



Investing in Quality Under Autonomous and Induced Learning

Dogan A. Serel , Maqbool Dada , Herbert Moskowitz & Robert D. Plante

To cite this article: Dogan A. Serel , Maqbool Dada , Herbert Moskowitz & Robert D. Plante (2003) Investing in Quality Under Autonomous and Induced Learning, IIE Transactions, 35:6, 545-555, DOI: [10.1080/07408170304415](https://doi.org/10.1080/07408170304415)

To link to this article: <http://dx.doi.org/10.1080/07408170304415>



Published online: 29 Oct 2010.



Submit your article to this journal [↗](#)



Article views: 32



View related articles [↗](#)



Citing articles: 17 View citing articles [↗](#)

Investing in quality under autonomous and induced learning

DOGAN A. SEREL¹, MAQBOOL DADA^{2,*}, HERBERT MOSKOWITZ² and ROBERT D. PLANTE²

¹Faculty of Business Administration, Bilkent University, 06533 Bilkent, Ankara, Turkey

²Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907-1310, USA

E-mail: dada@mgmt.purdue.edu

Received April 2000 and accepted June 2002

The reduction of variability in product performance characteristics is an important focus of quality improvement programs. Learning is intrinsically linked to process improvement and can assume two forms: (i) autonomous learning; and (ii) induced learning. The former is experientially-based, while the latter is a result of deliberate managerial action. Our involvement in quality and capacity planning with several major corporations in different industries suggested that it would be instructive to devise a model that would prescribe an optimal combination of autonomous and induced learning over time to maximize process improvement. We thus propose such a model to investigate the optimal quality improvement path for a company given that quality costs depend on both autonomous and induced types of learning experienced on a number of quality characteristics. Several properties of an optimal investment path are developed for this problem. For example, it is shown that decisions maximizing short-term gains may actually lead to suboptimal resource utilization decisions when total costs associated with a longer planning horizon are taken into account. Numerical examples are used to assess the sensitivity of the optimal investment plan with respect to changes in several model parameters.

1. Introduction

A significant portion of quality related costs is incurred due to variation in process output. Thus, manufacturing companies strive to continually improve processes via reduction in process variation. An important mechanism for reducing process variation is for the manufacturer to commit to a quality improvement philosophy and strategy that fosters continuous process *learning* and improvement.

When considering the relationship between learning and process improvement, it is useful to view organizational learning as being either autonomous or induced (Levy, 1965). *Autonomous learning* is associated with learning by doing and captures the efficiency gained through repetitive implementation of tasks and experience. *Induced learning* is generated by conscious managerial or engineering actions that improve the efficiency of the system through changes in the technology, the underlying processes, and physical or human capital. Some specific examples of such actions are engineering design changes, and personnel training programs (Adler and Clark, 1991).

Our interest was motivated by our involvement in improvement activities with several large companies. One was a major pharmaceutical corporation faced with the problem of capacity planning. It was found that process learning in this organization's manufacturing operations resulted in

an annual increase in capacity of 15%. A second was a major consumer electronics manufacturer who undertook induced learning investments via six sigma type quality training using project execution teams at one of its manufacturing plants in Latin America that produced an average of 3000 units per day. These induced investments in quality training yielded “mature quality levels” in 2 weeks from product start-up (defined as at least 90% first-pass yield), resulting in increased productivity and reduced process costs. This was a 94% improvement in the time to achieve process maturity *vis-à-vis* the time to do so when improvements were based on autonomous learning, which was 9 months. Moreover, savings in manufacturing costs were greater than 11 million dollars over the 2-week period.

In developing its induced investment strategy, the consumer electronics manufacturer had to determine an appropriate sequence of investments in each of the three major stages of its manufacturing process: (i) Automatic Component Insertion (ACI); (ii) Manual Assembly (MA); and (iii) Soldering (S). The induced learning program was launched by holding a plant-wide 3-day active learning workshop that focused on statistical quality concepts and analysis tools. This was followed by pre-defined, targeted team-based projects to make specific improvements in each of the three major stages of the manufacturing process. For example, at ACI, defects such as incorrect lead length and epoxy contamination were identified, and, corrective actions determined and implemented. At MA, reversing parts was determined to be a major contributor to poor quality

*Corresponding author

and corrected. At S, too much or too little solder was identified as a major quality issue, and also corrected.

One of the challenges faced by both the above pharmaceutical and consumer electronics companies was determining what should be the optimal sequence of investments in autonomous and induced learning for process improvement and concomitant capacity increase to improve quality, reduce costs, and reduce cycle times, resulting in a capacity increase. Motivated, in part, by this decision problem, we investigate the dynamic behavior of manufacturing quality costs as a function of variance reducing investments which are realistically associated with multiple quality characteristics. In this context, the main goal is to select the most promising/beneficial areas for focusing quality improvement efforts, given the options available in each time period. Based on contemporary learning and quality cost theories, we develop a dynamic model of relating investment and learning analytical formulation that provides answers to questions such as: What is the optimal investment in learning path to minimize expected quality costs? Is there an easy-to-prescribe optimal investment policy that is robust under fairly general conditions? How are quality improvement decisions influenced by autonomous and induced types of learning? The multi-period model to be presented relies on the fundamental notion that improvements in quality are realized through gradual and continuous decreases in process variation over time.

1.1. Relationship to the learning literature

The link between the learning curve and quality improvement activities has been explored quite extensively. Fine (1986) developed a quality-based learning model in which the quality level, represented by economic conformance to tolerances, was a management-controlled decision variable. By dynamically changing the economic conformance level, management controls the cumulative production of conforming output, which determines the unit cost of production. Thus, a model based on quality-weighted volume replaces the well-known volume-based learning curve. Higher quality levels lead to higher percentages of conforming output as well as faster rates of reduction in unit production cost. Fine and Porteus (1989) studied a different quality improvement model that included only induced learning with stochastic rewards. Kini (1994) incorporated the influences of both good and defective items on the learning rate. Zangwill and Kantor (1998) described how various forms of the learning curve such as power and exponential functions can be treated in a unified manner. Moskowitz *et al.* (1997) and Plante (2000) formulated single-period models to determine target levels for quality improvement in the presence of induced learning.

Most learning models in the literature have commonly considered only autonomous learning, and explored optimal production policies that minimize production costs (Mazzola and McCardle, 1997). However, several recent

papers have addressed both autonomous and induced learning simultaneously. For example, Li and Rajagopalan (1998) differentiated between the “productivity knowledge” and “quality knowledge” gains resulting from learning efforts by building a model in which both autonomous and induced learning activities influence the changes in the accumulated levels of productivity and quality knowledge. Lapre *et al.* (2000) proposed a learning curve for the waste rate of a manufacturing process, which includes both autonomous and induced learning.

There have also been recent empirical studies on autonomous and induced learning. Ittner (1996) investigated the relationship between the expenditures on quality improvement activities and the costs associated with product defects. Mukherjee *et al.* (1998) proposed the following two main dimensions for the knowledge gained via quality improvement projects: (i) operational learning which refers to the acquisition of “know-how”; and, (ii) conceptual learning which is defined as the acquisition of “know-why”. By analyzing the quality improvement projects undertaken by a steel wire manufacturer, they attempted to assess the impact of these learning dimensions on the waste rate of the production process. Empirical research by Li and Rajagopalan (1997) concluded that quality improvement activities led to identifying inefficiencies in the production process, and such a knowledge gain resulted in increased productivity. Analyzing data pertaining to 12 manufacturing plants and consistent with our proposed modeling approach, Ittner *et al.* (2001) find that production quality was influenced by both autonomous and induced learning.

1.2. Rationale of our modeling approach

Building on the literature, we simultaneously; (i) consider multiple learning curves in a manufacturing environment; (ii) use the variability of quality characteristics as a performance metric (cf. Zangwill and Kantor, 1998); and (iii) quantify the quality-related costs based on the combined effect of these metrics. More specifically, we employ Taguchi’s “quality loss function” to estimate quality-related costs (Taguchi and Clausing, 1990). For this model, the ideal state of a quality characteristic that maximizes user satisfaction is called the target value (Kackar, 1985), and, all deviations from this target value incur some cost. Taguchi’s loss function thus implies that quality costs are incurred whenever the quality (performance) characteristic is not on its target, even if the product conforms to specifications. The traditional quality cost theory dating back to Juran (1951), on the other hand, does not consider it as a costly outcome when the performance of a product falls in the interval between the lower and upper specification limits. As compared to the traditional approach, Taguchi’s loss function places emphasis on reducing the variability of the performance characteristic as the key element of modern quality management practice. Since the performance characteristics are usually modeled as random variables, the overall quality

level of production is inversely related to performance variation around the target value. Thus higher variation implies higher manufacturing costs. Consequently, reducing variability in the performance characteristics will simultaneously reduce both quality costs to the users, and production costs to the manufacturer with concomitant reductions in scrap and rework costs, and improvements in productivity, cycle time, service call rates, etc. (Kackar, 1985).

Consistent with Taguchi's loss function approach and unlike that of Fine (1986), we relate the reductions in quality costs to decreases in process output variation over time. Process variation reduction is a continuous effort that is influenced by the accumulation of process knowledge (MacKay and Steiner, 1997). We draw from the learning literature to specify this relationship between the level of process variation at a specific point in time and the level of "learning" achieved up to that point. Rather than relating the cumulative production volume to the cost of production, our learning curve describes the relationship between process quality and elapsed time. Use of a learning curve as a function of time has been deployed in practice; e.g., Schneiderman (1988) and Stata (1989) report that this form of a learning curve, which we employ in our model, was actually being applied at Analog Devices.

Our work is an initial attempt to: (i) introduce Taguchi's loss function concept in a dynamic multi-period framework; and, (ii) incorporate multiple quality characteristics and their interdependencies into a theory of quality-based learning. We also introduce a general modeling approach to incorporate induced learning effects into the autonomous learning curve. The basic idea is that investments in induced learning induces the organization to make forward leaps along the original autonomous learning curve.

2. Modeling framework and key assumptions

To evaluate the quality of a product, more than one quality characteristic is often monitored. In such cases, a multivariate quality loss function is appropriate to use. Let $\mathbf{Y} = (Y_1, \dots, Y_p)$ be the vector of quality characteristics, and, $\mathbf{T} = (T_1, \dots, T_p)$ be the vector of target values associated with those characteristics. Each of the quality characteristics, Y_i , may be directly related to specific inputs and raw materials used in the process. Rather than invoking the standard assumption of process independence, we allow some generality in our model by allowing for positive correlation between characteristics. Then, the expected quality loss function for p characteristics, $E[L]$, is given by (Pignatiello, 1993; Kapur and Cho, 1996):

$$E[L] = \sum_{i=1}^p k_i [(\mu_i - T_i)^2 + \sigma_i^2] + \sum_{i=1}^{p-1} \sum_{j=i+1}^p k_{ij} [\rho_{ij} \sigma_i \sigma_j + (\mu_i - T_i)(\mu_j - T_j)], \quad (1)$$

where μ_i is the mean of Y_i , σ_i^2 is the variance of Y_i , ρ_{ij} is the correlation coefficient between Y_i and Y_j , k_i is the Taguchi loss coefficient associated with Y_i , and k_{ij} is the Taguchi loss coefficient associated with Y_i and Y_j , $i \neq j$. The Taguchi loss coefficients k_i and k_{ij} can be estimated using a regression approach (Kapur and Cho, 1996). Given fixed loss coefficients, the expected quality loss can be decreased by reducing the variances of the characteristics, the biases $(\mu_i - T_i)$, and products of biases. We will focus on learning strategies that influence the variations around the target levels of the respective performance characteristics. The reduction of bias is an independent process, (i.e., dual response (Vining and Myers, 1990)) and is not considered. From a practical perspective, to increase process capability and to create an environment conducive to continuous learning: (i) it is generally more beneficial (and difficult) to invest in reducing process variation than bias (i.e., adjusting the process mean); (ii) knowledge gained from reduction in process variation can be used to reduce bias; and, (iii) often, if not usually, bias can be easily mitigated or eliminated by adjusting process settings. Hence, in the interest of exposition and pragmatism, the terms involving biases in (1) will not be considered, and the expected quality loss per unit of output at time t will therefore be assumed as

$$E[L(t)] = \sum_{i=1}^p k_i \sigma_i^2(t) + \sum_{i=1}^{p-1} \sum_{j=i+1}^p k_{ij} \rho_{ij} \sigma_i(t) \sigma_j(t). \quad (2)$$

As the notation in (2) indicates, all k_i , k_{ij} , and ρ_{ij} are assumed time-invariant. It will also be assumed that quality-related costs are incurred continuously over time, and reduction in (expected) quality costs will be realized by autonomous and induced learning in Y_i that will decrease σ_i^2 , which in turn will decrease $E[L]$. Additional assumptions in our model are as follows:

A1: Autonomous learning for the quality characteristic is described by a traditionally deployed exponential relationship, i.e.,

$$\sigma_i^2(t) = \sigma_i^2(0) e^{-b_i t}, \quad i = 1, \dots, p, \quad b_i > 0, \quad t \geq 0, \quad (3)$$

where b_i is the learning rate associated with Y_i . The exponential-type learning curve is a common assumption in learning research (Zangwill and Kantor, 1998). Equation (3) implies that, as time increases, σ_i^2 decreases at a decreasing rate.

A2: A decision maker has a total of N investment opportunities to accelerate the process of variance reduction of the Y_i 's. We partition the timeline into equally spaced intervals (periods), and plausibly assume that investments in induced learning can be made at the beginning of these periods. The length of the period is such that only one investment in each characteristic is possible per period. (This is plausible from the viewpoint of process management, focus, and resource limitations.) Without loss of generality, we assume that the length of each period is equal to one under a suitably chosen time unit.

A3: An investment in Y_i reflects induced learning by shifting the variance to that of s_i periods later. For example, if we invest in Y_i at the beginning of period j , the new σ_i^2 at time $j + 1 + t$ will be equal to σ_i^2 (prior to investing) at time $j + 1 + t + s_i$, $t \geq 0$. The parameter s_i remains the same for every investment in Y_i , which is consistent with the observation that the decrease in variance generated by an investment becomes smaller as the current variance decreases. An alternative interpretation of this assumption is as follows: Investment in induced learning at time j creates a downward jump on the current learning curve. The variances on the interval between $j + 1$ and $j + 1 + s_i$ on the current curve are skipped over, and the remaining portion of the current curve continues from time $j + 1$ into the future. Thus, each induced learning investment generates a leap down the original autonomous learning curve. This can be related to the concept of *forgetting* (as is discussed by Argote and Epple (1990)), which is modeled by positing that some of the cumulative experience is lost, and thus some earlier points on the learning curve, say for the prior s_i periods are revisited. Our assumption that induced learning accelerates the process down the learning curve “mirrors” the forgetting model; that is, the process variance reflects induced learning that would result from an additional s_i periods of autonomous experience in the process.

Further, assumption A3 implies that each investment in Y_i reduces all future variances by $(1 - e^{-b_i s_i})\%$ from their projected levels prior to the investment. A similar assumption is made for the return from investments in the quality improvement model of Marcellus and Dada (1991). Thus, investment in induced learning can be regarded as a capital investment that yields benefits over multiple periods. Because of autonomous learning, inefficiencies existing in the process decrease with time, and induced learning investments made in later periods yield smaller returns.

For a general learning curve associated with a performance metric (such as cost, defect rate, or variance), an ideal value (goal) for the performance metric can be specified. For example, the hypothetically best possible value for the variance of a quality characteristic is zero. Each improvement made to the system brings it closer to this performance goal. The difference between the current performance level and the performance goal represents what remains to be removed from the system before it can operate optimally (Zangwill and Kantor, 1998). Some researchers have assumed that the marginal improvement rate of the performance metric is proportional to this gap between the performance goal and the current performance level (Levy, 1965; Zangwill and Kantor, 1998; Lapre *et al.*, 2000). In our model each additional investment reduces the variance of the quality characteristic and also the current distance from the ideal target value of zero. Hence, our assumption that returns from investments diminish with additional investments can be seen as consistent with the modeling approaches cited in the literature discussed above.

A4: A finite horizon of n periods is assumed (we refer to it as planning horizon). For now, we assume n is sufficiently

large so that returns from all investments prescribed by the optimal investment plan start to be generated before the end of the n th period. Later, a more explicit lower bound on n will be derived.

A5: All correlation coefficients between the quality characteristics are non-negative. This condition frequently holds in practice, and also, the analysis of the model is more tractable under this correlation structure. (The metrics can also be redefined to satisfy this assumption.)

Finally, although we do not take into account investment costs explicitly, by limiting the total number of investment actions to N , we implicitly assume that a decision maker has a fixed budget of $\$N \times L$ where L is the cost of each individual investment. Thus, the sensitivity of an optimal investment plan with respect to the budget can be explored by changing the value of N . Later, we will discuss the implications of the model when we remove the limit on the available number of investments.

The objective in our model is to minimize the total undiscounted quality-related costs per unit of output over n periods, subject to the constraint that at most N investments in induced learning are possible. The manufacturer’s multi-stage decision problem is essentially determining the optimal number and timing of investments in induced learning in each characteristic.

3. A dynamic programming formulation

For ease of exposition, we first assume that there are two quality characteristics, i.e., $p = 2$. Let $\sigma_i^2(u)$ be the variance of Y_i at time u under a feasible investment policy, i.e., it reflects the effects of both autonomous and induced types of learning. Then, using (2), the total expected cost at time t for the remaining horizon, $TC(t)$, is:

$$TC(t) = \int_t^n E[L(u)]du = k_1 \int_t^n \sigma_1^2(u)du + k_2 \int_t^n \sigma_2^2(u)du + k_{12}\rho_{12} \int_t^n \sigma_1(u)\sigma_2(u)du. \quad (4)$$

The policy that minimizes $TC(0)$ also maximizes the total benefits (rewards) generated by investments in induced learning. Thus, instead of directly minimizing $TC(0)$, we treat the problem as a specialized capital budgeting problem and determine the optimal investment actions over time that maximize the total expected rewards at time 0. The benefit that a particular investment yields depends on when it and other investments are made.

We formulate a Dynamic Programming (DP) model to determine the optimal investment path. At each stage, two decision alternatives (invest, not invest) are available for each of two quality characteristics. Hence, at each state j , $j = 0, 1, 2, \dots$, there are two state variables, $n_1(j)$, and $n_2(j)$ that represent the cumulative number of investments in each quality characteristic prior to state j . The initial state at stage 0 is $n_1(0) = n_2(0) = 0$. Since the final state

and stage when the last investment is made are not prescribed, it is convenient to apply forward DP to find the optimal policy by successively finding the optimal policies for sub-problems of stage 1, 2, 3, ...

Let $G_j(m_1, m_2)$ be the maximum total benefits (rewards) accumulated prior to stage j given that m_1 investments in Y_1 and m_2 investments in Y_2 have been made over stages 0 through $j - 1$. Also let $f_1(m_1, m_2, j)$ be the immediate reward associated with the movement from state $(m_1 - 1, m_2)$ at stage $j - 1$ to the state (m_1, m_2) at stage j , $f_2(m_1, m_2, j)$ be the immediate reward associated with the movement from state $(m_1, m_2 - 1)$ at stage $j - 1$ to the state (m_1, m_2) at stage j , and $f_{12}(m_1, m_2, j)$ be the immediate reward associated with the movement from state $(m_1 - 1, m_2 - 1)$ at stage $j - 1$ to the state (m_1, m_2) at stage j .

Thus the recurrence relation is as follows:

$$G_j(m_1, m_2) = \max\{G_{j-1}(m_1 - 1, m_2) + f_1(m_1, m_2, j), \\ G_{j-1}(m_1, m_2 - 1) + f_2(m_1, m_2, j), \\ G_{j-1}(m_1 - 1, m_2 - 1) + f_{12}(m_1, m_2, j), \\ G_{j-1}(m_1, m_2)\}, \\ G_0(0, 0) = 0, \quad j = 1, 2, \dots \quad (5)$$

The recurrence relation (5) conforms to the standard form of forward DP with an additive objective function. To address boundary conditions, we impose the following restrictions on the state variables and the stage index: $G_i(m_1, m_2)$ is not evaluated when $m_1 > i$, or $m_2 > i$, or $m_1 + m_2 > N$. Also, $G_i(-1, m_2) \equiv G_i(0, m_2)$, $G_i(m_1, -1) \equiv G_i(m_1, 0)$, $f_1(m_1, m_2, i) \equiv 0$ when $m_1 = 0$, $f_2(m_1, m_2, i) \equiv 0$ when $m_2 = 0$, and $f_{12}(m_1, m_2, i) \equiv 0$ when $m_1 = 0$ or $m_2 = 0$. The immediate reward for a particular state transition is the expected savings in quality-related costs for the remaining horizon starting from the beginning of the next stage. Notice that the immediate reward includes the cost savings not only in the next period, but also in all remaining periods. Hence, using (4) and assumptions *A1* and *A3*,

$$f_1(m_1, m_2, j) \\ = (TC(j) \text{ given } n_1(j) = m_1 - 1, n_2(j) = m_2, \\ \text{and no further investment after period } j) \\ - (TC(j) \text{ given } n_1(j) = m_1, n_2(j) = m_2, \\ \text{and no further investment after period } j), \\ = k_1 \sigma_1^2(0) \left[\int_j^n \exp(-b_1(u + s_1(m_1 - 1))) du \right. \\ \left. - \int_j^n \exp(-b_1(u + s_1 m_1)) du \right] \\ + k_{12} \rho_{12} \sigma_1(0) \sigma_2(0) \left[\int_j^n \exp(-b_1(u + s_1(m_1 - 1))/2) \right. \\ \times \exp(-b_2(u + s_2 m_2)/2) du \\ \left. - \int_j^n \exp(-b_1(u + s_1 m_1)/2) \right. \\ \left. \times \exp(-b_2(u + s_2 m_2)/2) du \right]. \quad (6)$$

At first, the definition of $f_1(\cdot)$ may appear counter-intuitive since the difference in total cost is computed by assuming no further investment after period j . Clearly, further investments after period j may be prescribed in the optimal solution to the problem. Between $j = 1$ and $j = n - 1$, $f_1(\cdot)$ is computed as if $j - 1$ is the last decision stage in the problem, and the optimal investment policy for $(j + 1)$ -period problem is determined. At the end of the recursive procedure, when the stage $j = n$ is reached, the maximal expected rewards for the n -period problem has already been computed. Note that each investment changes the expected costs in all of the future periods. Our formulation ensures that the return from a new investment at any stage is incorporated into the total reward function after adjusting it by the effects of all previous investments made up to that stage.

Evaluating the integrals in (6), we have

$$f_1(m_1, m_2, j) \\ = T_1 \exp(-b_1(j + s_1 m_1))(1 - \exp(-b_1(n - j))) \\ + T_{12}(\exp(b_1 s_1/2) - 1) \exp(-b_1(j + s_1 m_1)/2) \\ \times \exp(-b_2(j + s_2 m_2)/2)(1 - \exp(-(b_1 + b_2) \\ \times (n - j)/2)),$$

where

$$T_1 = k_1 \sigma_1^2(0)(\exp(b_1 s_1) - 1)/b_1, \quad \text{and} \\ T_{12} = 2k_{12} \rho_{12} \sigma_1(0) \sigma_2(0)/(b_1 + b_2). \quad (7)$$

Similarly,

$$f_2(m_1, m_2, j) \\ = T_2 \exp(-b_2(j + s_2 m_2))(1 - \exp(-b_2(n - j))) \\ + T_{12}(\exp(b_2 s_2/2) - 1) \exp(-b_1(j + s_1 m_1)/2) \\ \times \exp(-b_2(j + s_2 m_2)/2)(1 - \exp(-(b_1 + b_2)(n - j)/2)),$$

where $T_2 = k_2 \sigma_2^2(0) (\exp(b_2 s_2) - 1)/b_2$, and

$$f_{12}(m_1, m_2, j) \\ = T_1 \exp(-b_1(j + s_1 m_1))(1 - \exp(-b_1(n - j))) \\ + T_2 \exp(-b_2(j + s_2 m_2))(1 - \exp(-b_2(n - j))) \\ + T_{12}(\exp((b_1 s_1 + b_2 s_2)/2) - 1) \exp(-b_1(j + s_1 m_1)/2) \\ \times \exp(-b_2(j + s_2 m_2)/2)(1 - \exp(-(b_1 + b_2)(n - j)/2)).$$

It can be shown that the search for the optimal policy can be reduced to a search among policies in which investments start in the first period and continue without any interruption in subsequent periods. In other words, to determine the optimal investment path, we only need to consider the decisions in the first N stages. Hence, when $p = 2$, in an optimal policy, the investment in induced learning will terminate at the beginning of stage M , where $[N/2] \leq M \leq N$, where $[N/2]$ is the smallest integer larger than or equal to $N/2$. Consequently, in this case of two attributes only $O(N)$ distinct policies need be enumerated. The optimal number

of investments in each variable can only be determined after comparing the costs of all possible policies satisfying the above property. It should be noted that the challenging aspect of the problem is not the number of possible policies, but the computation of costs associated with these policies. Since the magnitude of rewards from a particular investment depends on the previous investment history, a sequence of interdependent computations is required to determine the costs of alternative investment policies. All feasible investment policies are composed of N investments. The total benefit associated with a particular policy is computed by summing the marginal benefits generated by individual investments. Equation (7) illustrates how these marginal benefits can be computed. Hence, although a DP formulation is not strictly necessary once we know the general form of the optimal policy, we would still need to determine the total benefit of a particular policy based on those marginal benefit expressions which are part of our DP formulation. The DP formulation with its recurrence relations helps us (especially when there are an arbitrary number of attributes) to enumerate and evaluate all feasible investment policies in a systematic and structured manner.

Thus, we can rewrite the recurrence relation in (5) as

$$\begin{aligned} G_j(m_1, m_2) &= \max\{G_{j-1}(m_1 - 1, m_2) + f_1(m_1, m_2, j), \\ &G_{j-1}(m_1, m_2 - 1) + f_2(m_1, m_2, j), \\ &G_{j-1}(m_1 - 1, m_2 - 1) + f_{12}(m_1, m_2, j)\}, \\ G_0(0, 0) &= 0. \end{aligned} \quad (8)$$

The integer values for the state variables satisfy the constraints: $n_1 \leq i$, $n_2 \leq i$, $n_1 + n_2 \leq N$, $n_1 + n_2 \geq i$, and $i \leq N$ for all $G_i(n_1, n_2)$ in (8). We can also restate assumption $A4$ regarding the length of the planning horizon: $n \geq N + 1$.

Notice that the optimal solutions and associated expected rewards to the problems with $1, 2, \dots, N - 1$ investment opportunities are also determined as a by-product when the DP algorithm finds the optimal solution to the problem with N investments. If each investment costs $\$L$, we can decide whether an additional investment is worthwhile by comparing it against the computed increase in the expected reward from increasing the total number of investments by one.

3.1. Optimal policy structure

Although we have stated that it is suboptimal to not make an investment in one period and then invest in a later period, it is not obvious which variables should be invested in which time periods. The following lemmas, stated without proof, further characterize the optimal policy, and indicate that if the optimal number of investments in each variable is known, it is not difficult to match these investments with time periods.

Lemma 1. *If it is not optimal to invest in both Y_1 and Y_2 at stage t , then it is not optimal to do so at stage $t + k$, $k \geq 1$.*

Lemma 2. *If it is optimal to invest in only Y_1 (Y_2) at stage t , it is not optimal to invest in only Y_2 (Y_1) at stage $t + k$, $k \geq 1$.*

Combining Lemma 1 and Lemma 2, the form of the optimal policy is determined as:

Corollary 1. *Invest in both Y_1 and Y_2 for the first p_1 periods, then stop investing in one of the variables, and continue to invest in the other variable in the next p_2 periods, $p_1, p_2 \geq 0$.*

Notice that since it is always optimal to make N investments, $2p_1 + p_2 = N$, and that when $N = 1$, it is optimal to invest in only one process at stage 1.

Corollary 1 implies that, regardless of the initial state and learning parameters, it is always beneficial to start investing in a quality characteristic immediately rather than deferring the investment. Thus, if the adherence of an organization to continuous improvement can be described by autonomous learning, additional improvement efforts should be spent at the beginning of the planning horizon. This was especially evident in our experience with a major consumer electronics manufacturer. By significantly investing in quality improvement during the start-up of a new product launch, the company achieved mature quality levels within *days*. Previously, it took 9 months to achieve such quality levels. We also note that our result is consistent with the model of Li and Rajagopalan (1998) who found that the optimal amount of induced learning efforts continuously decrease over time.

Clearly, Corollary 1 points out the collection of $G_j(\cdot)$ terms that actually need to be computed to identify the optimal policy. It can be observed that once j , m_1 , and m_2 are specified in $G_j(m_1, m_2)$, we can compute $G_j(m_1, m_2)$ in the left hand side of (8) directly from the values of (corresponding) $G_{j-1}(\cdot)$ and $f(\cdot)$ without actually searching the maximum of three terms found in the right-hand side of (8). For example, in order to compute $G_8(8, 4)$, we only need the values of $G_7(7, 4)$ and $f_1(8, 4, 8)$; similarly, the value of $G_6(6, 6)$ is found by summing $G_5(5, 5)$ and $f_{12}(6, 6, 6)$. Thus, it is possible to develop a slightly modified and more efficient solution algorithm which computes $G_j(m_1, m_2)$ terms by pruning those policies that are known to be not consonant with Corollary 1.

The model is also applicable to the case where there is only a single quality characteristic. To handle this case, we set ρ_{12} and b_2 to zero. Thus, the optimal policy has the same form for the single-characteristic case.

3.2. Sensitivity of the optimal policy with respect to N

We first show that, *ceteris paribus*, the optimal reward increases at a decreasing rate as the number of investments in a particular characteristic increase.

Lemma 3. *The optimal reward function $G(\cdot)$ is componentwise concave in the number of investments in each variable, n_i .*

Proofs of Lemmas 3 through 6 are given in the Appendix. In the rest of this sub-section, we further assume that:

A6: The correlation coefficient between the characteristics is zero; and

A7: The planning horizon is sufficiently long so that we can treat the immediate reward functions f_1, f_2 , and f_{12} as independent of n .

Then, Lemmas 4 and 5 can be used to accelerate the computation of an optimal policy and to examine its sensitivity with respect to the total number of investments.

Lemma 4. *Assume A6 and A7. When we keep the number of investments in Y_1 (Y_2) fixed, and increase the number of investments in Y_2 (Y_1), the incremental relative return from investing in Y_2 (Y_1) over investing in Y_1 (Y_2) will not increase.*

Lemma 5. *Assume A6 and A7. Let m_1 and m_2 (r_1 and r_2) be the optimal total number of investments in Y_1 and Y_2 for the problem with N ($N + 1$) total investment opportunities. Then, the following relationships between m_1, m_2, r_1 , and r_2 hold:*

$$r_1 \geq m_1, \quad r_2 \geq m_2. \tag{9}$$

The result that the optimal number of investments in each variable does not decrease as the total available number of investments increases is useful, since knowing the optimal solution for the N -investment problem reduces the search efforts for determining the optimal solution for the $(N + 1)$ -investment problem. Note that Lemma 5 does not imply that the optimal investment path for the N -investment problem also subsumes the optimal investment paths for the $1, 2, \dots, N-1$ -investment problems. A numerical example is provided in Table 1. For $N = 5$, the optimal investment path is: $(1,1) \rightarrow (2,2) \rightarrow (2,3)$, we invest in both variables in the stages 0 and 1, and then we invest only in Y_2 at stage 2. For $N = 4$, the optimal investment path is: $(1,1) \rightarrow (1,2) \rightarrow (1,3)$. The optimal investment paths are $(0,1) \rightarrow (0,2) \rightarrow (0,3)$ and $(0,1) \rightarrow (0,2)$ for $N = 3$ and $N = 2$, respectively.

Another consequence of Lemmas 3 and 5 is that the optimal total expected reward follows the law of diminishing returns as the total number of investments increases. Namely, the marginal benefit from each additional investment opportunity will get smaller as more investments are undertaken. Lemma 5 implies that the optimal set of investments in the $(N + 1)$ -investment problem contains the optimal set of investments in the N -investment problem plus one new investment. Thus, the number of investments in one of the variables increases by one as we go from the optimal plan for the N -investment problem to that for the $(N + 1)$ -investment problem. Because of Lemma 3, it follows that the optimal total rewards increase at a decreasing rate as investment opportunities increase. The property of diminishing returns points out that investments in quality

Table 1. Values of $G_j(m_1, m_2)$ by DP algorithm ($n = 400, k_1 = k_2 = 2, k_{12} = 1, \rho_{12} = 0.25, \sigma_1^2(0) = 3, \sigma_2^2(0) = 4, b_1 = 0.01, b_2 = 0.03, s_1 = s_2 = 3$)

j	m_1	m_2	$G_j(m_1, m_2)$
1	1	0	17.86
1	0	1	24.14
1	1	1	41.98
2	2	0	35.03
2	2	1	59.11
2	0	2	45.64
2	1	2	63.45
2	2	2	80.56
3	3	0	51.52
3	3	1	75.57
3	3	2	97.00
3	0	3	64.80
3	1	3	82.59
3	2	3	99.68
4	4	0	67.36
4	4	1	91.39
4	0	4	81.88
4	1	4	99.64
5	5	0	82.58
5	0	5	97.10

improvement are desirable up to the point where the marginal return from the next investment becomes less than the cost of the investment. This critical stopping point can be determined easily by comparing the cost of an investment with the increase in total rewards as we sequentially increase the total number of investment opportunities by one.

We have proven Lemmas 4 and 5 for the case of large n and zero correlation. Although we have been unable to show the equivalent results without these restrictions on n and ρ_{12} , our numerical experimentation indicates that the properties of Lemmas 4 and 5 also appear to hold for small n and nonzero correlation.

3.3. Extension to an arbitrary number of variables

Now, we discuss the DP formulation for $p > 2$. Let U be the set of all process variables in the model, i.e., $X_i \in U, i = 1, \dots, p$. At each stage j , we divide the variables into two disjoint groups, $U_1(j)$ and $U_2(j)$, defined as

- $U_1(j)$: Set of variables invested in stage j ;
- $U_2(j)$: Set of variables not invested in stage j .

Based on the partition above, we also define groups of variable indices

$$\begin{aligned}
 I &= \{i : X_i \in U\}, \\
 q(j) &= \{i : X_i \in U_1(j)\}, \\
 q'(j) &= \{i : X_i \in U_2(j)\}.
 \end{aligned}$$

It is not difficult to generalize (5) to the case $p > 2$. The immediate reward function now is a sum of terms that only contain one or two variables. This separability property, in

fact, enables us to extend the structural results for $p = 2$ to any value of p . The immediate reward associated with the movement from state $(m_i - 1 : i \in q(j - 1), m_i : i \in q'(j - 1))$ at stage $j - 1$ to the state $(m_i : i \in I)$ at stage j is given by

$$\begin{aligned}
 & f_{q(j-1)}(m_i : i \in I, j) \\
 &= \sum_{i \in q(j-1)} T_i \exp(-b_i(j + s_i m_i)) \times (1 - \exp(-b_i(n - j))) \\
 &+ \sum_{i \in q(j-1)} \sum_{k \in q'(j-1)} T_{ik} (\exp((b_i s_i + b_k s_k)/2) - 1) \\
 &\times \exp(-b_i(j + s_i m_i)/2) \exp(-b_k(j + s_k m_k)/2) \\
 &\times (1 - \exp(-(b_i + b_k)(n - j)/2)) \\
 &+ \sum_{i \in q(j-1)} \sum_{k \in q'(j-1)} T_{ik} (\exp(b_i s_i/2) - 1) \exp(-b_i(j + s_i m_i)/2) \\
 &\times \exp(-b_k(j + s_k m_k)/2) (1 - \exp(-(b_i + b_k)(n - j)/2))
 \end{aligned} \tag{10}$$

where $T_i = k_i \sigma_i^2(0) (\exp(b_i s_i) - 1)/b_i$, and,

$$T_{ik} = 2k_{ik} \rho_{ik} \sigma_i(0) \sigma_k(0)/(b_i + b_k), i \in I, k \in I, k > i.$$

Analogously to (7), the recurrence relation is

$$\begin{aligned}
 G_j(m_i : i \in I) &= \max_{q(j-1) \in I} \{G_{j-1}(m_i - 1 : i \in q(j - 1), \\
 & m_i : i \in q'(j - 1)) + f_{q(j-1)}(m_i : i \in I, j)\}.
 \end{aligned}$$

The results in Section 3.1 can be shown for $p > 2$ if all T_{ik} are assumed to be non-negative. Lemmas 1 and 2 can be combined and generalized to the following lemma:

Lemma 6. *If it is not optimal to invest in the subset $q(j)$ or any other subset containing $q(j)$ at stage j , the optimal subset to invest is not $q(j)$ at any later stage.*

The following corollary follows from Lemma 6:

Corollary 2. *If it is optimal to invest in all elements of $q(j)$ in a stage, it is also optimal to invest in all elements of $q(j)$ in all previous stages.*

Thus, the optimal investment policy for $p > 2$ is similar to that for $p = 2$. We start with investing in a set of variables, then, successively the variables are dropped from the investment set one-by-one.

Finally, the rationale behind Lemmas 3 and 5 are also applicable to the case $p > 2$, and thus Section 3.2 can be extended to $p > 2$.

4. Numerical examples

The following numerical examples provide insight into why additional structural properties are difficult to obtain. For the problems described in Table 2, n_i^* denotes the optimal total number of investments in Y_i , and G^* is the maximal total cost savings computed from solving the DP model. We also present the cost savings, G_1 and G_2 , if all N investments are made only in Y_1 and Y_2 , respectively. Consider the example in the third row of Table 2. Absent any investments

Table 2. Sensitivity of the optimal policy with respect to b_1 and s_1 ($N = 6, n = 30, k_1 = k_2 = k_{12} = 1, \sigma_1^2(0) = 3, \sigma_2^2(0) = 2, b_2 = 0.02, s_2 = 2$)

b_1	s_1	ρ_{12}	C_0	G_1	G_2	G^*	n_1^*	n_2^*
0.01	3	0.5	152.46	13.35	11.13	13.79	4	2
0.04	3	0.5	121.76	28.58	10.53	28.58	6	0
0.09	3	0.5	94.20	27.39	9.85	27.46	5	1
0.05	1	0	91.73	9.73	8.24	10.46	4	2
0.05	5	0	91.73	30.26	8.24	30.26	6	0
0.05	7	0	91.73	34.74	8.24	34.86	5	1

in induced learning, the total expected cost at time zero, C_0 , is the sum of quality costs for the n -period horizon under autonomous learning only:

$$\begin{aligned}
 C_0 &= k_1 \sigma_1^2(0) (1 - \exp(-b_1 n)) b_1^{-1} \\
 &+ k_2 \sigma_2^2(0) (1 - \exp(-b_2 n)) b_2^{-1} \\
 &+ 2k_{12} \rho_{12} \sigma_1(0) \sigma_2(0) (1 - \exp(-(b_1 + b_2)n/2)) \\
 &\times (b_1 + b_2)^{-1}.
 \end{aligned}$$

In our example, $C_0 = 94.20$. Using the recurrence relation given by (8), we determine the optimal investment path as: $(0,0) \rightarrow (1,1) \rightarrow (2,1) \rightarrow (3,1) \rightarrow (4,1) \rightarrow (5,1)$. The cost savings associated with this policy: $G_5(5,1) = 27.46$.

We observe that a myopic policy that selects the decision yielding the highest immediate reward at each stage is not optimal in this example. Fine and Porteus (1989) refer to the myopic policy as the last chance policy since it prescribes the optimal decision if there is only one last chance to invest. In our example, the immediate reward at any state is maximized if we can invest in both variables. Hence, a myopic policy would prescribe the following path: $(0,0) \rightarrow (1,1) \rightarrow (2,2) \rightarrow (3,3)$. However, $G_3(3,3) = 25.05$, and thus, not surprisingly, the myopic policy does not maximize total rewards in this case. This result indicates that the practice of across-the-board process improvement strategies (a form of a myopic non-optimal policy) advocated by managers is certainly questionable, which was also found by Moskowitz *et al.* (1997). Essentially, a short-term approach to quality improvement may lead to investment decisions that are non-optimal under a long-term perspective.

According to our numerical experimentation, the optimal investment plan is very sensitive to the value of n (Table 3). The optimal plan stabilizes as n increases, and short planning horizons may create *nervousness* in the system when an investment plan needs to be revised as N varies. This suggests that the planning horizon should be selected sufficiently long so that even if the investment budget changes later, it will likely remain in the vicinity of the optimal plan.

Some learning models in the literature (e.g., Dorroh, *et al.*, 1994) include a salvage value for the amount of learning achieved during the planning horizon. Although we have not made the optimal policy depend on the variances of characteristics at the end of the planning horizon, it can

Table 3. Sensitivity of the optimal policy with respect to planning horizon (n) ($N = 6$, $k_1 = k_2 = k_{12} = 1$, $\rho_{12} = 0$, $\sigma_1^2(0) = 2$, $\sigma_2^2(0) = 4$, $b_1 = 0.06$, $b_2 = 0.02$, $s_1 = s_2 = 3$)

n	C_0	G_I	G_2	G^*	n_I^*	n_2^*
10	51.29	6.41	7.10	9.52	3	3
30	118.06	14.85	23.42	24.82	2	4
40	140.44	16.50	29.44	30.34	1	5
300	232.84	18.49	56.46	56.46	0	6

be seen that increasing n reduces the portion of expected returns from investments that are realized after the n th period. Thus, one may consider that increasing n is equivalent to a lower salvage value for learning that is induced during the n -period quality improvement program.

Regarding the sensitivity of the optimal solution with respect to changes in learning rates, our computational experience suggests that it is hard to draw a general conclusion. We observed that, as the autonomous learning rate for a variable (i.e., b_i) increases, depending on the values of other parameters, the optimal number of investments in that variable may either increase or decrease. The first three rows of Table 2 show the changes in n_i^* and G^* as b_1 changes. A similar pattern can be observed for the impact of induced learning parameters in the last three rows of Table 2. The higher benefits from induced learning in Y_i (i.e., higher s_i) may sometimes lead to a lower number of investments in Y_i . The managerial implication is that, reliable estimation of the parameters will generally be needed to glean the maximum possible benefits from induced learning investments.

5. Concluding remarks

The formulation of an investment model, incorporating autonomous and induced learning, is intended to illustrate how a manufacturing company might plan its future quality improvement actions according to the learning curve characteristics associated with the variables that affect its outgoing product quality. Today's increasingly competitive product markets pressure companies to base their decisions regarding quality on a long-run horizon. Investments in quality should be carefully planned and executed after evaluating the trade-offs associated with alternative uses of funds. Our study draws attention to the effects of the learning curves associated with the quality characteristics on the optimal allocation of quality improvement efforts. One particularly important insight gained from our study is that investment decisions should be made under a long-run perspective. The decisions that maximize benefits in the short-run are not necessarily optimal when a longer planning horizon is considered. We also show that delaying investments in quality is never beneficial.

Our work can be extended in various directions. A direct research extension is inserting some uncertainty into the problem, for example, by making the outcomes of induced

learning investments stochastic. It may also be interesting to explore the robustness of the optimal investment plan and expected rewards with respect to imprecisely estimated model parameters. Finally, our model might be applied to investigate induced learning in activities other than those associated with quality.

Acknowledgement

The authors are grateful to two anonymous reviewers for their insightful comments.

References

- Adler, P.A. and Clark, K.B. (1991) Behind the learning curve: a sketch of the learning process. *Management Science*, **37**, 267–281.
- Argote, L. and Epple, D. (1990) Learning curves in manufacturing. *Science*, **247**, 920–924.
- Dorroh, J.R., Gulledge, T.R. and Womer, N.K. (1994) Investment in knowledge: a generalization of learning by experience. *Management Science*, **40**, 947–958.
- Fine, C. (1986) Quality improvement and learning in productive systems. *Management Science*, **32**, 1301–1315.
- Fine, C. and Porteus, E.L. (1989) Dynamic process improvement. *Operations Research*, **37**, 580–591.
- Ittner, C.D. (1996) Exploratory evidence on the behavior of quality costs. *Operations Research*, **44**, 114–130.
- Ittner, C.D., Nagar, V. and Rajan, M.V. (2001) An empirical examination of dynamic quality-based learning models. *Management Science*, **47**, 563–578.
- Juran, J.M. (1951) *Quality Control Handbook*, McGraw-Hill, New York, NY.
- Kackar, R.N. (1985) Off-line quality control, parameter design, and the Taguchi method. *Journal of Quality Technology*, **17**, 176–188.
- Kapur, K.C. and Cho, B. (1996) Economic design of the specification region for multiple quality characteristics. *IIE Transactions*, **28**, 237–248.
- Kini, R.G. (1994) Economics of conformance quality. Ph.D. dissertation, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA, USA.
- Lapre, M.A., Mukherjee, A. S. and Van Wassenhove, L. N. (2000) Behind the learning curve: linking learning activities to waste reduction. *Management Science*, **46**, 597–611.
- Levy, F.K. (1965) Adaptation in the production process. *Management Science*, **11**, B136–B154.
- Li, G. and Rajagopalan, S. (1997) The impact of quality on learning. *Journal of Operations Management*, **15**, 181–191.
- Li, G. and Rajagopalan, S. (1998) Process improvement, quality, and learning effects. *Management Science*, **44**, 1517–1532.
- MacKay, R.J. and Steiner, S.H. (1997) Strategies for variability reduction. *Quality Engineering*, **10**(1), 125–136.
- Marcellus, R.L. and Dada, M. (1991) Interactive process quality improvement. *Management Science*, **37**, 1365–1376.
- Mazzola, J.B. and McCardle, K.F. (1997) The stochastic learning curve: optimal production in the presence of learning curve uncertainty. *Operations Research*, **45**, 440–450.
- Moskowitz, H., Plante, R.D. and Tang, J. (1997) Allocation of variance targets among suppliers. CMME working paper series, Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907, USA.
- Mukherjee, A.S., Lapre, M.A. and Van Wassenhove, L.N. (1998) Knowledge driven quality improvement. *Management Science*, **44**, S35–S49.

- Pignatiello, J.J. (1993) Strategies for robust multiresponse quality engineering. *IIE Transactions*, **25**, 5–25.
- Plante, R. (2000) Allocation of variance reduction targets under the influence of supplier interaction. *International Journal of Production Research*, **38**, 2815–2827.
- Schneiderman, A.M. (1988) Setting quality goals. *Quality Progress*, **21**(4), 51–57.
- Stata, R. (1989) Organizational learning—the key to management innovation. *Sloan Management Review*, Spring, 63–74.
- Taguchi, G. and Clausing, D. (1990) Robust quality. *Harvard Business Review*, Jan.–Feb. 65–75.
- Vining, G.G. and Myers, R.H. (1990) Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology*, **22**, 38–45.
- Zangwill, W.I. and Kantor, P.B. (1998) Toward a theory of continuous improvement and the learning curve. *Management Science*, **44**, 910–920.

Appendix

Proof of Lemma 3. To simplify the notation, we omit the stage index implied by the function arguments. Recall that Corollary 1 implies $j = \max(m_1, m_2)$ in $G_j(m_1, m_2)$. The concavity of $G(\cdot)$ in n_1 is equivalent to

$$G(m_1 + 1, m_2) - G(m_1, m_2) > G(m_1 + 2, m_2) - G(m_1 + 1, m_2). \quad (\text{A1})$$

Consider the left-hand side of (A1):

$$\begin{aligned} G(m_1 + 1, m_2) - G(m_1, m_2) &= f_1(m_1 + 1, m_2, m_1 + 1) \quad \text{if } m_2 \leq m_1 + 1, \\ &= f_1(m_1 + 1, m_2, m_2) \quad \text{if } m_2 > m_1 + 1. \end{aligned}$$

Similarly for the right-hand side of (A1):

$$\begin{aligned} G(m_1 + 2, m_2) - G(m_1 + 1, m_2) &= f_1(m_1 + 1, m_2, m_1 + 1) \quad \text{if } m_2 < m_1 + 2, \\ &= f_1(m_1 + 2, m_2, m_2) \quad \text{if } m_2 \geq m_1 + 2. \end{aligned}$$

Hence, depending on the values of m_1 and m_2 , we can rewrite (A1) as

$$f_1(m_1 + 1, m_2, m_2) > f_1(m_1 + 2, m_2, m_2) \quad \text{if } m_2 \geq m_1 + 2. \quad (\text{A2})$$

$$f_1(m_1 + 1, m_2, m_1 + 1) > f_1(m_1 + 2, m_2, m_1 + 2) \quad \text{if } m_2 \leq m_1 + 1. \quad (\text{A3})$$

The direction of inequalities in (A2) and (A3) follow from (6). Concavity of $G(\cdot)$ in n_2 can be shown analogously. ■

Proof of Lemma 4. Suppose we keep the number of investments in Y_1 fixed. In order to prove Lemma 4 it is sufficient to demonstrate the following inequality:

$$G(m_1, m_2 + 1) - G(m_1 + 1, m_2) \geq G(m_1, m_2 + 2) - G(m_1 + 1, m_2 + 1). \quad (\text{A4})$$

Rewrite (A4) and define L and R such that:

$$\begin{aligned} L &\equiv G(m_1 + 1, m_2 + 1) - G(m_1 + 1, m_2) \\ &\geq R \equiv G(m_1, m_2 + 2) - G(m_1, m_2 + 1). \end{aligned} \quad (\text{A5})$$

It can be observed that

$$\begin{aligned} L &= f_2(m_1 + 1, m_2 + 1, m_1 + 1) \quad \text{if } m_1 + 1 > m_2 + 1, \\ &= f_2(m_1 + 1, m_2 + 1, m_2 + 1) \quad \text{if } m_1 + 1 \leq m_2 + 1, \\ R &= f_2(m_1, m_2 + 2, m_1) \quad \text{if } m_1 > m_2 + 2, \\ &= f_2(m_1, m_2 + 2, m_2 + 2) \quad \text{if } m_1 \leq m_2 + 2. \end{aligned}$$

We will show that $L \geq R$ always. First, we consider the case $m_1 > m_2 + 2$. Then, substituting $\rho_{12} = 0$,

$$\begin{aligned} L &= f_2(m_1 + 1, m_2 + 1, m_1 + 1) \\ &= T_2 \exp(-b_2[m_1 + 1 + s_2(m_2 + 1)]), \quad \text{and} \\ R &= f_2(m_1, m_2 + 2, m_1) = T_2 \exp(-b_2[m_1 + s_2(m_2 + 2)]). \end{aligned}$$

Clearly $L = R$ if $s_2 = 1$, and $L > R$ if $s_2 > 1$. Now assume that $m_2 < m_1 \leq m_2 + 2$. In this case,

$$\begin{aligned} L &= T_2 \exp(-b_2[m_1 + 1 + s_2(m_2 + 1)]), \quad \text{and} \\ R &= T_2 \exp(-b_2[m_2 + 2 + s_2(m_2 + 2)]). \end{aligned}$$

Again, $L \geq R$ in this scenario. Finally, consider the case $m_1 \leq m_2$ in which L and R are given by

$$\begin{aligned} L &= T_2 \exp(-b_2[m_2 + 1 + s_2(m_2 + 1)]), \quad \text{and} \\ R &= T_2 \exp(-b_2[m_2 + 2 + s_2(m_2 + 2)]). \end{aligned}$$

We again observe that $L \geq R$. This concludes the proof of Lemma 4 in the case that the number of investments in Y_1 is kept fixed. We apply same reasoning to prove Lemma 4 in the case where we fix the number of investments in Y_2 . ■

Proof of Lemma 5. To show that (9) is true, it is sufficient to show that (A6) or (A7) is not optimal:

$$r_1 = m_1 - 1, \quad r_2 = m_2 + 2. \quad (\text{A6})$$

$$r_1 = m_1 + 2, \quad r_2 = m_2 - 1. \quad (\text{A7})$$

Once (A6) and (A7) are shown not to be optimal, it can be shown similarly that other combinations of r_1 and r_2 that do not satisfy (9) also are not optimal.

We will refer to the problem with N learning investment opportunities as the N -investment problem. Now suppose we need to make the last investment decision. At this point, marginal returns from investing in Y_1 and Y_2 determine which variable is selected for investment. For the N -investment problem, when $m_1 - 1$ investments in Y_1 and m_2 investments in Y_2 have been already made, we know that it is optimal to invest in Y_1 next. Now consider the last investment decision in the $(N + 1)$ -investment problem. Lemma 4 implies that if investing in Y_1 is preferable to investing in Y_2 given $m_1 - 1$ previous investments in Y_1 and m_2 previous investments in Y_2 , it will also be preferable given $m_1 - 1$ previous investments in Y_1 and $m_2 + 1$ previous investments in Y_2 . This implies that (A6) cannot be optimal for the $(N + 1)$ -investment problem. On the other

hand, if m_1 and $m_2 - 1$ investments in Y_1 and Y_2 have been made in the N -investment problem, the next optimal investment is in Y_2 . Because of Lemma 4, if investing in Y_2 is preferable to investing in Y_1 given m_1 investments in Y_1 and $m_2 - 1$ investments in Y_2 , it will also be preferable given $m_1 + 1$ investments in Y_1 and $m_2 - 1$ investments in Y_2 . This implies that (A7) cannot be optimal for the $(N + 1)$ -investment problem. ■

Proof of Lemma 6. Suppose we are at state $(m_i : i \in I)$ at stage j . Let $r(j)$ be any subset of I that includes all elements of $q(j)$, and let $s(j)$ be the subset consisting of all elements of $r(j)$ that are not in $q(j)$. Similarly to (A3), the following relationship holds:

$$\begin{aligned} & f_{r(j)}(m_i + 1 : i \in r(j), m_i : i \in r'(j), j + 1) \\ & > f_{s(j)}(m_i + 1 : i \in s(j), m_i : i \in s'(j), j + 1) \\ & + f_{q(j)}(m_i + 1 : i \in r(j), m_i : i \in r'(j), j + 2). \end{aligned} \quad (\text{A8})$$

(A8) can be verified by using (10). We can compare the left- and right-hand sides of (A8) term-by-term. The comparison of the terms involving only one variable is straightforward. For the terms resulting from covariances, because of separability, we can consider each two-variable combination in isolation, for which the direction of inequality in (A8) holds in a similar manner to (A3). Since the direction of inequality holds for every covariance term associated with a pair of variables and the total reward function is additive, (A8) is also satisfied when all covariance terms are considered together. Hence, any policy not conforming to Lemma 6 is not optimal. ■

Biographies

Dogan A. Serel is an Assistant Professor at the Faculty of Business Administration, Bilkent University, Turkey. He received his Ph.D. in Management from Purdue University. His main research interests are in inventory management and quality control.

Maqbool Dada teaches operations management at the Krannert Graduate School of Management at Purdue University. He received his Ph.D. in Management from the Sloan School of Management at MIT. His research interests include inventory systems, pricing models, service systems, and international operations management.

Herbert Moskowitz is the Lewis B. Cullman Distinguished Professor of Manufacturing Management and is Director of the Dauch Center for the Management of Manufacturing Enterprises at the Krannert Graduate School of Management, Purdue University. His area of specialization is management science and quantitative methods with emphasis on manufacturing and technology, total quality management, quality improvement tools, and judgment and decision making. He has been at Purdue since 1970 and has had visiting appointments at the University of Mannheim, West Germany, the University of British Columbia, the London Business School, and the Wharton School. He holds a B.S. degree in Mechanical Engineering, an M.B.A. and a Ph.D. in Management from UCLA. He is the co-author of five texts and has published about 140 articles in the areas of decisionmaking, optimization, management science, and quality control in academic journals. He is an Associate Editor of *Operations Research* and has been an Associate Editor of *Decision Sciences*, the *Journal of Behavioral Decision Making*, the *Journal of Interdisciplinary Modeling and Simulation* and a Special Associate Editor of *Management Science*.

Robert D. Plante is the Senior Associate Dean and James Brooke Henderson Professor of Management at the Krannert School of Management, where his research interests include the development of state-of-the-art statistical quality control and improvement models. His efforts have focused on the following classes of problems: (i) robust product/process design; (ii) screening procedures for process control and improvement; (iii) statistical/process/dynamic process control models; and (iv) specialized process improvement problems. He has more than 50 research publications in these areas which have appeared in numerous journals, including *Operations Research*, *Management Science*, *Decision Sciences*, *Journal of the American Statistical Association*, *International Journal of Production Research*, *The Accounting Review*, *Auditing: A Journal of Practice and Theory*, *Naval Research Logistics*, *IIE Transactions*, *Technometrics*, *Production and Operations Management*, *Computers and Operations Research*, *Journal of Quality Technology*, *European Journal of Operational Research*, *Manufacturing and Service Operations Management*, *Journal of the Operational Research*, *Information Systems and Operational Research*, and *Journal of Business and Economic Statistics*.

Contributed by the Design of Experiments and Robust Designs Department