MANAGEMENT SCIENCE
Vol. 38, No. 1, January 1992
Printed in U.S.A.

PRODUCTION LOT SIZING WITH MACHINE BREAKDOWNS*

HARRY GROENEVELT, LILIANE PINTELON AND ABRAHAM SEIDMANN
William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, New York 14627
Katholieke Universiteit Leuven, Leuven, Belgium
William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, New York 14627

Economic lot sizing and batching models often assume reliable manufacturing facilities. In this research, we focus on the effects of machine breakdowns and corrective maintenance on the economic lot sizing decisions. Two production control policies are proposed for coping with these stochastic interferences. The first policy assumes that production of the interrupted lots is not resumed after a breakdown. Instead, the on-hand inventory is depleted before a new cycle is initiated. Under the second policy studied here, production is immediately resumed after a breakdown, if the current on-hand inventory is below a certain threshold level. It is shown that this control structure is optimal among all stationary policies. We show that under both policies the optimal lot sizes will always be bigger than the ones in a corresponding deterministic case, and that the optimal lot size increases with the failure rate. We also provide exact optimal and closed form approximate lot sizing formulas and derive tight bounds on the average cost per unit time for the approximations. In addition, we present various structural properties for these policies and operational insights relevant to such management decisions as machine replacement or maintenance schedules.
(PRODUCTION PLANNING; LOT SIZING; MACHINE BREAKDOWNS)

1. Introduction

A major trend in discrete part manufacturing has been the introduction of JIT deliveries of parts and components along a series of vertically integrated manufacturing facilities. One of the major impediments for the successful operation of such a tightly coupled organization is formed by breakdowns in bottleneck resources. In this paper we investigate the impact of equipment breakdowns on the operating policy in a basic manufacturing environment. The specific operational decisions we focus on are production lot sizing and the determination of safety stocks. We study these problems in the context of the classic EMQ model (Harris 1915, see Erlenkotter 1989). The purpose of our analysis is to gain insight into the effects of the occurrence of breakdowns on the lot sizing decision. These insights can enhance and support such resource allocation decisions as corrective maintenance efforts, machine replacements and processing rates determination. In a subsequent paper (Groenevelt et al. 1992) we consider the problem of simultaneously determining lot sizing and safety stock policies when repair times are significant.

The EMQ model (economic order quantity/economic manufacturing or production quantity) is probably the oldest inventory control model. Osteryoung et al. (1986) report that these models are still widely used in industry, although the assumptions (constant demand rate, known inventory holding and setup cost, no shortages allowed, unit production cost independent of the lot size, constant production rate, infinite horizon) necessary to justify their use are rarely met. Numerous research efforts have been undertaken

* Accepted by Awi Federgruen; received November 8, 1989. This paper has been with the authors 2 months for 2 revisions.

0025-1909/92/3801/0104$01.25

to extend these models to conform more closely with real-world situations, see, e.g., Silver and Peterson (1985), Naddor (1984), and Nahmias (1989) for recent surveys. The study by Lev et al. (1981) focuses on how EMQ models are affected by cost strategies or extraneous influences which change parameters. The literature provides models that deal with the effect of learning and forgetting (Axsäter and Elmaghraby 1981, Muth and Spremann 1983), quantity discounts (Sethi 1984, Kuzdrall and Brittney 1982), limited storage capacity (Murdeshwar and Sathe 1985), deteriorating items (Shah 1977), inflation (Jesse et al. 1983), nonlinear holding costs (Brown et al. 1986), purchase price fluctuations (Goyal 1979), joint purchaser/vendor replenishment decisions (Banerjee 1986), etc. Other studies modify the basic EMQ model in order to cope with a varying demand pattern (Ritchie and Tsado 1986, Resh et al. 1976, Silver and Meal 1969, Snyder 1977, Wagner and Whitin 1958) or a finite horizon (Schwarz 1977, Lev et al. 1979). Setup cost reduction, inspired by the Japanese process-oriented manufacturing philosophy, has received recent attention (Billington 1987, Porteus 1985). The above cited EMQ studies ignored the imperfections in the production process (quality issues) and equipment (machine breakdowns). Billington (1986) focuses on quality improvement through lot size reduction. Optimal order quantities are derived for two output inspection policies in Lee and Rosenblatt (1985), optimal order quantity and inspection schedule are simultaneously determined in Lee and Rosenblatt (1987). The impact of inspection costs on the lotsize is also studied by Schwaller (1988). Porteus (1986) shows that for a process that can go out-of-control with a given probability each time it produces an item, it is better to produce smaller-than-EMQ lot sizes and explores different options for investing in quality improvement and setup cost reduction. Rosenblatt and Lee (1986) analyze the case where the system deteriorates during the production process and produces some proportion of defective items. The optimal lot size is again smaller than the classical EMQ lot size. Different deterioration processes are considered. See Yano and Lee (1989) for a recent survey. Other studies investigating the single-machine, single-product lot sizing decision with a random processing rate include Karmarkar (1987), and Zipkin (1986).

Our research also starts from the classical EMQ model, but assumes that the equipment is subject to stochastic breakdowns. The effect of these breakdowns on the optimal lot size and reorder level decisions is examined for two different operating policies. The first, the no-resumption (NR) policy, is discussed in §2 for the general failure distribution case. The case of exponential failures is elaborated upon in §3. The second, the abort/resume (AR) policy, is presented in §4 and further analyzed in the Appendix. §5 summarizes the insights gained from the major trade-offs associated with these policies.

2. The No-Resumption Policy with General Machine Breakdowns (NR-G)

2.1. *Assumptions and Notation*

In this section we develop a modified EMQ model with general machine breakdowns and no resumption (NR-G). Under this policy, when a breakdown takes place the interrupted lot is aborted and a new one is to be started only when all available inventory is depleted. Maintenance of the production unit is carried out after a failure or after a predetermined time interval (here the production time for the predetermined lot size), whichever occurs first. Each maintenance action restores the system to the same initial working conditions. Starting a new lot after a breakdown incurs a setup cost plus a corrective maintenance charge. We assume deterministic demand at a constant rate, and a continuous finite production rate. The inventory holding cost rate is proportional to on hand inventory, and the unit variable production cost is independent of the lot size.

Figure 1 illustrates the difference in the on-hand inventory sample path between the

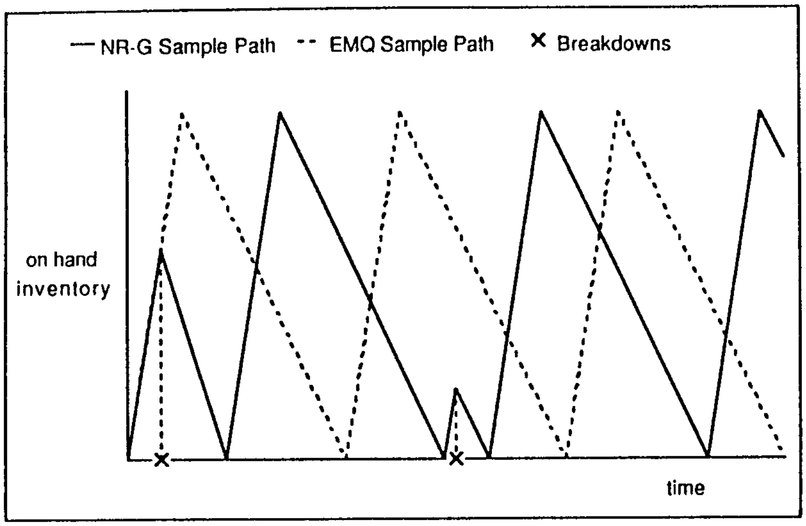

FIGURE 1. EMQ Model and NR-G Sample Paths: Inventory Level vs. Time.

classical EMQ model and the NR-G policy studied here. Under the EMQ model, all cycles are of equal length. However, two different cycle types occur for the NR-G policy. The first is a cycle without breakdowns which looks like an ordinary EMQ cycle. The second is a shorter cycle as a result of a machine breakdown.

Table 1 summarizes the notation used in formulating this model. Note that neither setups nor corrective maintenance are assumed to be time consuming in the model. In many flexible production systems setups are indeed very small and can be neglected. Instantaneous maintenance interventions are very common for production systems with modular design, where maintenance consists of replacing the unit or module in question by a new one. Repair of the failed unit happens "off line" and does not affect the production system. Another situation in which maintenance times can be (approximately) neglected is when the probability of a breakdown shortly after a setup takes place is very small (e.g., because preventive maintenance is performed as part of the setup procedure). With high probability there will then be enough inventory build-up to allow the corrective maintenance intervention to take place and do a setup while available inventory is running down.

2.2. *Mathematical Formulation*

Our first objective is to determine the optimal lot size Q^*, i.e., the lot size that minimizes $\bar{C}$, the (long-run) average cost per unit of time. Note that the inventory process has

TABLE 1
Notation Summary for the NR-G Model

d:	deterministic, constant demand rate (units/day)
p:	deterministic, constant production rate (units/day)
h:	inventory holding cost ($/day/unit)
S:	production setup cost ($)
M:	cost of corrective maintenance ($)
Q:	maximum or target lot size (units)
T:	random variable denoting time-to-breakdown (days)
$F(t)$:	cumulative density function of T
$f(t)$:	probability density function of T; $f(t) = F'(t)$
$r(t)$:	failure rate of T; $r(t) = f(t)/(1 - F(t))$

renewal epochs at the start of every production run. Hence by the well-known renewal reward theorem (see, e.g., Ross 1981) the long-run average cost can be found by taking the ratio of the expected cost per renewal cycle and the expected duration of a renewal cycle. We find by conditioning on the failure time:

$$E[\text{Cost per cycle}] = \int_0^\infty E[\text{Cost per cycle} \mid T = t] f(t)\,dt$$

$$= \int_0^{Q/p} E[\text{Cost per cycle} \mid T = t] f(t)\,dt + \int_{Q/p}^\infty E[\text{Cost per cycle} \mid T = t] f(t)\,dt$$

$$= \int_0^{Q/p} \left\{ S + M + \frac{1}{2} h(p-d)\frac{p}{d} t^2 \right\} f(t)\,dt + \int_{Q/p}^\infty \left\{ S + \frac{1}{2} h \frac{p-d}{pd} Q^2 \right\} f(t)\,dt$$

$$= S + M \cdot F(Q/p)$$

$$+ \frac{1}{2} h(p-d)\frac{p}{d}\left\{ (Q/p)^2(1 - F(Q/p)) + \int_0^{Q/p} t^2 f(t)\,dt \right\} \quad \text{and} \quad (2.1)$$

$$E[\text{Duration of a cycle}] = \int_0^\infty E[\text{Duration of a cycle} \mid T = t] f(t)\,dt$$

$$= \int_0^{Q/p} \frac{p}{d} t f(t)\,dt + \int_{Q/p}^\infty (Q/d) f(t)\,dt$$

$$= (p/d)\int_0^{Q/p} t f(t)\,dt + (Q/d)(1 - F(Q/p)). \quad (2.2)$$

Hence the long-run average cost per unit of time is given by

$$\bar{C}(Q) = \frac{E[\text{Cost per cycle}]}{E[\text{Duration of a cycle}]}$$

$$= \frac{S + M \cdot F\left(\frac{Q}{p}\right) + \frac{1}{2} h(p-d)\frac{p}{d}\left\{ \left(\frac{Q}{p}\right)^2 \left(1 - F\left(\frac{Q}{p}\right)\right) + \int_0^{Q/p} t^2 f(t)\,dt \right\}}{(p/d)\int_0^{Q/p} t f(t)\,dt + (Q/d)(1 - F(Q/p))}. \quad (2.3)$$

An optimal lot size Q^* is one that minimizes the cost expression (2.3).

3. The No-Resumption Model with Exponential Failure Distribution (NR-E)

3.1. *Optimal Lot Size for Exponential Failure Distribution*

The cost expression (2.3) is fairly complex and for general failure distributions does not allow a closed formula expression to be derived for the optimal lot size. Of course it is always possible to solve numerically for Q^*. However, for the exponential failure distribution some fairly simple formulas can be derived. In this section we will derive these formulas to obtain insights into the trade-offs involved in the production system operations. Hence we assume here that $F(t) = 1 - e^{-\lambda t}$, where λ = failure rate when production is in progress. Note that the failure rate $r(t) = \lambda$ is constant over time. A constant failure rate is considered to be an appropriate approximation to describe the failure pattern of a system consisting of many components, each subject to an individual pattern of breakdown or malfunction (see, e.g., Barlow and Proschan 1965). We will assume throughout the rest of the paper that $\lambda > 0$. For $\lambda = 0$ our model reduces of course to the standard EMQ model.

For the case of exponential failures, formula (2.3) then gives

$$\bar{C}(Q) = \frac{\frac{d\lambda}{p} S}{1 - e^{-(\lambda Q/p)}} + \frac{d\lambda}{p} M + \frac{h(p-d)}{\lambda} \left\{ 1 - \frac{\frac{\lambda Q}{p} e^{-(\lambda Q/p)}}{1 - e^{-(\lambda Q/p)}} \right\}. \quad (3.1)$$

Differentiating (3.1) yields

$$\frac{\partial \bar{C}(Q)}{\partial Q} = \frac{e^{-(\lambda Q/p)}}{[1 - e^{-(\lambda Q/p)}]^2} \left\{ -\frac{d\lambda}{p} S + \frac{h(p-d)}{\lambda} (e^{-(\lambda Q/p)} + \lambda Q/p - 1) \right\}. \quad (3.2)$$

By setting $\partial \bar{C}(Q)/\partial Q = 0$ and simplifying the resulting equation it is easy to verify that the optimal lot size Q^* is the *unique* nonnegative solution of the nonlinear equation:

$$e^{-(\lambda Q/p)} + \frac{\lambda Q}{p} = 1 + \frac{d\lambda^2 S}{hp(p-d)}. \quad (3.3)$$

Note that Q^* is well defined because the left-hand side of (3.3) viewed as a function of Q is increasing for $Q \geq 0$ and assumes all values in the range $[1, \infty)$. Note also that $\bar{C}(Q)$ assumes its global minimum at Q^* since $\bar{C}(Q)$ is a continuously differentiable function of Q with a derivative that changes sign only once (at Q^*) from negative to positive.

3.2. *Properties of the Lot Sizing Model with Exponential Failures*

In this subsection we will derive some structural properties of the modified lot sizing model with exponential failures.

Property 1. The long-run average corrective maintenance cost is independent of the lot size Q.

PROOF. The second term of expression (3.1) gives the long-run average maintenance cost.

Property 1 is easily explained by the fact that, as the failure rate is constant, the number of breakdowns in a given period (e.g., a year) is only determined by the value of that failure rate and the total production time required. It makes no difference how that production time is split up in lots.

Property 2. The minimum long-run average total cost in NR-E is given by

$$\frac{h(p-d)}{p} Q^* + \frac{d\lambda}{p} M,$$

where Q^* satisfies (3.3).

PROOF. Use (3.3) to eliminate the factor $e^{-(\lambda Q/p)}$ from the cost expression (3.1) and simplify.

Property 2 specifies a *linear* relationship between the optimal operating cost and the optimal lot size Q^*. Such a linear relationship is well known for the ordinary EMQ model. Intuitively, we would expect the optimal operating cost to increase with the failure rate. This is confirmed by Property 3.

Property 3. The optimal lot size and the optimal objective value are increasing functions of the failure rate λ.

PROOF. To simplify the notation somewhat, we will write Q^* and $Q^{*\prime}$ instead of $Q^*(\lambda)$ and $Q^{*\prime}(\lambda)$, and we define $b = dS/hp(p-d)$. In view of Property 2 it is sufficient to prove that $Q^{*\prime} > 0$ for $\lambda > 0$. Note that Q^* is defined implicitly by (3.3). By taking the derivative w.r.t. λ in (3.3) we find after some algebra

$$\frac{\lambda Q^{*\prime}}{p}(1 - \exp\{-\lambda Q^*/p\}) = 2\lambda b - \frac{Q^*}{p}(1 - \exp\{-\lambda Q^*/p\}).$$

It follows that $Q^{*\prime} > 0$ if and only if

$$\begin{aligned} 0 &\le 2\lambda^2 b - \frac{\lambda Q^*}{p}(1 - \exp\{-\lambda Q^*/p\}) \\ &= 2\left(\exp\{-\lambda Q^*/p\} + \frac{\lambda Q^*}{p} - 1\right) - \frac{\lambda Q^*}{p}(1 - \exp\{-\lambda Q^*/p\}) \\ &= \left(2 + \frac{\lambda Q^*}{p}\right)\exp\{-\lambda Q^*/p\} + \frac{\lambda Q^*}{p} - 2, \end{aligned} \tag{3.4}$$

where the first equality follows again from (3.3). Now consider the function $\phi(z) = (2 + z)e^{-z} + z - 2$. Note that $\phi'(z) = 1 - e^{-z} - ze^{-z}$, $\phi''(z) = ze^{-z}$ (≥ 0 for $z \ge 0$), and so $\phi'(z) \ge \phi'(0) = 0$ for $z \ge 0$, and similarly $\phi(z) \ge \phi(0) = 0$ for $z \ge 0$. The property follows since $\phi(\lambda Q^*/p)$ is just the RHS of (3.4). QED

Property 4. When $\lambda \downarrow 0$ (i.e., the system approaches perfect reliability),

$$Q^* \to \sqrt{\frac{2Sdp}{h(p-d)}}$$

(i.e., the optimal lot size approaches the ordinary EMQ), and

$$\bar{C}(Q) \to \frac{Sd}{Q} + \frac{1}{2}h\frac{p-d}{p}Q$$

(i.e., the long-run average cost function approaches the long-run average cost function in the standard EMQ model).

PROOF. To show

$$\bar{C}(Q) \to \frac{Sd}{Q} + \frac{1}{2}h\frac{p-d}{p}Q,$$

take the limit for $\lambda \downarrow 0$ in the right-hand side of (3.1) and apply de l'Hospital's rule. To show $Q^* \to \sqrt{2Sdp/h(p-d)}$, note that for all $\lambda \ge 0$:

$$1 + \frac{1}{2}\left(\frac{\lambda Q}{p}\right)^2 - \frac{1}{6}\left(\frac{\lambda Q}{p}\right)^3 \le e^{-(\lambda Q/p)} + \frac{\lambda Q}{p} \le 1 + \frac{1}{2}\left(\frac{\lambda Q}{p}\right)^2$$

and we can therefore conclude from (3.3) that

$$\frac{1}{2}\left(\frac{\lambda Q^*}{p}\right)^2 - \frac{1}{6}\left(\frac{\lambda Q^*}{p}\right)^3 \le \frac{d\lambda^2 S}{hp(p-d)} \le \frac{1}{2}\left(\frac{\lambda Q^*}{p}\right)^2$$

and hence

$$(Q^*)^2 - \frac{1}{3}\frac{\lambda}{p}(Q^*)^3 \le \frac{2dpS}{h(p-d)} \le (Q^*)^2. \tag{3.5}$$

From the second inequality in (3.5) it follows that $Q^* \ge \sqrt{2Sdp/h(p-d)}$. Since Q^* is an increasing function of the failure rate, there exists a constant Y such that $Q^* \le Y$ for $0 < \lambda \le 1$. Hence for $0 < \lambda \le 1$ we have

$$(Q^*)^2 \le \frac{2dpS}{h(p-d)} + \frac{1}{3}\frac{\lambda}{p}(Y)^3 \to \frac{2dpS}{h(p-d)} \quad \text{as} \quad \lambda \downarrow 0. \qquad \text{QED.}$$

An immediate consequence of Properties 3 and 4 is also:

Property 5. The optimal lot size for NR-E is always larger than the optimal lot size in the corresponding EMQ model.

To bound the error of using the Economic Manufacturing Quantity instead of Q^* we show

Property 6. $\bar{C}(\text{EMQ})/\bar{C}(Q^*) \le 1.02$.

PROOF. Let $a = \lambda^2 dS/hp(p-d)$. It is easy to verify that

$$\bar{C}(Q) = \frac{h(p-d)}{\lambda}\gamma_a\left(\frac{\lambda Q}{p}\right) + \frac{d\lambda}{p}M, \qquad \text{where}$$

$$\gamma_a(z) = \frac{1 + a - (1+z)e^{-z}}{1 - e^{-z}}.$$

Since the term $(d\lambda/p)M$ is nonnegative and independent of Q, it follows that

$$\bar{C}(\text{EMQ})/\bar{C}(Q^*) \le \gamma_a\left(\frac{\lambda\cdot\text{EMQ}}{p}\right) \Big/ \gamma_a\left(\frac{\lambda Q^*}{p}\right) = g(a).$$

Hence $g(a)$ is an upper bound for the ratio of the cost per unit time of EMQ over the optimal NR-E policy. Note that $(\lambda\cdot\text{EMQ})/p = \sqrt{2a}$, $\lambda Q^*/p = z^*(a)$, where $z^*(a)$ denotes the (unique) solution of the equation

$$e^{-z} + z = 1 + a, \tag{3.6}$$

and it is not hard to show that $\gamma_a(z^*(a)) = z^*(a)$. We then find

$$g(a) = \frac{1 + a - (1+\sqrt{2a})e^{-\sqrt{2a}}}{z^*(a)(1 - e^{-\sqrt{2a}})}. \tag{3.7}$$

Formulating a formal proof of the 2% error bound is tedious but not difficult. One merely needs to establish tight enough lower bounds on $z^*(a)$. For instance, for $a \ge 3$, it is not hard to show that $z^*(a) \ge a + 0.9813$. [By differentiating the equation defining $z^*(a)$ with respect to a, it follows that $\partial z^*(a)/\partial a > 1$, and since $z^*(3) > 2.9813$ the inequality follows.] Hence $g(a) \le z^*(a)g(a)/(a + 0.9813) \le 1.019$ for $a \ge 3$ as can be verified with elementary calculus. We also know that $\sqrt{2a} \le z^*(a)$, and hence $g(a) \le z^*(a)g(a)/\sqrt{2a} \le 1.018$ for $0 < a \le 0.005$. On intermediate intervals $z^*(a)$ can be

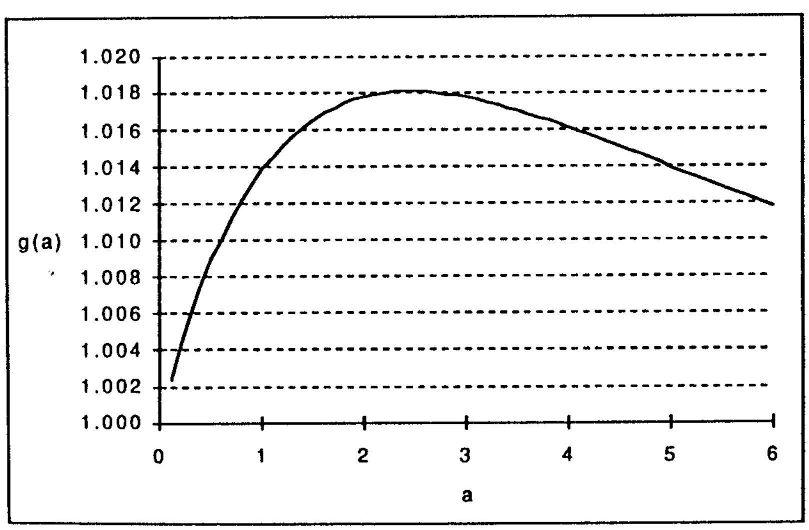

FIGURE 2. Upper Bound Cost Ratio Function $g(a)$.

TABLE 2
Numerical Input Data

	Figure 3	Figure 4	Figure 5	Figure 6	Figure 7	Figure 8	Figure 9
λ (l/day)	0/0.2/0.75	0.75	varies	varies	varies	varies	0/0.2/0.75
d (units/day)	30	30	30	30	30	30	varies
p (units/day)	35	35	35	35	35	35	35
h ($/units/day)	75	75	10; 75	10	75	75	75
S ($)	450	450	450	450	450	450	450
M ($)	1000	1000	1000	1000	1000	1000	1000
Q (units)	varies	varies	Q^*	EMQ, Q^*	Q^*	EMQ, Q^*	Q^*

bounded by linear functions to obtain the desired bound on $g(a)$. The function $g(a)$ is depicted in Figure 2. The maximum cost ratio bound is about 1.018 and occurs when $a \approx 2.5$.

3.3. *Numerical Results*

This section briefly discusses some numerical results for the NR-E model. Table 2 summarizes the numerical input data used to generate Figures 3 to 9.

Figure 3 pictures the total cost as function of the lot size Q for $\lambda = 0, 0.2, 0.75$. For larger values of λ the optimum value of Q shifts (slightly) to the right (see also Figure 5). In addition, for $\lambda > 0$ the cost curves flatten out, due to the fact that for large values of the (target) lot size Q the *actual* lot size will almost certainly be determined by equipment failure. As is to be expected, the cost curve flattens out quicker for larger values of λ. In the ordinary EMQ model ($\lambda = 0$) the cost curve has an asymptote with positive slope. The behaviour of the separate cost components as a function of λ is depicted in Figure 4 for one of the cases of Figure 3 ($\lambda = 0.75$). It illustrates the fact that the corrective maintenance cost is independent of the lot size. The inventory holding cost increases (of course) as the lot size increases, *but contrary to the* EMQ *model*, the holding costs flatten out with increasing lot size. This is due to the fact that the actual lot size is limited by the occurrence of failures. In addition, the setup cost per time unit tends to a positive constant (instead of approaching zero as in the EMQ model) when the (target) lot size increases.

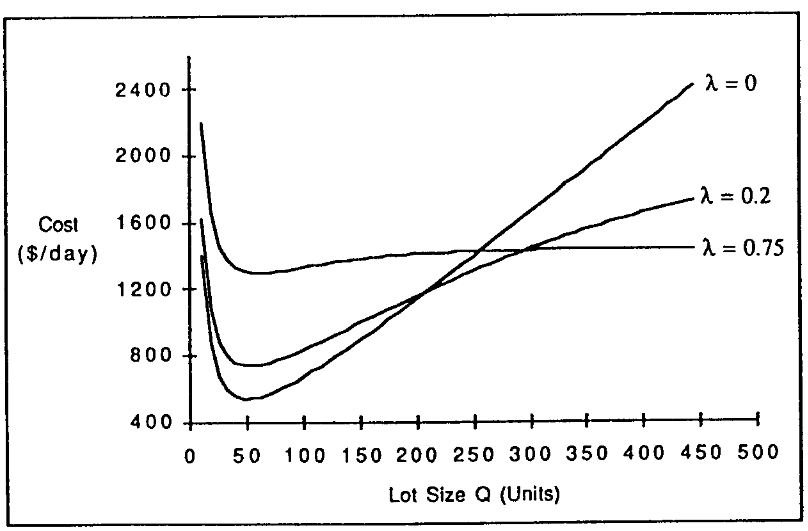

FIGURE 3. Long-Run Average Cost as a Function of Lot Size.

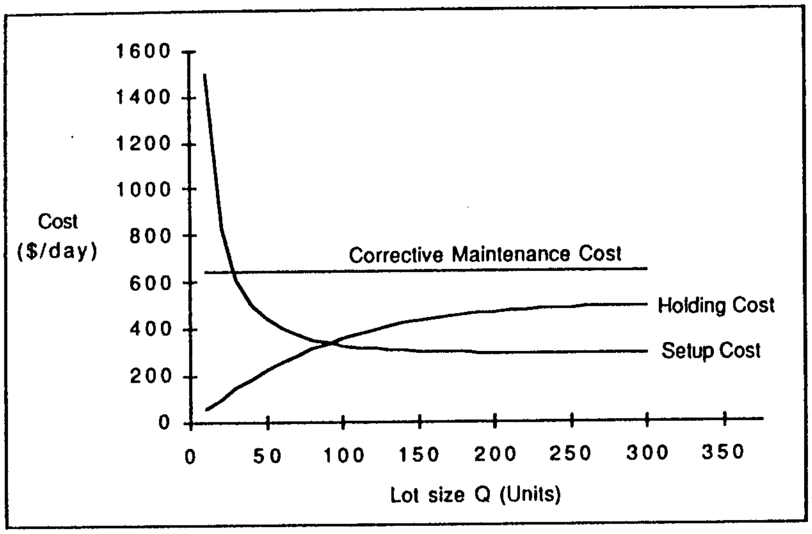

FIGURE 4. Cost Components as a Function of Lot Size.

Figure 5 gives the optimal lot size Q^* as a function of the exponential failure rate λ for two different values of the inventory holding cost. Comparison of these lot sizes with the corresponding EMQ lot size ($\lambda = 0$) shows that the higher λ the larger the difference between the optimal lot sizes for the two models, as predicted in Properties 3 and 4. This difference gets larger when the inventory holding cost parameter h decreases, and can become very large indeed. In light of this, the 2% cost ratio bound in Property 5 is quite surprising. It can be explained by noting that the target lot size is not necessarily the lot size achieved. This is illustrated in Figure 6, in which we depict target lot sizes of Q^* and EMQ as well as the average actual lot sizes achieved when using these target lot sizes as a function of the failure rate for the small holding cost case. As is evident, even though Q^* becomes more than twice as large as EMQ, the difference in the average actual lot size achieved is much smaller. Figure 6 also illustrates another important qualitative insight to be obtained from this model: when the failure rate is high, the actual value of the target lot size does not matter much; it is not the prescribed operating policy that

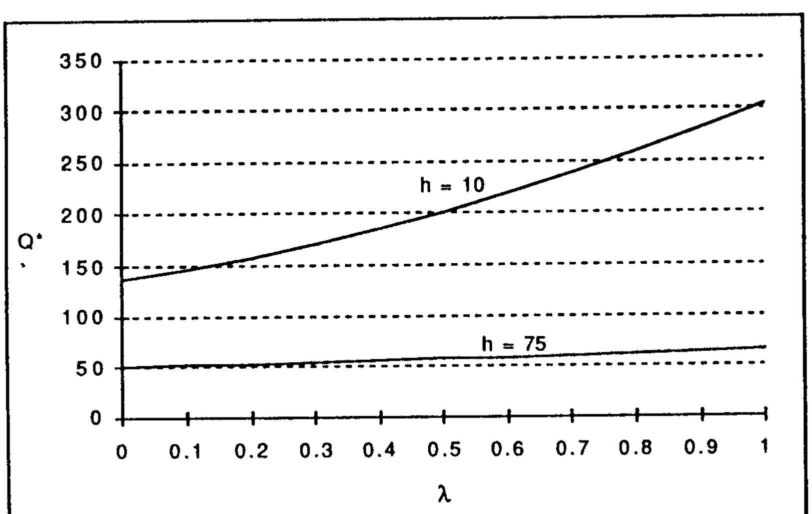

FIGURE 5. Optimal Lot Size Q^* as a Function of the Failure Rate λ.

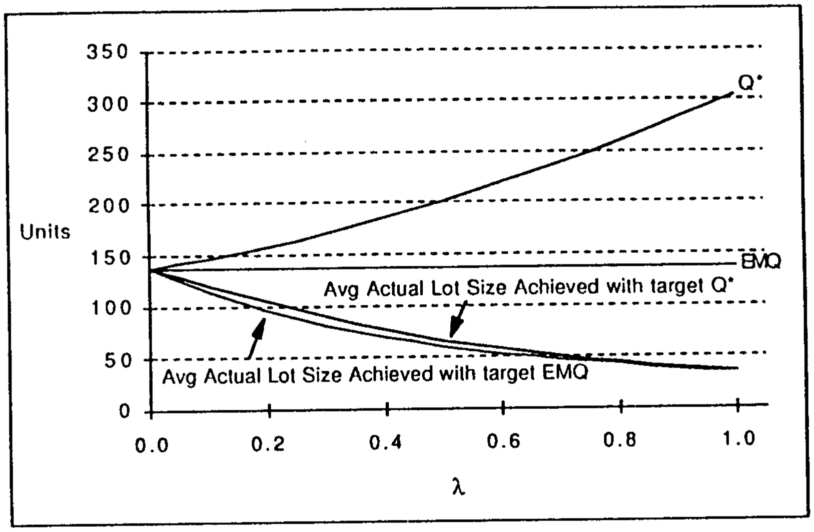

FIGURE 6. Average Actually Achieved and Target Lot Sizes as a Function of the Failure Rate λ.

determines the activities on the production floor but instead the reliability level of the equipment.

Figure 7 shows the total cost per unit time of the optimal policy and its components as a function of λ. As expected, the corrective maintenance costs increase linearly with λ due to the more frequent breakdowns, and the total costs rise slightly faster than linearly with an increasing failure rate (as can be seen from Property 2 and Figure 6). As λ increases, the average inventory holding costs decrease slightly. This decrease is the net result of two opposing trends: as the failure rate increases, more cycles will be aborted before the target lot size Q^* is reached, resulting in less inventory on average. On the other hand, as the failure rate increases, the target lot size Q^* itself increases, resulting in a tendency to hold more inventory. Setup cost increases significantly (a new setup is necessitated by every breakdown). The same holding and setup cost curves are plotted in Figure 8 as well to compare them with the costs incurred if the EMQ is used as the target lot size. It is clear that the increase in setup cost incurred when using EMQ instead

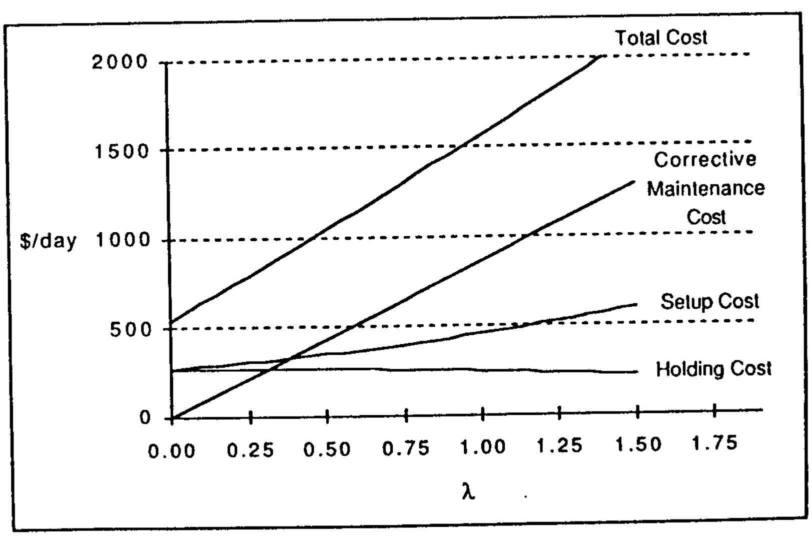

FIGURE 7. Long-Run Average Cost Components as a Function of the Failure Rate λ.

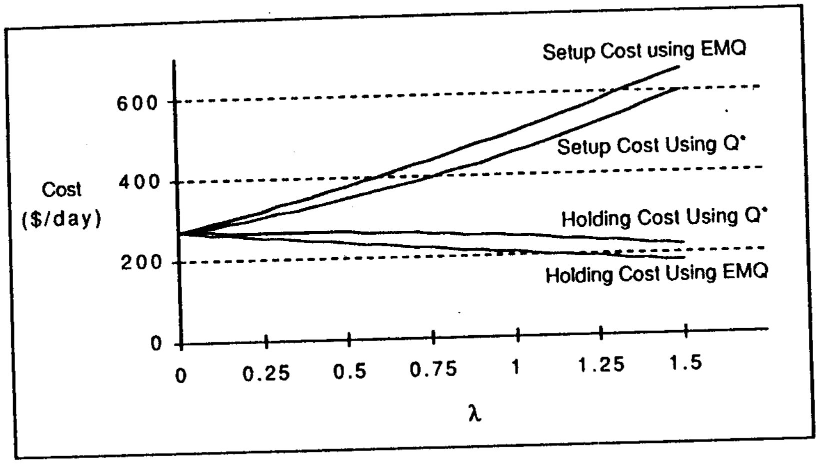

FIGURE 8. Comparison of Cost Components for Target Lot Sizes Q^* and EMQ.

of the optimal value Q^* is relatively small but not negligible. Its effect is also mostly cancelled out by a decrease in holding cost as expected from Property 6.

Figure 9 shows the optimal long-run average cost as a function of the demand rate d. When the failure rate is low ($\lambda = 0$), the curve has a maximum at $\frac{1}{2}p$ as in the EMQ model. When the failure rate increases, this maximum shifts to the right because higher demand requires more time spent producing and hence increased corrective maintenance costs. Note that when $\lambda = 0$ the total cost decreases to zero at full utilization of the facility ($d = p = 35$) due to the need to operate the machine continuously. In this case no setups are needed and no inventory is accumulated. On the other hand, when $\lambda > 0$ and $d = p$, the total cost per time unit is strictly positive due to both corrective maintenance and setup costs.

3.4. *Some Conclusions for the Exponential Failure Distribution*

Choosing too small a lot size would mean high operating costs due to numerous setups (see the very steep cost function for small lot sizes in Figure 3). Too large a lot size means a cost penalty due to increasing inventory costs (see Figures 3 and 4). This trade-off

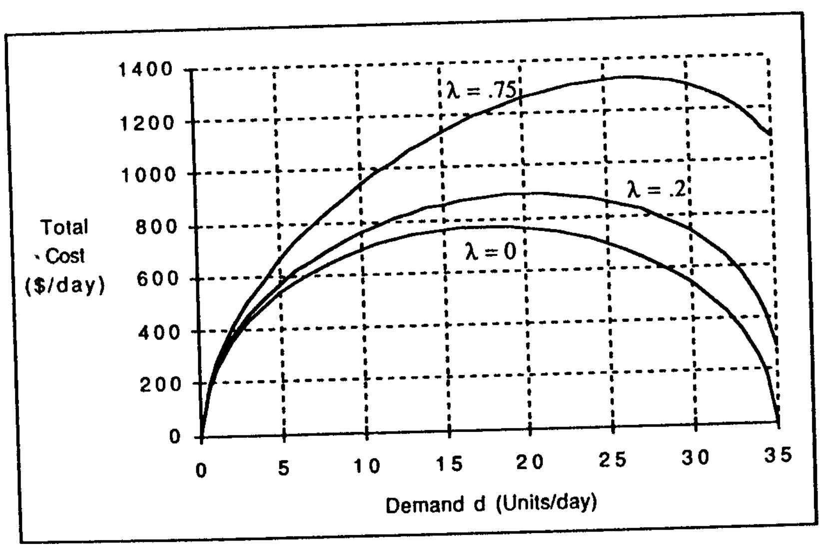

FIGURE 9. Minimum Total Average Cost as a Function of the Demand Rate d.

between setup and inventory cost is the classical trade-off in the EMQ model, but is complicated for the NR-E policy because of the disruptive effects of breakdowns and corrective maintenance costs. For example, it is not trivial to predict the optimal lot size, even in the simple case of a constant failure rate. Property 4 shows that the optimal lot size with exponential failures $Q^* \geq$ EMQ for all values of λ. On the other hand, Property 5 shows that the penalty for using EMQ instead of Q^* as the target lot size is very small for the case of exponential failures. These apparently conflicting results can be reconciled by observing that the lot size does not affect the average corrective maintenance costs at all (Property 1), and by the facts that Q^* is close to the EMQ lot size when the failure rate is small, and that when the failure rate is large the average cost function flattens out so much (see Figure 3) that the target lot size used does not matter much (as long as it is at least as large as EMQ). As Figure 6 shows, the average actual lot size produced becomes almost independent of the target lot size with an increasing failure rate.

4. The Abort/Resume Policy with Exponential Failures (AR-E)

In the models of §2 and §3 it is assumed that the current production run will always be aborted upon a failure. It is intuitively clear that this is optimal if the cost of resumption equals the ordinary setup cost S. In many cases, however, the cost of resuming the production run after a failure, R, might be substantially lower than S. This is the case, e.g., when the existing setup does not need to be interrupted to repair the machine. Then it may very well be cheaper to resume after a failure at a cost of R rather than to let inventory drop back to zero before starting a new run at a cost of S. In this section we will study this abort/resume policy, again under the assumption of exponential failures.

We will analyze a family of stationary production policies with two parameters: a minimum lot size Q_1, and a maximum lot size $Q_1 + Q_2$. Hence a production run is terminated either when the total amount produced during the run hits $Q_1 + Q_2$ or when a breakdown occurs while the total production during the run exceeds Q_1. On the other hand, production is resumed immediately after a breakdown when the total production so far during the run is less than Q_1. A typical on-hand inventory sample path is depicted in Figure 10. In the Appendix we prove that for the exponential failure model this two-parameter family contains a policy with long-run average cost not exceeding the long-run average cost of the best stationary policy.

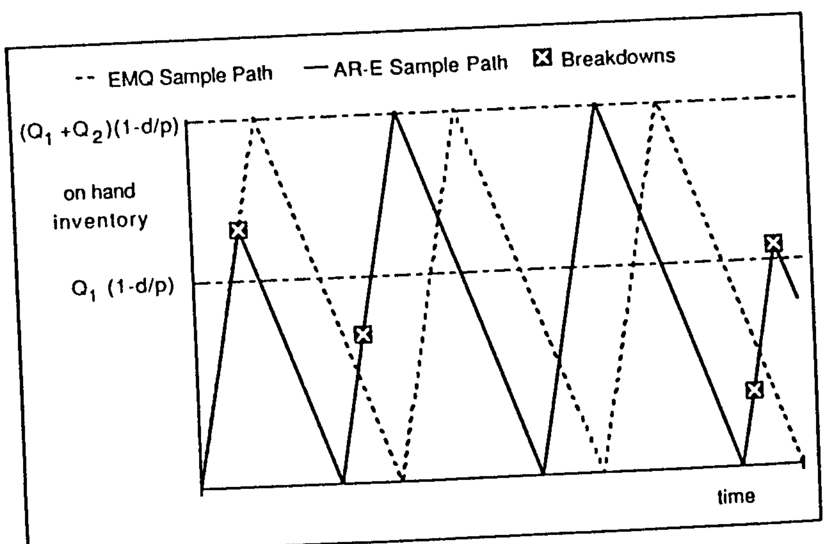

FIGURE 10. On-Hand Inventory Sample Paths for EMQ Model and AR-E Model.

We will next derive an expression for $\bar{C}(Q_1, Q_2)$, the long-run average total cost as a function of the decision variables Q_1 and Q_2. Since we have exponential failures, the long-run average corrective maintenance costs are again easily seen to be equal to $(d\lambda/p)M$, independent of the values of the decision variables. Similar to the analysis in §2 we find

$$\bar{C}(Q_1, Q_2) = \frac{d\lambda}{p} M + \frac{E[\text{setup, resumption and holding cost per cycle}]}{E[\text{duration of a cycle}]}$$

$$= \frac{d\lambda}{p} M + \frac{S + \frac{\lambda Q_1}{p} R + \frac{1}{2} h(p-d) \frac{p}{d} E[\tau^2]}{\frac{p}{d} E[\tau]}, \tag{4.1}$$

where τ = the amount of time spent producing during a cycle. The probability distribution of τ is given by:

$$\Pr\left\{\tau = \frac{Q_1 + Q_2}{p}\right\} = e^{-\lambda Q_2/p},$$

$$\Pr\left\{\tau \le \frac{Q_1}{p} + t\right\} = \begin{cases} 1 - e^{-\lambda t} & \text{if} \quad 0 \le t < \frac{Q_2}{p}, \\ 0 & \text{if} \quad t < 0. \end{cases}$$

We thus find with some calculus

$$E[\tau] = \frac{Q_1 + Q_2}{p} e^{-\lambda Q_2/p} + \int_0^{Q_2/p} \left(\frac{Q_1}{p} + t\right) \lambda e^{-\lambda t} dt = \frac{1}{\lambda}\left(1 + \frac{\lambda Q_1}{p} - e^{-\lambda Q_2/p}\right),$$

$$E[\tau^2] = \left(\frac{Q_1 + Q_2}{p}\right)^2 e^{-\lambda Q_2/p} + \int_0^{Q_2/p} \left(\frac{Q_1}{p} + t\right)^2 \lambda e^{-\lambda t} dt$$

$$= \frac{1}{\lambda^2}\left(\left(\frac{\lambda Q_1}{p}\right)^2 + 2\left(1 + \frac{\lambda Q_1}{p}\right)(1 - e^{-\lambda Q_2/p}) - 2\frac{\lambda Q_2}{p} e^{-\lambda Q_2/p}\right),$$

and finally

$$\bar{C}(Q_1, Q_2) = \frac{d\lambda}{p} M + \frac{S + \frac{\lambda Q_1}{p} R + \frac{hp(p-d)}{2\lambda^2 d}\left(\left(\frac{\lambda Q_1}{p}\right)^2 + 2\left(1 + \frac{\lambda Q_1}{p}\right)(1 - e^{-\lambda Q_2/p}) - 2\frac{\lambda Q_2}{p} e^{-\lambda Q_2/p}\right)}{\frac{p}{d\lambda}\left(1 + \frac{\lambda Q_1}{p} - e^{-\lambda Q_2/p}\right)}. \tag{4.2}$$

To simplify notation in what follows, let $z_i = \lambda Q_i/p$, $i = 1, 2$, and define

$$a = \lambda^2 dS/hp(p-d), \qquad k = R/S.$$

Then we can focus our attention for optimization purposes on the function

$$\gamma_{a,k}(z_1, z_2) = \frac{a(1 + kz_1) + \frac{1}{2}z_1^2 + (1 + z_1)(1 - e^{-z_2}) - z_2 e^{-z_2}}{1 + z_1 - e^{-z_2}} \tag{4.3}$$

since

$$\bar{C}(Q_1, Q_2) = \frac{d\lambda}{p} M + \frac{h(p-d)}{\lambda} \gamma_{a,k}\left(\frac{\lambda Q_1}{p}, \frac{\lambda Q_2}{p}\right).$$

Note that $a > 0$ and $0 \leq k \leq 1$. The case $k = 0$ corresponds to zero resumption costs and hence it will be optimal to set Q_1 = EMQ, $Q_2 = 0$, i.e., follow the traditional EMQ lot sizing policy. In case $k = 1$ the resumption cost is equal to the normal setup cost. Intuitively this means that we should never resume after a failure, and the policy $Q_1 = 0$, $Q_2 = Q^*$ from the previous section is optimal. These results follow rigorously from the next property that gives the optimal policy values for the general model.

Property 4.1. The function $\gamma_{a,k}(z_1, z_2)$ is minimized by (z_1^*, z_2^*) where z_2^* is the *unique* solution of the equation in x:

$$x + e^{-x} = 1 + ak, \quad \text{and} \tag{4.4}$$

$$z_1^* = -(z_2^* - ak) + \sqrt{(z_2^* - ak)^2 + 2a(1 - k)}\,. \tag{4.5}$$

PROOF. We will first show that (z_1^*, z_2^*) is the unique stationary point of $\gamma_{a,k}$. Taking partial derivatives yields

$$\frac{\partial}{\partial z_1}\gamma_{a,k}(z_1, z_2) = \frac{1 + ak + z_1 - e^{-z_2} - \gamma_{a,k}(z_1, z_2)}{1 + z_1 - e^{-z_2}}, \tag{4.6}$$

$$\frac{\partial}{\partial z_2}\gamma_{a,k}(z_1, z_2) = (z_1 + z_2 - \gamma_{a,k}(z_1, z_2))\frac{e^{-z_2}}{1 + z_1 - e^{-z_2}}. \tag{4.7}$$

After setting the partial derivatives equal to zero and solving we obtain (4.4) and (4.5) as the unique solution in the nonnegative orthant. One can check the second-order conditions and verify that (z_1^*, z_2^*) is a local minimum, but since $\gamma_{a,k}$ is not convex it is more complicated to show that (z_1^*, z_2^*) is also a global minimum.

Since $\gamma_{a,k}(z_1, z_2) \to \infty$ as $(z_1, z_2) \to (0, 0)$, there exists an $\epsilon > 0$ such that $\gamma_{a,k}(z_1, z_2) > \gamma_{a,k}(z_1^*, z_2^*)$ when $0 \leq z_1 + z_2 \leq \epsilon$. Using (4.6) and (4.7) it is also not hard to show that

$$\frac{\partial}{\partial z_1}\gamma_{a,k}(z_1, z_2) \geq 0 \qquad \text{for all} \qquad z_1 \geq 3(a + 1)$$

and that

$$\frac{\partial}{\partial z_2}\gamma_{a,k}(z_1, z_2) \geq 0 \qquad \text{for all} \qquad z_2 \geq a + 1.$$

A minimizer of $\gamma_{a,k}$ over the set $D = \{(z_1, z_2) | 0 \leq z_1 \leq 3(a + 1), 0 \leq z_2 \leq a + 1, z_1 + z_2 \geq \epsilon\}$ exists, since $\gamma_{a,k}$ is continuous over D and D is compact. This minimizer also minimizes $\gamma_{a,k}$ over the entire nonnegative orthant. From the above it is also clear that the minimizer *cannot* be on one of the boundaries corresponding to $z_1 + z_2 = \epsilon$, $z_1 = 3(a + 1)$, or $z_2 = a + 1$. We conclude the proof by showing that if the global minimizer is on one of the other two boundaries it satisfies (4.4) and (4.5).

Suppose $(0, \bar{z}_2)$ is a minimum of $\gamma_{a,k}$ on D. Then

$$\frac{\partial}{\partial z_1}\gamma_{a,k}(0, \bar{z}_2) \geq 0,$$

and hence we obtain with (4.6):

$$1 + ak - e^{-\bar{z}_2} \geq \gamma_{a,k}(0, \bar{z}_2), \qquad \text{and} \tag{4.8}$$

$$\frac{\partial}{\partial z_2}\gamma_{a,k}(0, \bar{z}_2) = 0,$$

so (with (4.7))

$$\bar{z}_2 = \gamma_{a,k}(0, \bar{z}_2). \tag{4.9}$$

It is not hard to check that (4.9) implies $\bar{z}_2 = 1 + a - e^{-\bar{z}_2}$, and so (4.8) and (4.9) can only be satisfied simultaneously when $k = 1$, and in this case $(0, \bar{z}_2) = (z_1^*, z_2^*)$. The case of a minimum $(\bar{z}_1, 0)$ can similarly be shown to occur only when $k = 0$, and then $(\bar{z}_1, 0) = (z_1^*, z_2^*)$. QED

It is informative to consider several special cases of Property 4.1. When $k = 0$ (i.e., $R = 0$), there is no reason to abort a production run after a failure since resumption is free. Hence the ordinary EMQ model applies, and with (4.4) and (4.5) we see that the optimal policy $(z_1^*, z_2^*) = (\sqrt{2a}, 0)$ indeed coincides with the EMQ policy. When $k = 1$ (i.e., $R = S$), it is just as expensive to resume a run after a failure as it is to start a new run. One would expect that in that case it is optimal never to resume, and this is indeed what (4.5) implies in this case. Moreover, since there is no resumption, we are now back in the no-resumption model of §3, and $\lambda z_1^*/p$ indeed gives the optimal value for Q_1^* (compare (3.3) and (4.4)).

Numerically calculating the optimal z_1^* and z_2^* from (4.4) and (4.5) is not hard, but it is of some value to have a simple, closed form approximation. We propose the following approximation for (z_1^*, z_2^*):

$$z_1^h = \sqrt{2a} - \sqrt{2ak}, \qquad z_2^h = \sqrt{2ak}. \tag{4.10}$$

The value for z_2^h is obtained by replacing the left-hand side in (4.4) by $1 + \frac{1}{2}x^2$, the first few terms of its power series expansion. The value for z_1^h is then set so that the maximum lot size equals the EMQ. This heuristic has several desirable properties: when $k = 0$ (i.e., $R = 0$), the heuristic policy coincides with the EMQ policy and is thus optimal. When $k = 1$ (i.e., $R = S$), we have $z_1^h = 0$, $z_2^h = \sqrt{2a}$, which is the approximately optimal policy shown in §3 to have a cost error of less than 2%. Numerical investigations show that the maximum cost error of the approximation (4.10) is less than 4.2%, and occurs when $a \approx 50.23$, $k \approx 0.5925$.

Note 1. The average cost per unit time of the optimal NR-E policy of §3 can be up to a factor $1/k$ worse than that of the optimal abort/resume policy. Note that by (4.4) and (4.5) we have $z_2^* \leq 1 + ak$ and $z_1^* \leq \sqrt{2a}$. Using that $\gamma_{a,k}(0, z^*(a)) = z^*(a) \geq a$ (with $z^*(a)$ defined by (3.6)) we find

$$\frac{\gamma_{a,k}(0, z^*(a))}{\gamma_{a,k}(z_1^*, z_2^*)} \geq \frac{a}{z_1^* + z_2^*} \geq \frac{a}{1 + ak + \sqrt{2a}} \to \frac{1}{k} \quad \text{as } a \to \infty. \tag{4.11}$$

It is therefore important to consider the abort/resume decision carefully when $k \downarrow 0$ and $a \to \infty$. The latter increases with the machine utilization d/p, the ratio of setup over holding cost S/h, and with the square of the failure rate.

Note 2. Numerical investigations show that the cost of following the EMQ policy (i.e., using $z_1 = \sqrt{2a}$, $z_2 = 0$) can be up to 36.9% higher than the cost of the optimal AR-E policy. The worst case occurs for $k = 1$ and $a \approx 3.32$. When $k = 0.5$ and $a \approx 5.02$ the EMQ policy cost is still 14.7% greater than the AR-E policy cost.

5. Conclusions

Understanding the impact of machine breakdowns on the production lot sizing decisions poses some counterintuitive issues due to their conflicting influence on the various cost function components. For example, as machine reliability goes down, one would expect an increase in the optimal lot size in order to compensate for production runs aborted by breakdowns. On the other hand, the increase in breakdown rate also pushes down the average actual lot size produced and the average on hand inventories, so the net effect is hard to predict without a detailed analysis.

In this paper we present two novel extensions to the EMQ model. These extensions are aimed at incorporating stochastic machine breakdowns and deal with analyzing the optimal lot size and the associated reorder policy. The first extension assumes that a production run is aborted at a machine breakdown, and a new run is to be started when all on hand inventory is depleted. We derive a cost expression for a general failure distribution. For tractability reasons we then focused on the common case of exponential failures, resulting in the NR-E policy. In the second extension (AR-E) a production run interrupted by a breakdown can be either resumed immediately or aborted until all on-hand inventory is depleted. The decision to abort or resume depends on the amount of on-hand inventory and on the setup and resume costs. This policy is particularly applicable when there is no need to repeat the entire setup at each breakdown. We prove that the optimal stationary policy is of a threshold control type, and derive both exact and approximate optimal control parameters for both policies. Uniqueness is proved for the exact solutions, and sharp cost bounds are provided for the approximate solutions. In addition we present various structural properties and operational insights. It is shown that the long-run average corrective maintenance cost is independent of the lot size used; however, the failure rate significantly affects the optimal lot size. For both policies, the classical EMQ lot size serves as the limiting case when the failure rate approaches zero. Moreover, using the EMQ lot size under the NR-E policy will result in an average cost increment of 2% or less above the optimal average cost. It is shown that when the resume cost approaches the setup cost from below, the optimal AR-E policy approaches the optimal NR-E policy. In addition, the ratio of average cost per unit time of NR-E to AR-E is shown to be bounded by the ratio of setup to resume cost.

The results presented here have several operations management implications. One can use these models in order to determine production schedule, WIP level and buffer capacity requirements as a function of the machine reliability and throughput rate. The desired effort allocated to corrective and preventive maintenance activities can be assessed if one knows the relationship between the effort and the failure rate. Equipment selection is another issue to be evaluated as different machines are characterized not only by different processing rates and setup costs but also by different failure rates. Several extensions are noteworthy in this context. The case of multiple product types with cyclic production requires special attention as at each breakdown one may decide to either resume the production run or switch to another product. Other possible extensions include randomly distributed processing and repair rates, and the incorporation of random demand.[1]

[1] The authors would like to thank Professor Uday S. Karmarkar from the University of Rochester for stimulating discussions and useful comments. The authors also express their gratitude to the referees for constructive comments and suggestions.

Appendix A: Optimality Properties of the Threshold Policy

In this appendix we will show that the optimal threshold type policy of §4 is the optimal stationary policy for problem AR-E. To do this, we prove three properties that successively narrow down the set of stationary policies containing an optimal one.

Since policies with finite long-run average cost exist (e.g., the threshold policy of §4), we can restrict ourselves to policies that with probability 1 perform only a finite number of setups in any finite time interval. The following property shows that any policy that performs one or more major setups while inventory is strictly positive is dominated by a policy that always waits until inventory is depleted before doing a major setup. Note that Property A1 does not require an assumption of stationarity! Note also that Property A1 implies the optimality of a single number control structure for problem (NR-E).

Property A1. For any policy π that does a major setup at time T while inventory is strictly positive, there is a policy π' that does no major setups while inventory is strictly positive, such that, for all $t \geq T$, the total cost incurred up to time t under policy π' is strictly less than the total cost incurred up to time t under policy π.

PROOF. For any policy π, let $I_\pi(t)$ denote the inventory at time t, and let $C_\pi(t)$ be the total cost incurred up to time t. Let π be an arbitrary policy, and suppose that under π a major setup occurs at time T and that $I_\pi(T) > 0$. We construct a new policy π' as follows: up to time T, π' makes the same decisions as π. At time T, π' does not do a setup. Instead, under π' a major setup is not performed until all inventory is depleted at time $T + I_\pi(T)/d$. After that time, policy π' keeps on producing until $I_\pi(t) = I_{\pi'}(t)$, doing minor setups if required. It is clear that both policies meet the demand constraint. Whether or not the sample paths ever rejoin, we get that $C_{\pi'}(t) < C_\pi(t)$ for all $t \geq T$, since under policy π a (major or minor) setup is needed for every minor setup that takes place under policy π', the number of breakdowns up to time t is at least as great under policy π as under policy π' (because cumulative production is at least as great), and there is more inventory under policy π than under policy π' for at least a time interval of length $I_\pi(T)/d > 0$. The theorem now follows with induction. QED

From now on we will restrict ourselves to *stationary* policies that never perform a major setup while inventory is positive. We will call such policies stationary "return-to-zero" policies. Under any stationary return-to-zero policy, the system behavior can be described with a regenerative process with regeneration points $\{t \mid I(t) = 0\}$. It follows that the long-run average cost of corrective maintenance is equal to $d\lambda M/p$ with probability 1 for any stationary return-to-zero policy. Hence in what follows we can safely ignore corrective maintenance costs. Any *stationary return-to-zero* policy can be specified by a pair (J, A), where J is the maximum inventory where a production run will be terminated and $A \subset [0, J] = \{I \mid 0 \leq I \leq J\}$ is the "abort" set, i.e., the set of inventories such that a restart will not be performed when a breakdown occurs at time t and $I(t) \in A$. We will assume that the set A is measurable. To simplify notation, let $\xi = \lambda/(p-d)$. For a policy $\pi = (J, A)$ define:

$$a'_\pi(I) = \begin{cases} 1 & \text{if} \quad I \in A, \\ 0 & \text{if} \quad I \notin A, \end{cases}$$

$$a_\pi(I) = \int_{x=0}^{I} a'_\pi(x)dx \qquad (0 \leq I \leq J).$$

$$p_\pi(I) = \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \Pr\{\text{the maximum inventory during a cycle} \in (I - \epsilon/2, I + \epsilon/2) \mid \pi\}$$

$$= \xi a'_\pi(I) \exp\{-\xi a_\pi(I)\} \qquad \text{a.e.} \qquad (0 \leq I \leq J).$$

$$P_\pi(I) = \Pr\{\text{the maximum inventory during a cycle} = I \mid \pi\}$$

$$= \begin{cases} \exp\{-\xi a_\pi(J)\} & \text{if} \quad I = J, \\ 0 & \text{if} \quad I \neq J. \end{cases}$$

$$\tau(I) = \text{the length of a cycle with maximum inventory } I = \frac{p}{d(p-d)} I.$$

$$\gamma(I) = \text{the holding cost incurred during a cycle with maximum inventory } I = \frac{hp}{2d(p-d)} I^2.$$

$$\rho_\pi(I) = E[\text{restart costs incurred during a cycle with maximum inventory } I \mid \pi] = \xi R(I - a_\pi(I)).$$

$$k_\pi = E[\text{total cost incurred during a cycle} \mid \pi].$$

$$t_\pi = E[\text{length of a cycle} \mid \pi].$$

Note that $P_\pi(J) + \int_0^J p_\pi(I)dI = 1$, and that γ and τ don't depend on the policy π.

We find that the long-run average cost of policy π (excluding breakdown costs) can be written as

$$\bar{C}_\pi = \frac{k_\pi}{t_\pi} = \frac{S + \int_0^J p_\pi(I)[\gamma(I) + \rho_\pi(I)]dI + P_\pi(J)[\gamma(J) + \rho_\pi(J)]}{\int_0^J p_\pi(I)\tau(I)dI + P_\pi(J)\tau(J)}.$$

Finally, let $v_\pi(I) = \gamma(I) - \bar{C}_\pi \tau(I)$. Note that since $\bar{C}_\pi$ is fixed, $v_\pi(I)$ is a convex and quadratic function of I that assumes its minimum at $I = \bar{C}_\pi/h$.

The next property gives an optimality condition for the maximal inventory level J. In fact, it generalizes optimality condition (4.7).

Property A2. Let $\pi = (J, A)$ be an arbitrary stationary return-to-zero policy. If $hJ \neq \bar{C}_\pi$, then there always exists a stationary return-to-zero policy $\pi' = (J', A')$ such that $hJ' = \bar{C}_{\pi'}$ and $\bar{C}_{\pi'} < \bar{C}_\pi$.

PROOF. We will construct a sequence of policies $\pi_0 = \pi$, $\pi_1 = (J_1, A_1)$, $\pi_2 = (J_2, A_2), \ldots$, such that

$$\bar{C}_{\pi_n} = hJ_{n+1} \qquad (n \geq 0), \tag{A-1}$$

$$\bar{C}_{\pi_n} > \bar{C}_{\pi_{n+1}} \qquad (n \geq 0), \quad \text{and} \tag{A-2}$$

$$A_n = A_1 \cap [0, J_n] \qquad (n \geq 1). \tag{A-3}$$

Note from (A-1)–(A-3) that $\{J_n \mid n \geq 1\}$ is a decreasing sequence bounded from below by 0, so that $J' = \lim_{n\to\infty} J_n$ exists. Now let $\pi' = (J', A')$, where $A' = A_1 \cap [0, J']$. The proposition follows by noting that $\bar{C}_{\pi'} = \lim_{n\to\infty} \bar{C}_{\pi_n} = h \cdot \lim_{n\to\infty} J_n = hJ'$, and $\bar{C}_{\pi'} < \bar{C}_{\pi_0} = \bar{C}_\pi$.

To construct the sequence of policies, first consider the case that $n = 0$ and $hJ = hJ_0 < \bar{C}_\pi$. For this case, define $J_1 = \bar{C}_\pi/h$ and $A_1 = A \cup [J, J_1]$. Using $p_{\pi_1}(I) = p_\pi(I)$ $(0 \le I < J)$, $\rho_{\pi_1}(I) = \rho_\pi(I)$ $(0 \le I \le J)$, $\rho_{\pi_1}(I) = \rho_{\pi_1}(J)$ $(J \le I \le J_1)$, and $\int_{I=J}^{J_1} p_{\pi_1}(I)dI = P_\pi(J) - P_{\pi_1}(J_1)$, we find

$$(k_\pi - \bar{C}_\pi t_\pi) - (k_{\pi_1} - \bar{C}_\pi t_{\pi_1}) = P_\pi(J)[v_\pi(J) + \rho_\pi(J)] + \int_{I=0}^{J} p_\pi(I)[v_\pi(I) + \rho_\pi(I)]dI$$

$$- P_{\pi_1}(J_1)[v_\pi(J_1) + \rho_{\pi_1}(J_1)] - \int_{I=0}^{J_1} p_{\pi_1}(I)[v_\pi(I) + \rho_{\pi_1}(I)]dI$$

$$= P_{\pi_1}(J_1)[v_\pi(J) - v_\pi(J_1)] + \int_{I=J}^{J_1} p_{\pi_1}(I)[v_\pi(J) - v_\pi(I)]dI$$

$$+ P_{\pi_1}(J_1)[\rho_\pi(J) - \rho_{\pi_1}(J_1)] + \int_{I=J}^{J_1} p_{\pi_1}(I)[\rho_\pi(J) - \rho_{\pi_1}(I)]dI > 0,$$

since the first two terms are strictly positive and the last two terms equal zero. Using $k_\pi - \bar{C}_\pi t_\pi = 0$ we conclude that $\bar{C}_{\pi_0} = \bar{C}_\pi > k_{\pi_1}/t_{\pi_1} = \bar{C}_{\pi_1}$.

Next, consider the case that $hJ_n > \bar{C}_{\pi_n}$ for some $n \ge 0$. Define $J_{n+1} = \bar{C}_{\pi_n}/h$, and $A_{n+1} = A \cap [0, J_n]$. Using that $p_{\pi_{n+1}}(I) = p_{\pi_n}(I)$ $(0 \le I < J_{n+1})$, $\rho_{\pi_{n+1}}(I) = \rho_{\pi_n}(I)$ $(0 \le I \le J_{n+1})$, and $\int_{I=J_{n+1}}^{J_n} p_{\pi_n}(I)dI = P_{\pi_{n+1}}(J_{n+1}) - P_{\pi_n}(J_n)$, we find

$$(k_{\pi_n} - \bar{C}_{\pi_n} t_{\pi_n}) - (k_{\pi_{n+1}} - \bar{C}_{\pi_n} t_{\pi_{n+1}}) = P_{\pi_n}(J_n)[v_{\pi_n}(J_n) + \rho_{\pi_n}(J_n)] + \int_{I=0}^{J_n} p_{\pi_n}(I)[v_{\pi_n}(I) + \rho_{\pi_n}(I)]dI$$

$$- P_{\pi_{n+1}}(J_{n+1})[v_{\pi_n}(J_{n+1}) + \rho_{\pi_{n+1}}(J_{n+1})] - \int_{I=0}^{J_{n+1}} p_{\pi_{n+1}}(I)[v_{\pi_n}(I) + \rho\pi_{n+1}(I)]dI$$

$$= P_{\pi_n}(J_n)[v_{\pi_n}(J_n) - v_{\pi_n}(J_{n+1})] + \int_{I=J_{n+1}}^{J_n} p_{\pi_n}(I)[v_{\pi_n}(I) - v_{\pi_n}(J_{n+1})]dI$$

$$+ P_{\pi_n}(J_n)[\rho_{\pi_n}(J_n) - \rho_{\pi_{n+1}}(J_{n+1})] + \int_{I=J_{n+1}}^{J_n} p_{\pi_n}(I)[\rho_{\pi_n}(I) - \rho_{\pi_{n+1}}(J_{n+1})]dI > 0,$$

since all terms are nonnegative and at least the first strictly positive. Like in the previous case we conclude that $\bar{C}_{\pi_n} > k_{\pi_{n+1}}/t_{\pi_{n+1}} = \bar{C}_{\pi_{n+1}}$.

Conditions (A-1)–(A-3) and $hJ_n > \bar{C}_{\pi_n}$ $(n \ge 1)$ now follow easily with induction. QED

The third property shows that among the policies satisfying the conditions of Properties A1 and A2 there is an optimal policy with a convex abort set A that contains the maximal inventory level J.

Property A3. Let $\pi = [J, A]$ be a stationary return-to-zero policy satisfying $J = \bar{C}_\pi/h$. Define $A' = \{I \mid J - a_\pi(J) \le I \le J\}$, and let $\pi' = [J, A']$. Then $\bar{C}_{\pi'} \le \bar{C}_\pi$.

PROOF. Note that $a_\pi(J) = a_{\pi'}(J)$, $\rho_{\pi'}(J) = \rho_\pi(J)$, $P_{\pi'}(J) = P_\pi(J)$, and

$$\int_{I=0}^{J} p_\pi(I)\rho_\pi(I)dI = 1 - \exp\{-\xi a_\pi(J)\} = \int_{I=0}^{J} p_{\pi'}(I)\rho_{\pi'}(I)dI.$$

It follows that

$$(k_\pi - \bar{C}_\pi t_\pi) - (k_{\pi'} - \bar{C}_\pi t_{\pi'}) = P_\pi(J)[v_\pi(J) + \rho_\pi(J)] + \int_{I=0}^{J} p_\pi(I)[v_\pi(I) + \rho_\pi(I)]dI$$

$$- P_{\pi'}(J)[v_\pi(J) + \rho_{\pi'}(J)] - \int_{I=0}^{J} p_{\pi'}(I)[v_\pi(I) + \rho_{\pi'}(I)]dI$$

$$= \int_{I=0}^{J} [p_\pi(I) - p_{\pi'}(I)]v_\pi(I)dI$$

$$= \int_{I=0}^{J} \xi a'_\pi(I) \exp\{-\xi a_\pi(I)\} v_\pi(I)dI$$

$$- \int_{I=0}^{J} \xi a'_{\pi'}(I) \exp\{-\xi a_{\pi'}(I)\} v_\pi(I)dI. \qquad \text{(A-4)}$$

In order to show that the RHS of (A-4) is nonnegative, we consider the following problem:

$$(P) \qquad \underset{\alpha}{\text{minimize}} \quad \int_{I=0}^{J} \xi\alpha'(I) \exp\{-\xi\alpha(I)\} v_\pi(I)dI$$

subject to $\alpha(0) = 0$, $\alpha(J) = a_\pi(J)$, $0 \le \alpha'(I) \le 1$ a.e.,

and $\alpha(\cdot)$ is continuous and measurable on $(0, J)$.

The optimal solution to problem (P) is a_τ, since $J = \bar{C}_\tau/h$ and by definition $v_\tau(\cdot)$ is decreasing on $[0, \bar{C}_\tau/h]$. Also by definition, a_τ is a feasible solution to (P). This is sufficient for the RHS of (A-4) to be nonnegative. QED

Together, Properties A1–A3 imply

THEOREM A4. *Let π be any stationary policy for* AR-E, *and let π^* denote the optimal threshold policy of* §4. *Then the long-run average costs per unit time must satisfy $\bar{C}_{\pi^*} \leq \bar{C}_\pi$.*

References

AXSÄTER, S. AND S. E. ELMAGHRABY, "A Note on EMQ under Learning and Forgetting," *AIIE Trans.*, 13 (1981), 86–90.

BANERJEE, A., "A Joint Economic-Lot-Size Model for Purchaser and Vendor," *Decision Sci.*, 17 (1986), 292–311.

BARLOW, R. E. AND F. PROSCHAN, *Mathematical Theory of Reliability*, Wiley, New York, 1965.

BILLINGTON, P. J., "The Classic Economic Production Quantity Model with Setup Cost as Function of Capital Expenditure," *Decision Sci.*, 18 (1987), 24–42.

———, "Quality Improvement Through Reduced Lot Sizes in the Economic Order Quantity Model," *15th DSI Proc.*, 1986, 1159–1161.

BROWN, R. M., T. E. CONINE AND M. TAMARKIN, "A Note on Holding Costs and Lot-Size Errors," *Decision Sci.*, 17 (1986), 603–608.

ERLENKOTTER, D., "An Early Classic Misplaced: Ford W. Harris's Economic Order Quantity Model of 1915," *Management Sci.*, 35 (1989), 898–900.

GOYAL, S. K., "An Inventory Model for a Product for which Purchase Price Fluctuates," *New Zealand Oper. Res.*, 30, 1 (1979), 43–53.

GROENEVELT, H., L. PINTELON AND A. SEIDMANN, "Production Batching with Machine Breakdowns and Safety Stocks," *Oper. Res.*, 40 (1992) (to appear).

HARRIS, F. W., "What Quantity to Make at Once," *The Library for Factory Management. Vol. V. Operation and Costs*, A. W. Shaw, Chicago, 1915, 47–52.

JESSE, R. R., A. MITRA AND J. F. COS, "EOQ Formula: Is It Valid under Inflationary Conditions?" *Decision Sci.*, 14 (1983), 370–374.

KARMARKAR, U. S., "Lot Sizes, Lead Times and In-Process Inventories," *Management Sci.*, 33 (1987), 409–418.

KUZDRALL, P. AND R. BRITNEY, "Total Setup Lot Sizing with Quantity Discounts," *Decision Sci.*, 13, 1 (1982), 101–112.

LEE, H. L. AND M. J. ROSENBLATT, "Simultaneous Determination of Production Cycle and Inspection Schedules in a Production System," *Management Sci.*, 33, 9 (1987), 1125–1136.

——— AND ———, "Optimal Inspection and Ordering Policies for Products with Imperfect Quality," *IIE Trans.*, 17, 3 (1985), 284–289.

LEV, B., H. J. WEISS AND A. L. SOYSTER, "Inventory Models with Finite Horizons and Price Changes," *Oper. Res. Quart.*, 30, 1 (1979), 43–53.

———, ——— AND ———, "Optimal Order Policies when Anticipating Parameter Changes in EOQ Systems," *Naval Res. Logist. Quart.*, 28, 2 (1981), 267–279.

MCLEAVEY, D. W. AND S. L. NARASIMHAN, *Production and Inventory Control*, Allyn and Bacon, Boston, 1985.

MURDESHWAR, T. M. AND Y. S. SATHE, "Some Aspects of Lot Size Models with Two Levels of Storage," *Opsearch*, 22, 4 (1985), 255–262.

MUTH, E. J. AND K. SPREMANN, "Learning Effects in Economic Lotsizing," *Management Sci.*, 29 (1983), 264–269.

NADDOR, E., *Inventory Systems*, R. E. Krieger, Malabar, FL, 1984.

NAHMIAS, STEVEN, *Production and Operations Analysis*, Irwin, Homewood, IL, 1989.

OSTERYOUNG, J. S., D. E. MCCARTY AND W. J. REINHART, "Use of EOQ Model for Inventory Analysis," *Production and Inventory Management J.*, (3rd Q, 1986), 39–45.

PORTEUS, E. L., "Investing in New Parameter Values in the Discounted EOQ Model," *Naval Res. Logist. Quart.*, 33 (1985), 39–48.

———, "Optimal Lot Sizing, Process Quality Improvement and Setup Cost Reduction," *Oper. Res.*, 34 (1986), 137–144.

RESH, M., M. FRIEDMAN AND L. C. BARBOSA, "On a General Solution of the Deterministic Lot Size Problem with Time Proportional Demand," *Oper. Res.*, 24 (1976), 718–725.

RITCHIE, E. AND A. TSADO, "The Penalties of Using the EOQ: A Comparison of Lot-Sizing Rules for Linearly Increasing Demand," *Production and Inventory Management J.*, (1st Q, 1986), 12–19.

ROSENBLATT, M. J. AND H. L. LEE, "Economic Production Cycles with Imperfect Production Processes," *IIE Trans.*, 18 (1986), 48–55.

ROSS, S. M., *Introduction to Probability Models*, Academic Press, New York, 1981.

SCHWALLER, R. L., "EOQ under Inspection Costs," *Production and Inventory Management J.*, (3rd Q, 1988), 22–24.

SCHWARZ, L. B., "A Note on the Near Optimality of "5-EOQ's" Worth Forecast Horizons," *Oper. Res.*, 25 (1977), 533–536.

SETHI, S. P., "A Quantity Discount Model with Disposals," *Internat. J. AIIE Trans.*, 22, 1 (1984), 31–39.

SHAH, Y. K., "An Order-Level Lot-Size Model for Deteriorating Items," *AIIE Trans.*, 9 (1977), 108–112.

SILVER, E. A. AND R. PETERSON, *Decision Systems for Inventory Management and Production Planning*, Wiley, New York, 1985.

——— AND H. C. MEAL, "A Simple Modification of the EOQ for the Case of a Varying Demand Rate," *Production and Inventory Management J.*, 10, 4 (1969), 52–65.

SNYDER, R. D., "The Classical Economic Order Quantity Formula," *Oper. Res. Quart.*, 24 (1977), 125–127.

WAGNER, H. M. AND T. M. WHITIN, "Dynamic Version of the Economic Lot Size Model," *Naval Res. Logist. Quart.*, 5 (1958), 89–96.

YANO, C. A. AND H. L. LEE, "Lot Sizing with Random Yields: A Review," Technical Report 89-16, Dept. of Ind. and Oper. Eng., Univ. of Michigan, Ann Arbor, 1989.

ZIPKIN, PAUL, "Models for Design and Control of Stochastic, Multi-Item Batch Production Systems," *Oper. Res.*, 34 (1986), 91–104.