Houtum et al. is directly comparable to our study since the performance measure they work with is fill rate.

Organization of this report is as follows: first three subsections of section 2 are devoted to a detailed analysis of symmetric two-item case, to be extended to the asymmetric case again with two items in section 2.4. An outline of ongoing research on generalized models is given in section 3.

2 Two-Class System

The two-class system studied in this section is depicted in Figure 1. The case of single-server facility is considered in the first place as a building block of more generalized models. Processing time for the items is exponentially distributed with mean $\frac{1}{\mu}$ independently of the item type. There are no set-up costs or change-over times. S_i is the base-stock level for the inventories of item i. Demands for the items of type i occur according to independent Poisson processes with rate λ_i, $i = 1, 2$, and are met from the respective stock if there is available inventory (i.e., $\overline{n}_i > 0$). Upon demand arrival, a new item is released (failed item is sent to the repair shop) to be processed and to replenish stock. Requests that can not be satisfied immediately are backordered. n_i denotes the number of items to be processed and k_i is the number of backordered requests, $i = 1, 2$. Base-stock policies imply the following inventory balance equations: $n_i + \overline{n}_i = S_i + k_i$, $i = 1, 2$, (note that $\overline{n}_i \cdot k_i = 0$). Hence, $\mathbf{n} = (n_1, n_2)$ fully describes state of the system under consideration.

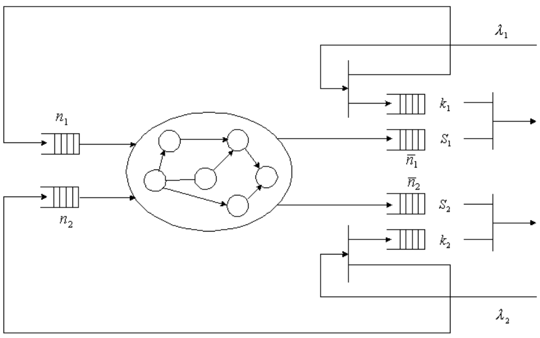


Figure 1: Two-class system.

Defining $f_m(\mathbf{n})$ as the minimum total cost over m periods when the initial state is $\mathbf{n}$, the recursive (multi-period) formulation given below is to determine the item type to be processed in state (n_1, n_2) when there are m periods to go.

$$\begin{aligned} f_m(n_1, n_2) &= c(n_1, n_2) + \frac{\lambda_1}{\tau} f_{m-1}(n_1 + 1, n_2) + \frac{\lambda_2}{\tau} f_{m-1}(n_1, n_2 + 1) \\ &\quad + \frac{\mu}{\tau} v_{m-1}(n_1, n_2), \\ f_0(n_1, n_2) &= 0, \end{aligned}$$

for all n_1, n_2, where $\tau = 2\lambda + \mu$ and $c(\mathbf{n}) = (\lambda_1 \cdot 1_{\{n_1 \geq S_1\}} + \lambda_2 \cdot 1_{\{n_2 \geq S_2\}})/(\lambda_1 + \lambda_2)$ and

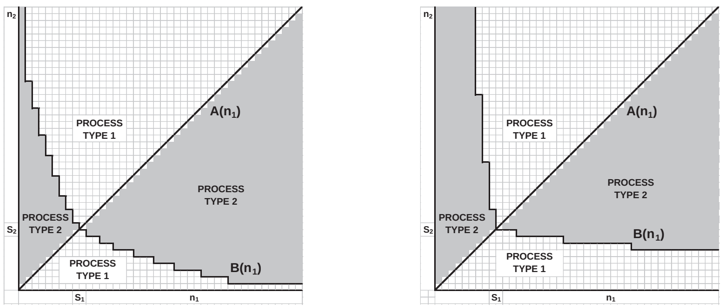
$$v_{m-1}(n_1, n_2) = \begin{cases} min\{f_{m-1}(n_1 - 1, n_2), f_{m-1}(n_1, n_2 - 1)\} & \text{if } n_2 > 0,\ n_1 > 0, \\ f_{m-1}(n_1 - 1, n_2) & \text{if } n_2 = 0, \\ f_{m-1}(n_1, n_2 - 1) & \text{if } n_1 = 0, \\ f_{m-1}(n_1, n_2) & \text{if } n_1 = 0,\ n_2 = 0. \end{cases}$$

Recall that preemption is allowed. Since it is always possible to redefine the time scale, without loss of generality, we assume that $\tau = 1$. The cost function, $c(n)$, is a weighted average of the stockout probabilities, it is introduced as an extension of the one in [2]. Minimizing the long-run average cost in the formulation above is, then, equivalent to maximizing weighted average of the fill rates. For the simple case studied in the following three subsections, the demand rates and base-stock levels are identical (letting $S = S_i$ for $i = 1, 2$).

2.1 Symmetric Case: Structure of the Optimal Policy

In order to solve the recursive formulation presented above, value-iteration algorithm is used. Based on the numerical experiments, structure of the optimal policy shows different characteristics over the following four regions of the state space: $n_1 < S_1$ and $n_2 < S_2$ (region I), $n_1 < S_1$ and $n_2 \geq S_2$ (region II), $n_1 \geq S_1$ and $n_2 \geq S_2$ (region III), $n_1 \geq S_1$ and $n_2 < S_2$ (region IV). LQ and Shortest Queue (SQ) policies are optimal in regions I and III, respectively. In the remaining regions II and IV, optimal policy is determined by a switching curve $B_m(n_1)$ that converges to $B(n_1)$ as m tends to infinity. $A_m(n_1)$ is the diagonal for all m, so is $A(n_1)$. See Figure 2 for two example cases (two different $\rho = (\lambda_1 + \lambda_2)/\mu$ values). This policy structure makes sense as clarified below.

- In region I, none of the items is in stockout. LQ policy is followed in order to avoid stockout for the item with higher risk of falling into the stockout region, i.e., the one with higher n_i.
- In region III, both of the items are in stockout. SQ policy is followed in order to eliminate stockout for the promising item to reach non-stockout region sooner, i.e., the one with smaller n_i.
- In regions II and IV, one of the items is in stockout. $B(n_1)$ is the threshold level to be away from region III and to reach region I.



(a) $\rho = 0.4$, $S = 9$. **(b)** $\rho = 0.9$, $S = 9$

Figure 2: Optimal scheduling policy.

2.2 Symmetric Case: Comparison of the Optimal Policy with LQ and FCFS Policies

In this section, performance of the optimal policy is compared with those of LQ and FCFS policies. To a certain extent, FCFS would also reflect the behavior of LQ policy because it is more probable that the next item to be processed would be of type i if $n_i > n_j$, $j \neq i$. As a matter of fact, LQ and FCFS policies perform almost equally well for a wide range of parameters, especially when they are compared to the optimal policy, as seen in Tables 1 and 2. Next, these two policies are compared to the optimal one for constant base-stock levels. LQ policy is optimal in region I. When ρ is small, probability of observing only a few items in the system would be high, i.e., the system would mostly be in region I. This explains why performance of the optimal policy is not strikingly dominant when ρ is small. On the other hand, for higher ρ values, visiting states outside region I is more probable making the difference between the optimal and the other policies apparent. The difference is due to handling states outside region I optimally, instead of persisting with LQ or FCFS.

In Table 1, the minimum base-stock levels required in order to achieve different target service levels (aggregate fill rates in the first column) can be seen under each of the optimal, LQ and

Table 1: Comparison of the Optimal, LQ and FCFS policies with minimum base-stock levels to satisfy target fill rates.

		OPTIMAL		FCFS			LQ		
FR	ρ	S	FR(%)	S	FR(%)	Error(%)	S	FR(%)	Error(%)
0.90	0.40	2	94.35	2	93.75	0.60	2	94.03	0.32
	0.60	3	93.64	3	92.13	1.51	3	92.68	0.96
	0.80	5	90.87	6	91.22	-0.35	6	91.91	-1.04
	0.90	9	90.04	12	91.00	-0.96	12	91.50	-1.46
	0.95	18	90.90	24	90.95	-0.05	23	90.28	0.62
0.95	0.40	3	98.73	3	98.44	0.29	3	98.68	0.05
	0.60	4	97.51	4	96.63	0.88	4	97.17	0.34
	0.80	7	96.22	8	96.10	0.12	8	96.67	-0.45
	0.90	13	95.71	15	95.07	0.64	15	95.48	0.23
	0.95	24	95.10	30	95.03	0.07	30	95.26	-0.16
0.99	0.40	4	99.72	4	99.61	0.11	4	99.72	0.00
	0.60	5	99.05	6	99.38	-0.33	6	99.59	-0.54
	0.80	10	99.01	12	99.23	-0.22	11	99.13	-0.12
	0.90	20	99.02	23	99.01	0.01	23	99.16	-0.14
	0.95	40	99.08	47	99.09	-0.01	46	99.08	0.00

FCFS policies. Errors are $FR_{OPT} - FR_{FCFS}$ and $FR_{OPT} - FR_{LQ}$. Under the optimal policies, the desired service levels are achieved with lower base-stock levels as compared to the other two policies. The decrease in investment for stock keeping units (sku) is more apparent for higher ρ values. This is explained by the increase in the probability of being in stockout (visiting the states in regions II, III and IV more frequently) when ρ is high. Recall that handling the stockout case in regions II, III and IV optimally puts forward the difference between the optimal and the other two policies. When $\rho = 0.9$ and $S = 8$, the difference between the fill rates under optimal and LQ policies is nearly 8%. This difference increases to 17% for $S = 1$, because smaller S values increase probability of stockout and cause the policies to be distinguishable (LQ and optimal policies are the same as long as the system is in region I). Comparisons of this sort can be made analyzing results in Table 2. Performances of the LQ and FCFS policies in terms of the required sku investment turn out to be almost the same (not only the example cases in Table 1, but in general for the extensive numerical experiments) in accordance with the intuitive comparison of LQ and FCFS policies to behave similarly on the average as pointed out in the preceding paragraph. Some error values are negative which means lower fill rate under the optimal policy, but note that these cases are observed when the optimal policy achieves the target service level with a smaller base-stock level.

2.3 Symmetric Case: Alternative Policies Generated by Heuristics

In this section, five heuristics are proposed to generate alternative policies. In devising these heuristics, the approach is to approximate $B(n_1)$ because a closed-form solution can not be derived for $B(n_1)$. (Note that $B(n_1)$ characterizes optimal policy together with $A(n_1)$, the latter of which turns out to be the diagonal $n_2 = n_1$ for the symmetric case). All alternative policies are of LQ and SQ types in regions I and III, respectively, as observed numerically for the optimal policy.

The reasoning behind heuristics 1 and 2 is given next. When the system is in state (n_1, n_2) in region II, with rate λ_2 the system moves to the states closer to the more costly region III; on the other hand, the system gets away from the costless region I with rate λ_1. In order to approximate $B(n_1)$, the choice for each state in region II would be between trying to get away from the more costly region III by processing type 2 items (while at the same time getting away from region I by arrivals of type 1 items) and trying to get closer to costless region I by processing type 1 items (while at the same time getting closer or even into region III by arrivals of type 2 items). Similar arguments could be raised also for the states in region IV.

Heuristic 1: $\frac{n_1-(S_1-1)}{\mu-\lambda_1}$ and $\frac{S_2-n_2}{\lambda_2}$ are the indices to be compared for state (n_1, n_2) in region II. The former can be interpreted as the expected number of steps (periods) to reach region I from state (n_1, n_2) (to cover distance $n_1 - (S_1 - 1)$) while type 1 is being processed. Similarly, interpretation of the latter is the expected number of steps (periods) to reach region III (to

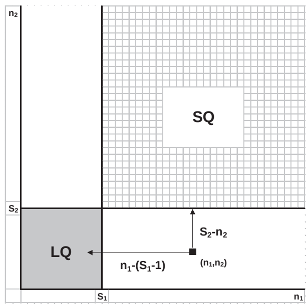
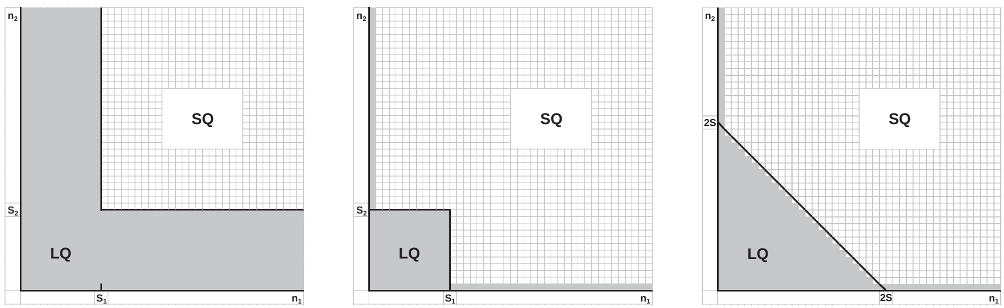


Figure 3: Distances from a state in region II to regions I and III.

consider distance $S_2 - n_2$) while type 1 is being processed. If the former is smaller, then we can process type 1 items because we expect to reach costless region I before stepping up the more costly region III. Otherwise, if $\frac{S_2-n_2}{\lambda_2}$ is smaller, then it is more probable that we will step up the more costly region III before reaching region I by processing type 1 items, so going away from the more costly region is preferable for the state under consideration. The distances to be covered are shown in Figure 3. Note that $\frac{S_1-n_1}{\lambda_1}$ and $\frac{n_2-(S_2-1)}{\mu-\lambda_2}$ are the indices to be compared for any state (n_1, n_2) in Region IV.

Heuristic 2: This heuristic is just a variation of heuristic 1. Indices are revised as $\frac{n_1-(S_1-1)}{\mu}$ and $\frac{S_2-n_2}{\lambda_2}$ for state (n_1, n_2) in Region II and as $\frac{S_1-n_1}{\lambda_1}$ and $\frac{n_2-(S_2-1)}{\mu}$ for state (n_1, n_2) in Region IV.



(a) Heuristic 3. **(b)** Heuristic 4. **(c)** Heuristic 5.

Figure 4: Heuristics.

Although heuristics 1 and 2 are easy to implement, use of time consuming value-iteration is unavoidable to compute steady-state performance measures (fill rate in this study). That is why three other heuristics are proposed as rather rough approximations with more regular structures so that we may either derive a closed-form steady-state distribution or at least devise a recursive scheme to calculate the steady-state probabilities easily. As a matter of fact, as noted in the first paragraph of this section, we come up with a recursive scheme for heuristic 3 but unfortunately not for heuristics 4 and 5.

Heuristic 3: SQ policy in region III and LQ policy in the remaining regions are followed as seen in Figure 4(a). The algorithm in [3] is used to find the steady-state probabilities in the region $\{(i, j)|\ i < S_1 - 1 \text{ or } j < S_2 - 1\} \cup (i = S_1 - 1,\ j = S_2 - 1)$ where LQ policy is followed, then the other steady-state equations can be solved recursively. This recursive scheme is skipped here to keep the presentation short.

Heuristic 4: LQ Policy in region I and SQ Policy in the remaining regions are followed as seen in Figure 4(b). Note that for the states where one of the queues is empty LQ policy is employed.

Heuristic 5: LQ policy in region $\{(n_1, n_2)|\ n_1 + n_2 < 2S\}$ and SQ policy in the remaining regions are followed as seen in Figure 4(c).

Note that heuristics are the same for the case $S = 1$; SQ when both n_1 and n_2 are positive and LQ when $n_1 = 0$ or $n_2 = 0$.

The observations that result from the numerical experiments (some example cases are given in Table 2) are as follows:

- Heuristics 1 and 2 perform very well since the actions chosen are the same as the optimal policy in regions I and III and in most of the states in regions II and IV. Heuristic 2 works better than heuristic 1 in almost all cases scanned.
- Heuristics 3 and 4 perform better (worse) than LQ and FCFS (heuristics 1 and 2) policies because the actions chosen are the same as (quite different from) the optimal policy only in regions I and III (II and IV).

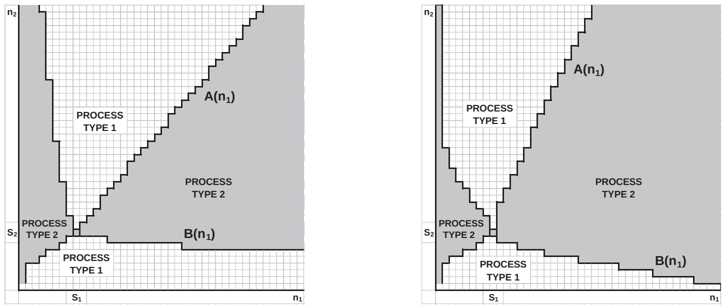
Table 2: Comparison of the performance (fill rate, %) of the Optimal, LQ, FCFS policies and heuristics.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.2	Optimal	89.07	98.85	99.88	99.99	100.00	100.00	100.00	100.00
	LQ	88.79	98.84	99.88	99.99	100.00	100.00	100.00	100.00
	FCFS	88.89	98.77	99.86	99.98	100.00	100.00	100.00	100.00
	Heuristic 1	89.07	98.85	99.88	99.99	100.00	100.00	100.00	100.00
	Heuristic 2	89.07	98.85	99.88	99.99	100.00	100.00	100.00	100.00
	Heuristic 3	89.07	98.85	99.88	99.99	100.00	100.00	100.00	100.00
	Heuristic 4	89.07	98.64	99.84	99.98	100.00	100.00	100.00	100.00
	Heuristic 5	89.07	98.82	99.88	99.99	100.00	100.00	100.00	100.00
0.6	Optimal	61.82	84.21	93.64	97.51	99.64	99.95	100.00	100.00
	LQ	55.32	81.55	92.68	97.17	99.59	99.94	100.00	100.00
	FCFS	57.14	81.63	92.13	96.63	99.38	99.89	99.99	100.00
	Heuristic 1	61.82	84.21	93.63	97.50	99.64	99.95	100.00	100.00
	Heuristic 2	61.82	84.15	93.62	97.50	99.64	99.95	100.00	100.00
	Heuristic 3	61.82	83.83	93.49	97.46	99.63	99.95	100.00	100.00
	Heuristic 4	61.82	82.61	91.58	95.92	99.11	99.83	99.99	100.00
	Heuristic 5	61.82	83.95	93.39	97.36	99.61	99.95	100.00	100.00
0.9	Optimal	35.46	53.20	63.86	71.24	81.26	87.71	93.47	97.19
	LQ	16.23	31.12	43.80	54.31	69.93	80.26	89.50	95.48
	FCFS	18.18	33.06	45.23	55.19	70.00	79.92	89.00	95.07
	Heuristic 1	35.46	53.20	63.79	71.08	80.97	87.37	93.17	97.01
	Heuristic 2	35.46	53.17	63.74	71.13	81.19	87.66	93.44	97.17
	Heuristic 3	35.46	46.65	56.36	64.48	76.61	84.64	91.83	96.48
	Heuristic 4	35.46	52.48	62.11	68.50	77.24	83.38	89.74	94.72
	Heuristic 5	35.46	53.20	63.79	71.08	80.90	87.25	93.05	96.92

2.4 Asymmetric Case

In this section, demand rates are allowed to be different for different item types. Structure of the optimal policy is investigated for two different cost functions: the weighted average introduced in section 2 and $c(\mathbf{n}) = (1_{\{n_1 \geq S_1\}} + 1_{\{n_2 \geq S_2\}})/2$. Note that the two cost functions are the same for the symmetric setting. In the latter, both item types are equally important. On the other hand, in the former, immediate delivery for the item type with higher demand rate is more important. In Figure 5, structure of the optimal policy is seen for two example cases.

Generating alternative policies for the asymmetric case is more complex than for the symmetric case because, for the symmetric setting, it is enough to approximate $B(n_1)$ only, $A(n_1)$ is the diagonal. However, $A(n_1)$ also needs to be approximated for the asymmetric case. To that end, i.e., to approximate $A(n_1)$, the idea used for heuristics 1 and 2 for the symmetric case, which is based on comparing the time required to cover distances of a state in regions II or IV to regions I and III, is extended to the remaining regions. Indices with a revision of heuristic 1 in section 2.3 are given in Table 3 for the equally weighted cost function. Revision of heuristic 2 is then obtained by using μ instead of $\mu - \lambda$ for calculating the first index in region II and the second one in region IV. In order to adjust heuristics for the weighted average cost function,



(a) $\rho = 0.4$, $S = 8$. **(b)** $\rho = 0.9$, $S = 8$.

Figure 5: Optimal scheduling policy: equally weighted cost function. ($\lambda_1 = 2\lambda_2$)

first and second indices in each region are multiplied by λ_1/λ and λ_2/λ, respectively. Note that, for equal demand rates, these heuristics turn into heuristics 1 and 2 in section 2.3.

Table 3: Indices for heuristic 1: equally weighted cost function.

Region IV	Region III
$\frac{S_1-n_1}{\lambda_1}, \frac{n_2-(S_2-1)}{\mu-\lambda_2}$	$\frac{n_1-(S_1-1)}{\mu-\lambda_1}, \frac{n_2-(S_2-1)}{\mu-\lambda_2}$
Region I	Region II
$\frac{S_1-n_1}{\lambda_1}, \frac{S_2-n_2}{\lambda_2}$	$\frac{n_1-(S_1-1)}{\mu-\lambda_1}, \frac{S_2-n_2}{\lambda_2}$

The heuristics are compared with the optimal and FCFS policies. For the equally weighted fill rate case, there is also a heuristic introduced in [3] and used here for comparisons: serve a customer of type 2 (1) when $n_2 - n_1 > \Delta$ ($n_1 - n_2 > \Delta$), for some predetermined constant $\Delta > 0$ when $\lambda_2 > \lambda_1$ ($\lambda_1 > \lambda_2$). The reader is referred to Table 4 for comparisons of some example cases.

Table 4: Comparison of the performance (fill rate, %) of the Optimal, FCFS policies and heuristics: equally weighted cost function.

ρ		$S=1$	$S=2$	$S=3$	$S=4$	$S=6$	$S=8$	$S=11$	$S=15$
0.2	Optimal	89.20	98.78	99.86	99.98	100.00	100.00	100.00	100.00
	FCFS	89.01	98.68	99.83	99.98	100.00	100.00	100.00	100.00
	Heuristic 1	89.20	98.78	99.85	99.98	100.00	100.00	100.00	100.00
	Heuristic 2	89.20	98.78	99.85	99.98	100.00	100.00	100.00	100.00
	$\Delta=1$	88.93	98.83	99.86	99.98	100.00	100.00	100.00	100.00
	$\Delta=2$	88.98	98.70	99.85	99.98	100.00	100.00	100.00	100.00
	$\Delta=3$	88.99	98.70	99.84	99.98	100.00	100.00	100.00	100.00
	$\Delta=5$	88.99	98.69	99.84	99.98	100.00	100.00	100.00	100.00
	$\Delta=8$	88.99	98.69	99.84	99.98	100.00	100.00	100.00	100.00
0.6	Optimal	63.42	84.56	93.59	97.38	99.58	99.94	100.00	100.00
	FCFS	58.33	81.94	91.90	96.26	99.15	99.80	99.98	100.00
	Heuristic 1	63.39	84.44	93.45	97.32	99.57	99.93	100.00	100.00
	Heuristic 2	63.39	84.51	93.49	97.34	99.57	99.93	100.00	100.00
	$\Delta=1$	56.48	83.21	93.02	97.17	99.55	99.93	100.00	100.00
	$\Delta=2$	57.72	82.01	93.05	97.12	99.54	99.93	100.00	100.00
	$\Delta=3$	58.33	82.00	92.15	96.95	99.49	99.92	100.00	100.00
	$\Delta=5$	58.80	82.03	91.88	96.27	99.32	99.88	99.99	100.00
	$\Delta=8$	58.96	82.05	91.81	96.13	99.07	99.77	99.98	100.00
0.9	Optimal	41.12	56.93	65.78	72.25	81.72	88.00	93.64	97.27
	FCFS	19.64	35.14	47.42	57.19	71.27	80.43	88.71	94.38
	Heuristic 1	41.07	56.76	65.31	71.54	80.88	87.22	93.07	96.96
	Heuristic 2	41.07	56.45	65.20	71.78	81.40	87.78	93.51	97.21
	$\Delta=1$	17.07	33.30	45.42	55.57	70.71	80.74	89.75	95.58
	$\Delta=2$	18.48	33.66	46.65	56.44	71.29	81.15	89.98	95.69
	$\Delta=3$	19.64	34.99	46.93	57.27	71.65	81.38	90.11	95.75
	$\Delta=5$	21.48	37.21	49.01	58.21	72.13	81.36	90.05	95.72
	$\Delta=8$	19.64	35.14	47.42	57.19	71.27	80.43	88.71	94.38

3 Extensions

An immediate generalization of the proposed heuristics is for multi-class setting. Another is to incorporate the (exponential) set-up times for switching from processing one type to another. In this case, the recursive formulation in section 2 is revised to have set-up information in the state description in addition to $\mathbf{n}$. Then, the revision of the heuristics is not as immediate as in the case of extension to multi-class systems. Preliminary numerical experiments for these two generalizations show that value-iteration can not be employed due to computational restrictions. We are currently working on a simulation model to compare performances of alternative policies. This would not allow us to figure out structure of the optimal policy, but to compute some other relevant performance measures.

A flaw of the setting we have in this study could be considered as the missing trade-off between inventory holding and backorder costs. Backordering is handled by maximizing fill rate, but this means only the fraction of time with stockout is penalized regardless of the number of backorders. As pointed out in section 2.2, use of optimal policy naturally comes up with a striking advantage in terms of the required number of skus to achieve target fill rates as compared to the use of LQ and FCFS policies. But, in fact, sku investment is not included in our recursive optimization model. In this respect, analysis to resolve the trade-off between inventory holding cost (or sku investment) and fill rate appears as a further research direction.

References

[1] HA, A. 1997. Optimal dynamic scheduling policy for a make-to-stock production system. *Oper. Res.* **45** 42-53.

[2] HOUTUM, G.J. VAN, ADAN, I.J.B.F AND WAL, J. VAN DER. 1997 The symmetric longest queue system. *Communications in Statistics: Stochastic Models* **13** 105-120.

[3] ZHENG, Y.S. AND ZIPKIN, P. 1990. A queueing model to analyze the value of centralized inventory information. *Oper. Res.* **38** 296-307.