Control Limits for Multi-Stage Manufacturing Processes with Binomial Yield (Single and Multiple Production Runs)

Miryam Barad; Daniel Braha

Control Limits for Multi-stage Manufacturing Processes with Binomial Yield (Single and Multiple Production Runs)

MIRYAM BARAD and DANIEL BRAHA

Tel-Aviv University, Israel

The problem of ordering a single or a series of production runs to meet a single make-to-order demand, where the various stages in the production process have binomial yields, is addressed. Use of additional procurement of part-finished products or reworking defective material previously made to supplement yields at any stage is considered. The costs of this and manufacture, disposal of surplus or shortages are assumed to be proportional to the numbers involved. It is shown that when a single production run is considered, the optimal policy is defined by two critical numbers (control limits) at each stage. The treatment of ordering a series of production runs follows from and builds on the single run analytical result by developing an approximation in which the problem is decomposed into a series of single runs. Set-up costs may or may not be incurred.

Key words: lot sizing, multi-stage processes, stochastic yield

INTRODUCTION

Stochastic yield in response to a deterministic make-to-order (single period) demand for non-defective units is a situation common to many industries, such as those with electronic or mechanical products, and difficulties in the complex area of microelectronics assembly have renewed interest in research into this problem. The implications of including yield randomness into production/inventory models have been addressed by Karlin[1], Shih[2], Lee and Yano[3] and Wein[4], amongst others (see Yano and Lee[5] for a survey of lot sizing with random yields). Three aspects of the modelling problem are noted as relevant here.

(1) The first aspect is the choice of the probability law used to describe the stochastic yield of the manufacturing process. Two major alternative sources are discrete distributions (usually with independence between the units produced being assumed) or continuous ones. After consideration, theoretical versus empirical could be the relevant classification to adopt.

(2) The second aspect of model design is the complexity of the manufacturing system being represented, whether it can be treated as single- or multi-stage.

(3) The number of runs that may be carried out in order to satisfy the given demand is the third relevant aspect of model design. If this is limited (to a single run, perhaps) then a short (or excess) quantity of good finished units may occur, and the resulting shortage (or overage) costs have to be tolerated. If additional runs may be made within the delivery time allowed, the total processing time for each run being relatively low, then these may be made to make up shortages, although for such a decision to be economic, the costs of the additional set-ups (e.g. of reordering parts and materials) have to be considered.

Initially, the problem was treated as that of 'Reject Allowance' (see Giffler[6], Levitan[7] for example) with the yield of a (single stage) system being assumed to have either a binomial or Poisson distribution. Single-stage models were also studied by, amongst others, Karlin[1], Shih[2], Erhard and Taube[8], Henig and Gershak[9]. Shih considered a continuous version of the 'Reject Allowance' whilst assuming that the yield was stochastically proportional to the input. Such a yield model has also been examined by Henig and Gershak[9] and Gershak et al.[10] amongst

---

Correspondence: M. Barad, Department of Industrial Engineering, Faculty of Engineering, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel

others. Klein[11] represented the 'Reject Allowance' single stage problem with N production runs as a Markovian process. Beja[12] extended the binomial formulation of the model by using an assumption that he called 'constant marginal production efficiency'.

In the models mentioned above, yield distributions were 'given'. Situations where appropriate investments could improve such distributions were modelled, for example, by Gerchak and Parlar[13] and by Cheng[14]. An early approach to the serial multi-stage (single period) problem, was addressed by Vachani[15] who also considered more than one production run to satisfy the given demand.

A recent treatment of the multi-stage, single period model, which assumed a stochastically proportional distribution of the yield as mentioned above, was presented by Lee and Yano[3]. In their model (set in the microelectronics area) a decision is made after each manufacturing stage (an inspection station is located after each stage) to determine how many non-defective semi-finished units should be transferred to the next stage and how many should eventually be disposed of. They assume that production, disposal and shortage costs are linear and that it is less expensive to hold inventory at one stage than to process it and hold the expected non-defective output at the following stage. Under these circumstances, the authors proved the optimality of a single critical number representing the optimal target input quantity at each stage. A multi-stage serial production system, where buffer stocks have target values and their expected deviations from these target values are minimized, was analysed in Reference 16. Wein[4] extended Lee and Yano's model[3] to allow (at given cost) rework of the defective items at some production stages. Her model was also set in the world of microelectronics, namely in the Application Specific Integrated Circuit (ASIC). It assumed that rework could always be carried out and that shortages were not allowed. Empirical models of yield distributions were reported.

Here, an extension of Lee and Yano's formulation is presented. First, similarly to their model, a single production run is considered. Then, a multiple production run formulation of the problem, building on the single run analytical result through a decomposition into a series of single-run models, is also addressed. The single-run model differs from Lee and Yano's formulation in several aspects. It assumes a binomial rather than a stochastically proportional (multiplicative) yield model. A significant contrast between these two models is the independence of the individual yields exhibited by the binomial model, versus perfect correlation as implied by the multiplicative model. Accordingly, in the binomial model, the variance of the yield rate (fraction non-defective) diminishes with the increased batch size, while in the proportional model it is not affected by the batch size.

Another distinctive feature of the present formulation is its allowing for purchasing of semi-finished products, or reworking some of the defectives, at intermediate stages. It is shown that the optimal policy here has two critical numbers at each stage. The second critical number is essentially a 'procure-up-to' point for non-defective semi-finished units.

A 'purchasing' situation fits a company that applies a modular product design policy coupled with a Just-In-Time (make-to-order) inventory policy. The manufacture of pneumatic power tools[17] is an example of modular product design where a broad array of models (built through extensive use of standardized component modules) are made to order and where 'casings', 'mechanisms' and some modular components can be obtained from outside suppliers. Suppliers and customers work closely to resolve issues pertaining to design and delivery. In this environment, different final products share some common stages in the course of their manufacturing processes (see also A3 in the next section). Manufacturing stages, common to some specific models, are more likely to be early stages that are fundamental to many products.

A 'reworking' situation is compatible with an environment such as ASIC production (as captured by Wein's model) where the profitability of rework alternatives has to be assessed. In this context it should be recalled that 'special' rework operations (usually conducted externally to the main production line), although sometimes very expensive as compared with 'regular' (on-line) manufacturing costs, may still not be successful. Hence, valid estimates of rework cost have to take into account that some rework operations may have to be repeated and that re-inspection and testing may be necessary. To be compatible with 'purchasing', the

rework price here should include the production costs already invested in the unit to be reworked.

Timeliness of procurement alternatives is another important factor in this decision. Waiting for externally purchased units to arrive or for rework to be completed will increase the lead time of the final product. It also seems important to note that, although the probability of their occurrence can be substantially reduced, when yield is stochastic, time is limited and rework is not possible, shortages cannot be avoided. Given the estimates of the purchasing (or rework) cost per unit at each stage, the profitability of these activities (at each stage) is affected by the number of additional non-defective units to be procured (if at all), calculated as the difference between the 'procure up to' point and the yield of the previous stage. Initial intermediate inventories are also considered.

The rest of the paper is organized as follows. The current model is described in the next section, followed by its detailed solution, which also contains general properties of the critical numbers (control limits), specific decisions as dictated by special procurement circumstances (unavailability, zero cost, limited quantity) as well as yield modelling and estimation. Numerical examples are then presented, being followed by the case where it is possible to carry out additional production runs. The paper ends with some concluding remarks and suggestions for future work.

MODEL FORMULATION (SINGLE PRODUCTION RUN)

As mentioned above, a discrete, N-stage system with a given demand, D, for non-defective finished units is considered. It is assumed that a single production run is to be carried out and that, at its end, a linear penalty cost for shortage, or a linear overage cost, may be incurred. The non-defective units at any stage i, $i = 1, 2, \ldots, N$, are independently generated with equal probability P_i (a stage dependent parameter). Hence, the number of non-defective units produced at stage i, obeys a binomial distribution with parameters P_i, $0 < P_i < 1$ and U_i, the input batch size at the same stage. After performing perfectly reliable, 100% inspection of the production output from the previous stage and discarding all the defective units, a decision has to be made regarding the number of units to be processed at the next stage. Our decision formulation allows for three alternatives.

(A1) Processing all the non-defective units available from the previous stage.

(A2) Reducing, with per unit cost, the available input by disposal of some units.

(A3) Increasing, with per unit cost, the available input by purchasing additional non-defective semi-finished units, or by reworking some of the defective units.

Units may be purchased from other companies, or may be leftovers (meant for other final products) which became superfluous as a result of applying A2. (With a view to reducing inventory space, eventually only leftover units common to many final products would be stored.) If leftover units are used, the costs invested in their disposal can be considered sunk costs and, accordingly, their assigned procurement cost is zero (see also 'Remarks').

The following notation will be used:

D = the known demand;

i = stage index chosen to represent the number of remaining stages until the end of the manufacturing process, $i = 1, 2, \ldots, N$;

y_i = the available input to stage i, which is equal to the yield (non-defective units) of the previous stage (stage $i + 1$ according to the current notation);

U_i = the decision variable at stage i, representing the input to the ith stage;

p_i = the probability of producing a non-defective unit at stage i;

$P_i(x|U_i)$ = the probability of a yield x arising from an input to the ith stage of U_i;

w_i = the manufacturing costs per unit (including inspection) at stage i;

h_i = the disposal cost per available non-defective unit at stage i;

r_i = the procurement cost per non-defective semi-finished unit at stage i;

h_o = the overage cost incurred by producing one finished unit above the demand;

π = the shortage penalty per unit of unsatisfied demand.

Two interrelated functions are defined.

$C_i(y_i)$ = the minimal total expected cost given that the available input to stage i is y_i and an optimal policy is carried out from stage i throughout the remaining stages.

$F_i(U_i)$ = the minimal total expected cost, given that the input to the ith stage is U_i and an optimal policy is carried out in the subsequent stages.

The connection between these functions based on our decision-modelling is expressed by the following equation.

For every $1 \leqslant i \leqslant N$,

$$C_i(y_i) = \min\left\{F_i(y_i), \min_{j<y_i}[F_i(y_i - j) + jh_i], \min_{k>0}[F_i(y_i + k) + kr_i]\right\} \tag{1}$$

where

$$F_i(U_i) = w_i U_i + \sum_{x=0}^{U_i} C_{i-1}(x)P_i(x|U_i) \tag{}$$

and

$$F_1(U_1) = w_1 U_1 + \pi \sum_{x<D}(D - x)P_1(x|U_1) + h_o \sum_{x>D}(x - D)P_1(x|U_1) \tag{2}$$

A basic condition, C1 is assumed (an analogous assumption was presented by Lee and Yano[10]), namely that it is less expensive to remove an acceptable unit before a given production stage i, than to invest in its production and remove it at the next stage, $i - 1$:

$$\text{C1: } h_i < w_i + p_i h_{i-1} \quad \text{for} \quad i = 1, 2, \ldots, N.$$

Our objective is to determine $C_N(y_N)$ as well as the optimal policy at any stage i, $i = 1, 2, \ldots, N$. The procedure is depicted in Figure 1.
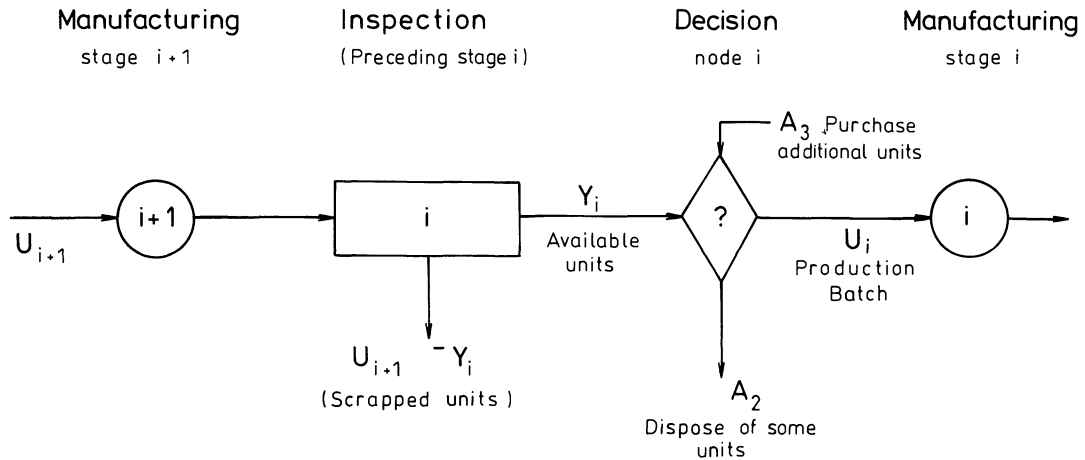

FIG. 1. *The general topology.*

THE SOLUTION APPROACH

The structure of the recurrence relations defined above suggests a dynamic programming formulation. First, the case of one more stage left to the end of the manufacturing process is considered.

The single stage formulation

$F_1(U_1)$ has an essential role in defining $C_1(U_1)$, since the optimal policy relies heavily on the discrete convexity of $F_1(U_1)$.

Definition 1

Discrete convexity — Assume a function $F(U)$ to be defined only on the non-negative integers. $F(U)$ is a discrete convex function if and only if $\Delta F(U)$ is a monotone increasing function of U, $\Delta F(U)$ being determined as: $\Delta F(U) = F(U + 1) - F(U)$.

Let us now examine some properties of $F_1(U_1)$.

Proposition 1

$F_1(U_1)$ is a 'discrete convex' function possessing a unique minimum value at $U_1 = U_1^*$.

Proof

For the proof see Appendix A.

The structure of the optimal policy

Proposition 2

Given y_1, the available input to stage 1, critical numbers L_1, M_1 satisfying $L_1 \leq U_1^* \leq M_1$ exist, defining the structure of the optimal policy (the input quantity U_1^0) as follows:

$$U_1^0 = \begin{cases} L_1 & \text{if } y_1 \leq L_1 \\ y_1 & \text{if } L_1 < y_1 < M_1 \\ M_1 & \text{if } y_1 \geq M_1 \end{cases} \qquad (3)$$

Proof

We divide the proof into two cases.

Case 1: $y_1 < U_1^*$

The discrete convexity of $F_1(U_1)$ implies that it is better to produce more than y_1. Hence, an additional amount, k, of the non-defective semi-processed units should be procured to supplement the currently available input y_1. The optimal procurement is determined by referring to the last term in (1). By substituting $y_1 + k$ with U_1, the following minimization problem is obtained:

$$\min_{U_1 \geq y_1} \{F_1(U_1) + r_1(U_1 - y_1)\} \qquad (4)$$

Let us define: $G_1(U_1) = F_1(U_1) + r_1(U_1 - y_1)$

Based on Proposition 1 regarding $F_1(U_1)$, it is easily proven that $G_1(U_1)$ too is a discrete convex function, having a unique minimum value denoted by L_1, with L_1 being the lowest U_1 satisfying $\Delta G_1(U_1) \geq 0$. To prove that $L_1 \leq U_1^*$, we first note that $\Delta F_1(U_1^*) > 0$ also implies $\Delta G_1(U_1^*) > 0$. The monotonicity of $\Delta G_1(U_1)$ completes the proof. Thus, the solution of (4) is:

$$U_1^0 = \begin{cases} L_1 & \text{if } y_1 \leq L_1 \\ y_1 & \text{if } L_1 < y_1 < U_1. \end{cases}$$

Case 2: $y_1 > U_1^*$

In this case it is more profitable to remove some of the above available input. Following an argument similar to the one used in case 1, the following minimization problem is obtained:

$$\min_{U_1 \leqslant y_1} \{F_1(U_1) + h_1(y_1 - U_1)\}. \tag{5}$$

Let us define: $H_1(U_1) = F_1(U_1) + h_1(y_1 - U_1')$.

As in Case 1, we conclude that $H_1(U_1)$ is a discrete convex function. Condition C1 guarantees that $H_1(U_1)$ attains its minimum at $U_1 = M_1$, M_1 being the lowest U_1 satisfying $\Delta H_1(U_1) \geqslant 0$.

To prove that $U_1^* \leqslant M_1$, let us assume that the contrary is true, i.e. that $U_1^* > M_1$. It is easily seen that in this case $\Delta H_1(M_1) < 0$, thus contradicting the condition $\Delta H_1(M_1) \geqslant 0$. Hence the optimal policy is:

$$U_1^0 = \begin{cases} y_1 & \text{if } U_1^* < y_1 < M_1 \\ M_1 & \text{if } y_1 \geqslant M_1. \end{cases}$$

This completes the proof of (3).

Corollary 1

(a) U_1^* is the lowest U_1 for which $\Delta F_1(U_1) \geqslant 0$.
(b) L_1 is the lowest U_1 for which $\Delta F_1(U_1) \geqslant -r_1$.
(b) M_1 is the lowest U_1 for which $\Delta F_1(U_1) \geqslant h_1$.

The N-stage formulation

Thus far, we have examined the case of one remaining stage to the end of the production process. By backward induction the N-stage problem will be solved.

Theorem 1

(1) $F_i(U_i)$ is a discrete convex function possessing a unique minimum value at $U_i = U_i^*$.
(2) Every stage possesses critical numbers L_i, M_i satisfying the following properties.
    (2.1) $L_i \leqslant U_i^* \leqslant M_i$.
    (2.2) U_i^* is the lowest U_i for which $\Delta F_i(U_i) \geqslant 0$.
    (2.3) L_i is the lowest U_i for which $\Delta F_i(U_i) \geqslant -r_i$.
    (2.4) M_i is the lowest U_i for which $\Delta F_i(U_i) \geqslant h_i$.
(3) The optimal input, U_i^0, is determined by the optimal policy:

$$U_i^0 = \begin{cases} L_i & \text{for } y_i \leqslant L_i \\ y_i & \text{for } L_i < y_i < M_i \\ M_i & \text{for } y_i \geqslant M_i. \end{cases} \tag{6}$$

The proof of this theorem follows similar arguments to those of the single stage case and is detailed in Appendix B where the discrete convexity of $F_i(U_i)$ is established.

The minimal total expected cost, $C_i(y_i)$ as a function of the available input at stage i, y_i, is portrayed in Figure 2.

Sufficient conditions for $L_i > 0$, $U_i > 0$ and $M_i > 0$ for $i = 1, 2, \ldots, N$ based on $\Delta F_i(0)$ are formulated in the following theorem.

Theorem 2

(1) $L_i > 0$ iff $\Delta F_i(0) + r_i < 0$.
(2) $U_i > 0$ iff $\Delta F_i(0) < 0$.
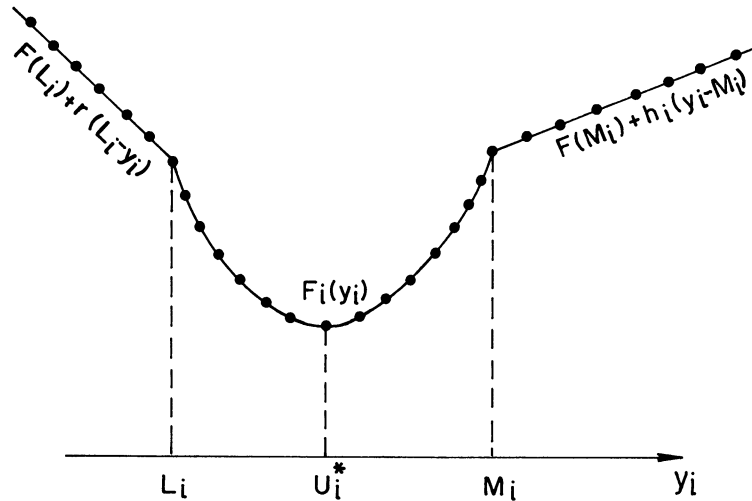(3) $M_i > 0$ iff $\Delta F_i(0) - h_i < 0$.

FIG. 2. *The expected total cost over the remaining stages as a function of the available input y_i.*

The recurrence equations that determine $\Delta F_i(0)$ and some special cases appear in Appendix C.

SOME REMARKS

1. Order relations among the critical numbers (control limits)

(a) At any given stage i, the value assumed by an upper control limit is no lower than that of the respective lower control limit.

(b) The lower control limits increase with the number of remaining stages until the end of the manufacturing process, i.e. $L_i \leqslant L_{i+1}$ for $i = 1, 2, \ldots, N - 1$, provided it is less costly to procure a non-defective unit at stage $i + 1$ and invest in its production at $i + 1$ than to procure it at the next stage i. This condition is expressed by inequality C2 and is presented below in terms of a lower threshold for r_i. (It also appears in Appendix C as special case 1 of Theorem 2.)

C2: $r_i > (r_{i+1} + w_{i+1})/p_{i+1}$ for $i = 1, 2, \ldots, N - 1$.

r_N denotes the raw material cost per unit.

The actual procuring price may be much higher, especially for a low i (a late manufacturing stage) where it is less likely to find other final products sharing common semi-finished parts. As the product nears completion, the economics of rework (if possible) tend to become increasingly attractive[18].

(c) In Theorem 1, the upper control limits exhibit the same relationships, $M_i \leqslant M_{i+1}$ for $i = 1, 2, \ldots, N - 1$. (See also special cases 1 and 2 in Appendix C.)

2. Degenerate control limits

This situation may arise when the lower control limit only, or when both control limits, assume a zero value for some stage.

(a) The implication of such an outcome concerning solely the lower control limit, L_i, is that the optimal policy discourages procurement at stage i. In Theorem 2, this occurs mainly if at the next stage, $i - 1$, the procuring cost r_{i-1}, for $i > 1$ is below the threshold C2 (see

Appendix C for recurrence equations of specific cases and also numerical examples in the next section). Unavailability of a specific semi-finished product may be modelled by arbitrarily assigning a high value to r_i, $(r_i \to \infty)$. Although originating from a somewhat different situation, the outcome will be the same, namely, a zero valued lower control limit at stage i.

(b) A different type of 'degeneration' arises when, at some stage i, both the lower control limit, L_i, and the input to the ith stage, U_i, (or eventually the upper control limit M_i as well) assume a zero value, $L_i = U_i = 0$ (or $L_i = U_i = M_i = 0$). In Theorem 2, this happens when it is cheaper to procure externally a non-defective unit at a later stage j $(j > i)$ than it is to produce it at stage i. Under these conditions the producer should revise the decision 'produce' versus 'supply' (see Appendix C and numerical examples).

3. Zero procurement cost of new material and semi-finished products

Suppose the available supply of raw material (y_N) is abundant and free. In this case the raw material cost per unit, r_N, assumes a zero value. Hence, by applying Theorem 1, it is easily seen that the lower and the higher critical numbers at the first manufacturing stage $i = N$ (i.e. N remaining stages) are equal to the optimal batch size, i.e. $L_N = M_N = U_N^0 = U_N^*$.

The lower critical number is equal to the optimal batch size $(L_i = U_i^*)$ at any other stage i, $i < N$, provided the cost of the semi-finished units is zero and their amount is not limited.

4. Limited supply of semi-finished procured units

(a) The general model was extended in Reference 19 to account for a limited quantity of semi-finished units that can be procured from an outside source. Briefly stated, let v_i designate a limited supply at stage i. Accordingly, the amount to be supplemented at stage i (for $L_i > y_i$), will be: $\min \{L_i - y_i, v_i\}$.

(b) Initial intermediate inventories (limited supply with zero procurement cost). An initial inventory of non-defective units, S_i, at stage i, $i = 1, 2, \ldots, N$ is a realization of zero cost semi-finished units with given limited supply. It is shown[19] that the structure of the optimal policy is now stated as follows:

$$U_i^0 = \begin{cases} L_i & \text{if } y_i + S_i < L_i \\ y_i + S_i & \text{if } L_i \leq y_i + S_i < U_i^* \\ U_i^* & \text{if } y_i \leq U_i^* \text{ and } y_i + S_i \geq U_i^* \\ y_i & \text{if } U_i^* \leq y_i \leq M_i \\ M_i & \text{if } y_i > M_i. \end{cases}$$

5. Yield modelling and estimation

Binomial models as such, and simple Poisson distributions that may be considered approximations of binomial models, play an important role in modelling yield. In the much-researched area of semiconductors, compound Poisson distributions are used as discrete models to describe clustering of defects, while Gamma distributions, representing continuous models, are successfully employed to the same end[20]. The advantage of the Poisson model is in its sound theoretical basis for inferring the defect density of existing products and for estimating the new product yield[21]. As detailed there, in practice the yield is measured as the average fraction of devices on a wafer that pass all tests, per area A_0, denoted $Y_{obs}(A_0)$. From this 'observed' yield a defect density is inferred, $D_{inf}(A_0)$, based on the assumed form of the area orientated distribution of defects, $D_{inf}(A_0) = f^{-1}[Y_{obs}(A_0)]$, where $f^{-1}[.]$ is the inverse function of that relating yield to defect density. This result is further used to 'estimate' the new product yield per area A, $D_{est}(A)$: $D_{est}(A) = \sigma D_{inf}(A_0)$, where σ is the magnitude of the scale factor. See also Cunningham[22] for an estimation and prediction of yield parameters.

Instead of considering an assumed theoretical distribution, empirical yield distributions are sometimes directly applied. It is noted that such yield distributions may represent mixtures of 'good' and 'bad' lots. While such mixed yield distributions are adequate for aggregated cost estimates, they may not be so for optimal decision making. To that end, it may be more advantageous to separate the yield distributions of 'good' and 'bad' lots since each may require a different optimal policy.

SOME NUMERICAL EXAMPLES

Let $D = 40$, $N = 4$, $h_o = 20$, $S_i = 0$ for $i = 1, 2, 3, 4$ and the other stage parameters be as follows.

Stage i	w_i	h_i	P_i
1	2	2	0.8
2	2	2	0.8
3	6	2	0.8
4	6	2	0.8

Three sets of procurement costs, r_i values, and two shortage penalty costs, $\pi = 52$ and $\pi = 100$ for each set are considered. The first set of procurement costs obey the conditions ensuring $L_i \leq L_{i+1}$ for $i = 1, 2, 3$ while the second and the third sets do not. The Normal Approximation to the binomial distribution is used to calculate the optimal parameters at each stage. The following results are obtained.

Set 1		$\pi = 52$			$\pi = 100$		
Stage i	r_i	L_i	U_i	M_i	L_i	U_i	M_i
1	27	47	52	52	50	53	53
2	19	54	66	69	57	69	71
3	9	64	77	79	67	81	83
4	1	79	**85**	90	83	**88**	94
Operating costs:		1364.13			1435.32		

Set 2		$\pi = 52$			$\pi = 100$		
Stage i	r_i	L_i	U_i	M_i	L_i	U_i	M_i
1	50	0	52	52	48	53	53
2	32	0	66	69	58	69	71
3	25	0	78	80	0	83	84
4	1	90	**91**	94	97	**98**	100
Operating costs:		1390.76			1485.74		

Set 3		$\pi = 52$			$\pi = 100$		
Stage i	r_i	L_i	U_i	M_i	L_i	U_i	M_i
1	50	0	52	52	48	53	53
2	13	60	66	69	64	69	71
3	9	0	77	80	0	81	83
4	1	0	**0**	0	0	**0**	0
Operating costs:		1136.53			1207.24		

The operating costs as presented do not consider the raw material costs. It is assumed that the manufacturing process starts with an optimal number of units entering stage 4.

Set 1. It is seen that both control limits increase with the number of stages remaining until the process ends. The operating costs are higher for the higher shortage penalty case and so are the respective control limits (lower and upper) at any given stage, as compared with those obtained for the lower shortage cost at the same stage. The widths of the control limits are about the same for the two penalty levels and do not seem to be affected by the stage number.

Set 2. Through the recurrence equations supporting Theorem 2, a procurement 'bargain' that is offered at stage 2 ($r_2 = 32$, $r_2 < (r_3 + w_3)/P_3$), has a different effect on the optimal decision policy depending on the shortage penalty level. The optimal policy does not take advantage of the 'bargain' when a low penalty is in effect. With the exception of Stage 4, all lower control limits, including Stage 2, assumed zero values. It seems that the purchasing price is still too high for the low level shortage (52). Under a higher shortage penalty (100), the optimal policy reacted as predicted, namely it displayed a zero lower control limit at Stage 3, which preceded Stage 2 where the bargain was offered. Then, at Stage 2, procurement was encouraged through the existence of a strictly positive lower control limit. A strictly positive lower control limit was also exhibited at Stage 1 where, compared with the relatively high shortage penalty, the procurement cost (50), was again economically attractive.

Set 3. Another procurement 'bargain' is offered at Stage 2. This time the conditions are such that the bargain is economically attractive even under a low-level shortage cost. This attractiveness is revealed in two ways: first (as expected) in terms of a positive lower control limit at Stage 2, but then also through a more unusual outcome. Both lower and upper control limits had zero values at Stage 4 (the first stage in manufacturing order). The meaning of this result is that it is not economical to produce at Stage 4, but only at the next Stage 3 (second stage in manufacturing order). However, procurement at Stage 3 is discouraged ($L_3 = 0$). Accordingly, a feasible solution (through an interpretation of this result) would be to start production only at Stage 2 (Stage 3 in manufacturing order). The low operating cost as calculated for this situation reflects promises of economic benefits, likely to occur from replacing the current decision to 'produce' at the first and the second manufacturing stages (remaining Stages 4 and 3 here) by a decision to 'procure'. Hence, for production to be optimal it should start at the manufacturing Stage 3 (remaining Stage 2 here) using as raw material the intermediate material at Stage 2. Under these circumstances, from a pragmatic perspective it becomes very important to ascertain the reliability of the supplier, with respect to both quantity and quality of the semi-produced product, which will thus become *the* raw material. Another possibility would be to consider drastic improvements (reduced manufacturing costs and/or yield) of Stages 4 and 3 so as to make them compete successfully with the price offered by the external supplier.

MULTI-PRODUCTION RUNS

Let us now consider the possibility that if, by the end of a production run, the number of final, non-defective units produced, falls short of the given demand, additional production runs may be carried out. Overlapping of production runs is not encouraged, since in this case the decision on the quantity to produce is complicated by having to rely on partial information about the unsatisfied demand. Hence, an additional production run will not be practical unless the length of the period up to the delivery date is several times that of the expected duration of a single production run.

A few authors[11,23-25] have considered additional production runs to satisfy a make-to-order demand for the single stage problem. In most of these papers it was assumed that the number

of non-defective units in a batch follows a binomial distribution. Bowman and Fetter (see Reference 23, pp. 324–330) considered set-up costs for each additional production run.

Here we are concerned with extending the single production run N-stage formulation, to a multiple production run problem, using a similar scenario. It is assumed that the maximum number of production runs (M) that can be carried out to satisfy the make-to-order demand is given. In practice, M can be assessed through dividing the (contracted) delivery period by the estimated duration of a single production run. The developed procedure is compatible with two versions. One version assumes that no set-up is needed for launching an additional run, while in the second version a positive set-up is assumed and, accordingly, an economic decision has to be made at the end of each production run on whether or not it is worthwhile to start a new one.

A decomposition approach

To amend the model for a multiple production run formulation, a function $F_i^t(U_i, D^t)$, representing the minimal total expected cost, given the remaining number of runs t, $(t = 1, 2, \ldots, M)$, the input to stage i, U_i, $(i = 1, 2, \ldots, N)$ and the unfulfilled demand D^t, was defined. As it was shown that, for $t > 1$, non-convex functions may be obtained[19], an optimal policy with a structure similar to the one formulated for the single run, could not be determined. Hence, the solution builds on the single run analytical result by developing an approximation in which the problem is decomposed into a series of single runs. The approximation assumes that $F_N^t(U_N^*, D^t)$ (the minimal total expected cost given that a production run intended to satisfy the given demand D^t, starts with an optimal input, U_N^*) is linear in D^t. If follows that at the end of any run t, $t > 1$, the shortage penalty π^t can be approximated as the optimal total expected operating cost of the next production run, $t - 1$, for a given unsatisfied demand of one unit $(D^{t-1} = 1)$. For $t = 1$ the shortage penalty assumes its original value, $\pi^1 = \pi$. To estimate π^t for $t = 2, 3, \ldots, M$, the system was first operated backward, run by run. Then the operating procedure is run forward to determine the 'optimal' decision variables and the expected cost at each run.

Algorithms were developed for each of the two versions of this procedure and experiments, with varying values of the parameters involved, were conducted. The insights obtained are as follows.

● The shortage penalties π^t (at the end of a given number of remaining production runs t, $t > 1$) expressed as ratios of the final shortage penalty, π, show (for all cases) an asymptotic decrease with the number of the remaining runs t. Accordingly, when the first version (no set-up) is run, the total expected cost decreases with the maximum number of production runs allowed, M.

● Similarly to the single production run, zero valued lower control limits (prohibiting procurement) were obtained for high procurement costs at intermediate stages.

● When only one additional production run is carried out, the resulting reduction in the total expected operating cost (expressed as a percentage) is higher for a higher shortage penalty. This reduction is only slightly affected by 'active' procurement at intermediate stages and decreases with the shortage penalty.

CONCLUDING REMARKS

● A decision model for a multi-stage process with discrete distribution of yield losses (related to the input levels), has been presented. Allowing for purchasing of intermediate products, or carrying out rework at intermediate stages, requires an optimal policy defined by two critical numbers (control limits).

● It is proved that the policy of removing non-defective units (one critical number) is a special case of the more general model as presented here. This structure of the optimal policy, extending the multi-stage model of Lee and Yano[3] (which proved the optimality of a single critical number), will also hold for the stochastic proportional yield model under

similar cost modelling, and can thus be considered as an extension of Wein's model[4] as well.

- As the multi-production run formulation of this problem shows that there is no simple 'optimal' policy, an approximate approach was developed to find 'good' control rules. This approximation enabled us to accommodate two modelling variations of the system (with and without a set-up cost for starting an additional production run).
- Further research on the discrete modelling of such systems may enhance the current assumptions.

Some related issues follow.

Quality improvement and yield modelling

The economic implications of improving the quality of the product at each stage may be taken into consideration as a possible enhancement of the model[26]. Such a possibility may especially be fruitful in wafer fabrication processes, where some authors report a U-shaped distribution of the yield[3], which may reasonably be a result of mixed batches of varying quality levels.

A more sophisticated mechanism for the yield such as clustering of defects (found appropriate in some production processes like IC fabrication[27]) may be also considered.

Economic location of inspection stations and inspection reliability

In the current work we assumed that preceding each manufacturing stage, a perfectly reliable inspection station was located. By removing these restrictions, the procedure may be modified to handle situations connected with the selective location of inspection stations[28], or coping with unreliable inspection. Optimal decisions should consider trade-offs between the cost of carrying out the inspections and the economic benefits of the information they provide, in a production system abiding by 'control limits', such as the above.

APPENDIX A

Proof of Proposition 1

From (2) we get:

$$F_1(U_1 + 1) = w_1(U_1 + 1) + \pi \sum_{x<D} (D - x)P_1[x|(U_1 + 1)] + h_o \sum_{x>D} (x - D)P_1[x|(U_1 + 1)].$$

(A1)

Using elementary probability theory, we obtain for any $x > 0$:

$$P_1[x|(U_1 + 1)] = p_1 P_1[(x - 1)|U_1] + (1 - p_1)P_1(x|U_1).$$

(A2)

Substituting (A2) in (A1) and rearranging yields:

$$\Delta F_1(U_1) = w_1 - p_1\pi + p_1(\pi + h_o)\Psi_1(U_1)$$

(A3)

where,

$$\Psi_1(U_1) = P_1[x \geq D|U_1].$$

(A4)

Since for every $U_1 \geq D$, the function $\Psi_1(U_1)$ is strictly increasing in U_1, on (A3) we conclude that $\Delta F_1(U_1)$ is a monotone increasing function of U_1.

To show that $F_1(U_1)$ attains a unique minimum, we note that

$$\lim_{U_1 \to \infty} \{\Psi_1(U_1)\} = 1,$$

thus implying

$$\lim_{U_1 \to \infty} \{\Delta F_1(U_1)\} = w_1 + p_1 h_0 > 0.$$

Hence, we conclude that $F_1(U_1)$ is not strictly decreasing, which completes the proof.

APPENDIX B

Proof of Theorem 1

Here we provide a proof of Part 1. Parts 2 and 3 of Theorem 1 can be established in a manner similar to that used in Proposition 2. The proof of Part 1 is by induction on the number of processing stages. For $i = 1$, Part 1 reduces to Proposition 1.

Assume that Part 1 holds for stages $1, 2, \ldots, i$.

By conditioning the expected cost $F_{i+1}(U_{i+1})$, on x and using the induction hypothesis, we obtain:

$$F_{i+1}(U_{i+1} + 1) = \sum_{x=0}^{L_i} [F_i(L_i) + r_i(L_i - x)] P_{i+1}(x|U_{i+1}) + \sum_{x=L_i+1}^{M_i-1} F_i(x) P_{i+1}[x|(U_{i+1} + 1)]$$

$$+ \sum_{x=M_i}^{U_{i+1}+1} [F_i(U_i) + h_i(x - M_i)] P_{i+1}(x|U_{i+1}) + w_{i+1}(U_{i+1} + 1). \qquad (B1)$$

Let us use the following property:

$$P_{i+1}[x|(U_{i+1} + 1)] = p_{i+1} P_{i+1}[(x - 1)|U_{i+1}] + (1 - p_{i+1}) P_{i+1}(x|U_{i+1}). \qquad (B2)$$

Substituting (B2) in (B1) and rearranging, we get:

$$\Delta F_{i+1}(U_{i+1}) = w_{i+1} + p_{i+1} \Psi_{i+1}(U_{i+1}) \quad \text{where,} \qquad (B3)$$

$$\Psi_{i+1}(U_{i+1}) = \left\{ -r_i \sum_{x=0}^{L_i-1} P_{i+1}(x|U_{i+1}) + \sum_{x=L_i}^{M_i-1} \Delta F_{i+1}(x) P_{i+1}(x|U_{i+1}) + h_i \sum_{x=M_i}^{U_i+1} P_{i+1}(x|U_{i+1}) \right\}.$$

$$(B4)$$

Before we establish the discrete monotonicity of $F_{i+1}(U_{i+1})$, let us prove that $\Psi_{i+1}(U_{i+1})$ is an increasing monotone function in U_{i+1}. Substituting (B2) in (B4) we get:

$$\Psi_{i+1}(U_{i+1} + 1) = (1 - p_{i+1}) \Psi_{i+1}(U_{i+1}) + p_{i+1} \left\{ -r_i \sum_{x=0}^{L_i-2} P_{i+1}(x|U_{i+1}^{\cdot}) \right.$$

$$\left. + \sum_{x=L_i-1}^{M_i-2} \Delta F_i(x + 1) P_{i+1}(x|U_{i+1}) + h_i \sum_{x=M_i-1}^{U_i+1} P_{i+1}(x|U_{i+1}) \right\}. \quad (B5)$$

Transferring $\Psi_{i+1}(U_{i+1})$ to the left-hand side while substituting $\Psi_{i+1}(U_{i+1})$ in the right-hand side and changing the summations' bounds, we obtain:

$$\Delta\Psi_{i+1}(U_{i+1}) = p_{i+1} \left\{ [\Delta F_i(L_i) + r_i] P_{i+1}[(L_i - 1)|U_{i+1}] \right.$$

$$- [\Delta F_i(M_i - 1) - h_i] P_{i+1}[(M_i - 1)|U_{i+1}]$$

$$\left. + \sum_{x=L_i}^{M_i-2} [\Delta F_i(x + 1) - \Delta F_i(x)] P_{i+1}(x|U_{i+1}) \right\}. \qquad (B6)$$

Applying the induction hypothesis we obtain:
(1) $\Delta F_i(L_i) + r_i > 0$
(2) $\Delta F_i(M_i - 1) - h_i < 0$
(3) $\Delta F_i(x + 1) - \Delta F_i(x) > 0$ for every non-negative integer x. Thus, we conclude that $\Delta\Psi_{i+1}(U_{i+1}) > 0$. By noticing that $\lim_{U_{i+1} \to \infty} \{\Delta F_{i+1}(U_{i+1})\} = w_{i+1} + p_{i+1} h_i > 0$, we finally get the desired result.

APPENDIX C

Recurrence equations for $\Delta F_i(0)$ and special cases of Theorem 2

$\Delta F_i(0)$ is determined by solving the following recurrence equations.
1. For $i = 1$ we get from (A3):

$$\Delta F_i(0) = w_1 - p_1 \pi.$$

2. For $i > 1$ three occurrences are possible. Hence, on (B3), we obtain:
   2.1 $L_{i-1} > 0$ [i.e. $\Delta F_i(0) + r_{i-1} < 0$] then,
   $\Delta F_i(0) = w_i - p_i \pi.$
   2.2 $L_{i-1} = 0$ [i.e. $\Delta F_i(0) + r_{i-1} > 0$]
   and
   $M_{i-1} > 0$ [i.e. $\Delta F_i(0) - h_{i-1} < 0$] then,
   $\Delta F_i(0) = w_i + p_i \Delta F_{i-1}(0).$
   2.3 $L_{i-1} = 0$ and $M_{i-1} = 0$ [i.e. $\Delta F_{i-1}(0) - h_{i-1} > 0$], then
   $\Delta F_i(0) = w_i + p_i h_{i-1}.$
The following are two special cases.

Special Case 1

Consider examining Stage i and assume that $L_{i-1} > 0$ (determined by applying Theorem 2 to Stage $i - 1$). It can be shown that

$L_i > 0$ if and only if $r_{i-1} > (w_i + r_i)/p_i$ $r_0 = \pi$ (which represents condition (C2))

$U_i > 0$ if and only if $r_{i-1} > w_i/p_i$ (C3)

$M_i > 0$ if and only if $r_{i-1} > (w_i - h_i)/p_i.$ (C4)

Special Case 2

Again consider examining Stage i and assume that $L_d = 0$ and $M_d > 0$ for every $d < i$ (determined by applying the recurrence equations to each stage $d < i$).

According to Theorem 2, $M_i > 0$ if and only if $\Delta F_{i-1}(0) - h_i < 0$.

Applying the recurrence equation as appropriate here, namely, $\Delta F_i(0) = w_i + p_i \Delta F_{i-1}(0)$ (case 2.2), using simple induction to calculate $\Delta F_{i-1}(0)$ and substituting in the above necessary condition for $M_i > 0$, we obtain,

$$-h_i + \sum_{j=0}^{i-1}\left[w_{i-1}\prod_{k=1}^{j} p_{i+1-k} \right] - \pi\prod_{k=1}^{j} p_k < 0 \Rightarrow M_i > 0 \quad \text{for every } i.$$

Hence, an analogous result to the one presented by Lee and Yano[3] is obtained. It should be clarified that in contrast to Reference 3, where intermediate procurement is not considered, this is a special case, obtained when the lower control limits assume zero values, i.e. procurement is prohibited.

REFERENCES

1. S. KARLIN (1958) One stage models with uncertainty. In *Studies in the Mathematical Theory of Inventory and Production.* (K. J. ARROW, S. KARLIN and H. SCARF, Eds) pp. 109–134. Stanford University Press, CA.
2. W. SHIH (1980) Optimal inventory policies when stockouts result from defective products. *Int. J. Prod. Res.* **18**, 677–686.
3. H. L. LEE and C. A. YANO (1988) Production control for multistage systems with variable yield losses. *Opns Res.* **36**, 269–278.
4. A. S. WEIN (1992) Random yield, rework and scrap in a multi-stage batch manufacturing environment. *Opns Res.* **40**, 551–563.
5. C. A. YANO and H. L. LEE (1992) Lot sizing with random yields: a review. Technical Report 89-16, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan.

6. B. GIFFLER (1960) Determining an optimal reject allowance. *Naval Res. Logist. Q.* **7**, 201–206.
7. R. E. LEVITAN (1960) The optimal reject allowance problem. *Mgmt Sci.* **6**, 172–187.
8. R. EHRHARDT and L. TAUBE (1987) An inventory model with random replenishment quantities. *Int. J. Prod. Res.* **25**, 1795–1804.
9. M. HENIG and Y. GERCHAK (1990) The structure of periodic review policies in the presence of random yield. *Opns Res.* **38**, 634–643.
10. Y. GERCHAK, R. G. VICKSON and M. PARLAR (1988) Periodic review production models with variable yield and uncertain demand. *IIE Trans.* **20**, 144–150.
11. M. KLEIN (1966) Markovian decision models for reject allowance problems. *Mgmt Sci.* **12**, 349–358.
12. A. BEJA (1977) Optimal reject allowance with constant marginal production efficiency. *Naval Res. Logist. Q.* **24**, 21–33.
13. Y. GERCHAK and M. PARLAR (1990) Yield variability, cost tradeoffs and diversification in the EOQ model. *Naval Res. Logist.* **37**, 341–354.
14. T. C. CHENG (1991) EPQ with process capability and quality assurance considerations. *J. Opl Res. Soc.* **42**, 713–720.
15. M. VACHANI (1970) Determining optimum reject allowances for multistage job-lot manufacturing. *AIIE Trans.* **2**, 70–77.
16. C. S. TANG (1990) The impact of uncertainty on a production line. *Mgmt Sci.* **36**, 1518–1531.
17. C. Y. WOO, L. B. SCHWARZ, A. C. SULLIVAN, C. J. BERGER, S. CHAND and D. SHORT (1993) Cases on the implementation of cell manufacturing technology—Ingersoll Rand Corporation: The Power Tool Division (Case A). In Working Paper Series 93-2-2, Center for the Management of Manufacturing Enterprises, Krannert School of Management, Purdue University, W. Lafayette IN, USA, 1-30.
18. T. A. WASSICK, C. B. FIELD and C. C. HU (1993) Repair and rework. In *Microelectronics Manufacturing Diagnostics Handbook*. (A. H. LANDZBERG, Ed.) pp. 205–230. Van Nostrand Reinhold, New York.
19. M. BARAD and D. BRAHA (1991) Optimal policies for coping with stochastic yield in multi-stage discrete processing systems: parts I and II. Working Paper, Department of Industrial Engineering, Tel Aviv University, Israel.
20. P. J. BONK, M. R. GUPTA, R. A. HAMILTON and A. V. S. SATYA (1993) Manufacturing yield. In: *Microelectronics Manufacturing Diagnostics Handbook*. (A. M. LANDZBERG, Ed.) pp. 9–35. Van Nostrand Reinhold, New York.
21. A. V. FERRIS-PRABHU (1990) A cluster-modified Poisson model for estimating defect density and yield. *IEEE Trans Semi Conductor Mftg* **3**, 54–59.
22. J. A. CUNNINGHAM (1990) The use and evaluation of yield models in integrated circuit manufacturing. *IEEE Trans. Semiconductor Mftg* **3**, 60–71.
23. E. H. BOWMAN and R. B. FETTER (1960) *Analysis for Production Management*. Irwin Homewood, Illinois.
24. S. WHITE (1965) Dynamic programming and systems of uncertain duration. *Mgmt Sci.* **12**, 37–67.
25. M. SEPEHRI, E. A. SILVER and C. NEW (1986) A heuristic for multiple lot sizing for an order under variable yield. *IIE Trans.* **18**, 63–69.
26. M. BARAD and G. BENNETT (1994) Optimal yield improvement strategies in multi-stage manufacturing processes. Working Paper, School of Mechanical and Manufacturing Engineering, UNSW, N.S.W. Australia 2033.
27. S. L. ALBIN and D. J. F. FRIEDMAN (1989) The impact of clustered defect distributions in IC fabrication. *Mgmt Sci.* **35**, 1066–1078.
28. M. BARAD (1990) A break-even quality level approach to location of inspection stations in a multi-stage production process. *Int. J. Prod. Res.* **28**, 29–45.