

**CAPACITY EXPANSION AND REPLACEMENT
WITH UNCERTAIN TECHNOLOGICAL BREAKTHROUGHS**

Sampath Rajagopalan
School of Business
University of Southern California
Los Angeles, CA 90089-1421

Medini R. Singh
Department of Industrial & Operations Engineering
University of Michigan
Ann Arbor, MI 48109-2117

Technical Report 92-3
January 1992

Abstract

A sequence of technological breakthroughs are anticipated but their magnitude and timing are uncertain. A firm, operating in such an environment, must decide how much capacity of the current best available technology to acquire to meet the future demand growth. It must also determine whether to replace any of the older vintages partially or completely, with the focus on obsolescence rather than deterioration as the motive for replacement. Such decisions are of the essence in electronics and other industries characterized by rapid technological change. Our analysis shows that it is optimal to purchase, dispose and replace capacity in amounts equal to the demand increments for an integral number of periods. Also, disposal of excess capacity need be considered only in a period when a new technology appears, and replacement decisions need be made only in periods when capacity is purchased to meet future demand. Using these results, a highly efficient regeneration point based dynamic programming recursion is presented. Computational experience suggest that it is possible to solve moderate size problems on a personal computer.

1 Introduction

The growth of electronics has had a significant impact on the pace of technological change in a variety of industries. In industries such as computers, telecommunications, medical equipment and machine tools, the silicon chip has resulted in new generations of products with superior performance. The chip industry itself undergoes even faster technological change. For example, successive generations of memory chips have appeared regularly every 3 to 5 years. (4-megabyte chip samples were first shipped in 1988, 16-megabyte samples were shipped in 1991 and 64-megabyte samples are expected in 1994, as reported in *Business Week* 1990). Less frequent but major innovations also occur in other industries such as automobiles and pharmaceutical. Managers face a difficult choice in such environments — whether to acquire (or upgrade to) the latest technology that appears, or to wait for the next innovation. By upgrading to the latest technology, the firm reaps the benefits of the latest technology, say in terms of lower operating costs. On the other hand, the firm may have to pay the acquisition costs and may then find that another innovation appears soon, making its previous technology investment relatively obsolete.

The other key tradeoff that a manager faces in making technology acquisition decisions is the issue of how much capacity of new technology to acquire and how much capacity of the older technology to dispose in any period. Scale economies in acquiring and disposing capacity and cost of carrying excess capacity need to be considered in making these decisions, in addition to the factors mentioned earlier. There is also an interesting interplay of scale economies and technological uncertainty. For instance, purchasing small capacity increments enables a firm to hedge against the uncertainties of rapid technological obsolescence.

In this paper, we focus on the impact of such uncertain technological breakthroughs, say in equipment technology, on the *capacity acquisition and replacement* decisions of a firm that uses this technology in its production process. As we have pointed out, the manager has to consider a number of complex tradeoffs in making these decisions. The objective of this paper is to develop and solve a realistic model of the technology acquisition process to help decision-makers in such environments. In particular, this paper was motivated by the problem faced by a firm that leases mobile magnetic resonance imaging (MRI) equipment to hospitals across the country. The MRI equipment are expensive (upto \$6 million) and major technological upgrades appear about once in

three years (Van Horssen 1991). The firm has over 50 MRI units and is often confronted with the issue of whether or not to upgrade and how many of the units to upgrade.

We consider a discrete time model with a sequence of breakthroughs that are separated by stochastic time intervals. When a breakthrough occurs, a new equipment vintage becomes available. We assume that this new vintage dominates the old one, that is, given a choice of which vintage to adopt, for any size expansion or replacement, the new one would be preferred. The level of technology achieved with a breakthrough is allowed to be one of several higher levels. In this model, we assume a finite number of possible technology levels, and that the inter-arrival time distribution between breakthroughs depends only on the last technology level achieved. Hence, we can model such issues as saturation in breakthroughs as higher technology levels are achieved.

We consider a model with purchase costs for the acquisition of capacity that are a function of the amount and vintage of capacity. The other costs considered are costs of carrying excess capacity and operating costs, both of which depend on the vintage. We also allow for disposal of older vintages in any period, with a salvage cost function that depends on the amount and vintage of capacity salvaged, and the new vintage available. The demand for capacity is deterministic and nondecreasing over time. In each period, the firm has to decide how much capacity of the current best vintage to acquire, and how much capacity of older vintages to replace. In the next section, we briefly review the relevant literature. In Section 3, we present the general model considered here and prove some key properties of the model that are used in subsequent sections to provide efficient stochastic dynamic programming recursions. In Section 4, we present the recursion for the case where capacity once used cannot be replaced and in Section 5 we present the recursion for the case where used capacity can be replaced. Results of a computational study are briefly discussed in Section 6. In Section 7, we point out some future research directions.

2 Review of the Literature

We briefly review some of the papers in the machine or technology replacement literature and the capacity expansion literature (Luss 1982) that address the key issues identified earlier.

The early equipment replacement models (refer Pierskella and Voelker 1976 for a survey), with

just one technology, simply consider the issue of optimal frequency of replacing machines to trade off the increased operating and maintenance costs with age of machine against the fixed costs of replacing machines. Sethi and Chand (1979, 1982) consider deterministic machine replacement models with a choice of machines available. They consider an improving technological environment over time, with better machines *available with certainty in successive time periods*. For such models, they present forward algorithms and planning horizon procedures. Jones, Zydiak, and Hopp (1991) consider a more general model with multiple (or parallel) machines, and fixed and variable costs associated with replacing the machines, with the number of machines constant over time. They present two key results to reduce the computational effort in determining an optimal replacement policy.

There have also been a number of papers in the replacement literature, wherein the time of appearance of a new technology (machine) is not deterministic. Goldstein, Ladany, and Mehrez (1988) consider a machine replacement problem with one anticipated technological breakthrough characterized by a constant hazard rate. They formulate a stationary dynamic programming model and develop a solution procedure, using the approach of Sethi and Chand (1979), to determine the replacement timing of the current technology machine. Nair and Hopp (1989) consider a similar model with one anticipated breakthrough but do not restrict themselves to constant hazard rates. Also, they use a forecast horizon approach to find the optimal decision. Balcer and Lippman (1984) present a rich model with a sequence of technological innovations that may or may not be adopted, with uncertainty in the time between discoveries, the size of each discovery as well as the future pace of discovery. While the stochastic evolution of technology in our paper is modeled in a manner similar to theirs, the focus of our paper is different from theirs in a number of important ways. First, the choices before the firm in Balcer and Lippman are to either replace the current technology with the new technology or do nothing. In our paper, the firm can decide the optimal amount of capacity of the new technology to purchase. This allows the firm to hedge against uncertainty in the future by buying smaller capacity increments. Hence, we do not have a *replace all or nothing* scenario. Second, Balcer and Lippman provide economic insights into the impact of expectations of future technological developments on the replacement decisions. The focus of our paper is operational rather than economic, in that we aim to develop efficient solution procedures to determine the optimal amount of capacity of different technologies to be purchased and replaced. However, we do present a number of key results that provide new insights into the

impact of future technological developments on capacity acquisition and replacement decisions. The stochastic evolution of technology is similar in Monahan and Smunt (1989), where they consider a Markov model with a sequence of technological innovations and run simulations to analyze the impact of uncertainty in interest rates and state of technology on the rate of technology acquisition. Cohen and Halperin (1986) considered a model with a choice of technologies available, where a technology is fully replaced by another technology for some fixed cost. They derive interesting analytical results on the types of technologies likely to be acquired over time.

Capacity expansion models, unlike the machine replacement models, permit consideration of scale economies. Further, in the environments considered in this paper with high technological obsolescence and uncertainty, permitting the incremental acquisition of capacity makes the model richer. While there is a vast literature on capacity expansion models (Freidenfelds 1981, Luss 1982), we are not aware of any papers that consider capacity expansion issues in the context of a sequence of random technological developments. We briefly review some related work in the capacity expansion literature that also addresses technology acquisition issues. Hinomoto (1965) considers a deterministic model of capacity expansion when facilities undergo technological improvement. Gaimon (1989) presents a dynamic game analysis to understand the impact of competitive forces on the acquisition of new technology capacity and disposal of old technology. Gaimon (p. 410) mentions at the outset that "... firms also run the risk of making an enormous investment in technology that may soon become obsolete." Klincewicz and Luss (1985) consider the issue of when to install facilities of fixed capacity using current technology, when there is spare capacity of the old technology. They provide procedures for determining the optimal timing decisions under linear and non-linear demand growth.

3 The Model and Some Properties

In this section, we first present a general model for the problem. We then generalize some of the regeneration properties present in simpler models (Wagner and Whitin 1958, Veinott 1966) to the model considered here. In each period, there is a certain demand for *additional* capacity. Let d_t denote the *increase* in demand for capacity in period t . We assume that $d_t \geq 0$. Many papers in the capacity expansion literature make this assumption (Luss 1982) and it appears to be reasonable in

dynamic environments, where the appearance of new technologies stimulates rapid demand growth.

The sequence of events in any period is as follows. The information regarding the appearance of a new technology becomes available at the beginning of a period. Based on this information, capacity disposal, if any, takes place followed by possible acquisition. We now define the decision variables for the model. Let x_{mtj} denote the amount of vintage m capacity acquired in period t and disposed in period j ($> t$). The index j in variable x_{mtj} insures that capacity disposed in a period must have been purchased earlier. Let Y_{mt} be the total vintage m capacity acquired in period t , i.e.,

$$Y_{mt} = \sum_{j>t} x_{mtj} \quad \forall m, t \quad (3.1)$$

Let Z_{mj} be the total vintage m capacity disposed in period j , i.e.,

$$Z_{mj} = \sum_{t<j} x_{mtj} \quad \forall m, j \quad (3.2)$$

Let the excess capacity of vintage m at the end of period t be denoted by I_{mt} . The firm does not allow capacity shortages. Also, let Z_{mt}^e denote the amount of excess or unused capacity of vintage m disposed in period t . It is useful to distinguish between *unused* and *used* capacity when the salvage cost depends upon whether the capacity has been used or not. For example, replacement of used capacity may involve disruption of production activities, in which case the salvage cost of used capacity will be higher than that of unused capacity. As another example, a car with no miles on its odometer will fetch a far higher price than a car with even just 100 miles on it. Now, the amount of unused capacity disposed in a period cannot exceed the amount available. That is,

$$Z_{mj}^e \leq I_{m,j-1} \quad \forall m, j \quad (3.3)$$

The equation balancing the demand and supply of capacity in each period is,

$$\sum_p (I_{p,t-1} + Y_{pt} - Z_{pt} - I_{pt}) = d_t \quad \forall t \quad (3.4)$$

Finally, the non-negativity conditions on variables are

$$x_{mtj}, I_{mt}, Y_{mt}, Z_{mt}, Z_{mt}^e \geq 0 \quad \forall m, t, j \quad (3.5)$$

The total vintage m capacity *utilized* in period t is,

$$= I_{m0} + \sum_{\tau \leq t} \sum_{j > \tau} x_{m\tau j} - I_{mt}$$

We now specify various costs in the model. The expected cost of acquiring technology m capacity in period t is represented by $f_{mt}(\cdot)$. Increasing values of m correspond to better technologies. The acquisition cost for technologies not yet available in a period is assumed to be infinite. The cost of carrying a unit of excess capacity of technology m in period t is denoted by $h_{mt}(\cdot)$. We assume the cost functions $f_{mt}(\cdot)$ and $h_{mt}(\cdot)$ are concave to reflect economies of scale; this is a common assumption in the capacity expansion literature (Luss 1982). The operating cost per unit capacity for vintage m , when used in period t , is c_{mt} . Note that $f_{mt}(\cdot)$, $h_{mt}(\cdot)$, and c_{mt} are all functions of time and of the vintage. This can be used to model phenomena such as declining purchase costs with time (due to wider acceptance of a technology), and lower operating costs for newer vintages.

When a new technology appears, there may be excess (unused) and used capacity of older technologies, both of which may be disposed fully or partially. Let $g_{pmt}(\cdot)$ denote the net salvage cost (may be negative) from disposal of *excess* capacity of technology p in period t , given that the latest technology available is m ($> p$). Similarly, $\bar{g}_{pmt}(\cdot)$ denotes the net salvage cost (potentially negative) from disposal of *used* capacity of technology p in period t , given that the latest technology available is m . The salvage costs $g_{pmt}(\cdot)$ and $\bar{g}_{pmt}(\cdot)$ are functions of time (t), the vintage disposed (p), and the best vintage currently available (m). Hence, in our model, the salvage cost changes due to the appearance of new technologies, rather than due to age. This is unlike in the machine replacement literature, where salvage values are a function of the age of the machine. This is because the focus in the machine replacement literature is the deterioration of machines with age, leading to an increase in operating and maintenance costs that induces replacement. However, the focus in this paper is on the increasing obsolescence of an older technology with the appearance of newer technologies, rather than on deterioration. In industries such as computers, where electronics has led to rapid technological changes, obsolescence is the primary motivation for replacement rather than physical deterioration. For example, the salvage value of a PC drops significantly when a more powerful PC is introduced, but is not very sensitive to when it was purchased or how long it was used.

We assume the following functional form for the salvage cost functions $g_{pmt}(\cdot)$:

$$g_{pmt}(z) = s_{pmt} \delta(z) - r_{pmt} z$$

where $\delta(z) = 1$ if $z > 0$, and $\delta(z) = 0$ if $z = 0$. The non-negative parameters s_{pmt} and r_{pmt} represent respectively, the setup cost and unit revenue from the disposal of technology type p , when m is the current best technology available. The cost of disposing used capacity, $\bar{g}_{pmt}(\cdot)$ is expressed similarly. Other papers in the literature, including those by Love(1973) and Chand and Morton(1982) have used similar salvage cost functions.

The probabilistic evolution of technology is modeled as a function of two factors: (i) the number of periods between two consecutive innovations and (ii) the new level of technology achieved with an innovation. Let us assume that the number of possible technology levels achievable within the potential problem horizon is M , with technology level $(m + 1)$ representing a clear improvement over technology m ; m is assumed to be integer-valued. In particular, we assume that in the chain of technological improvements, each successive vintage dominates all others that have appeared before. Given a choice, one would always prefer to acquire the newest vintage. We model the evolution of technology as a semi-Markov process (Ross 1983). Given that vintage m has just become available, let $Q_m(\cdot)$ and $q_m(\cdot)$ denote the cumulative distribution function (cdf) and probability distribution function (pdf) of the amount of time until the next vintage becomes available. We assume $q_m(0) = 0$. Hence the number of periods between the appearance of successive vintages depends on the vintage m last achieved, but are independent otherwise. The level of technology available, once a breakthrough occurs, changes according to a Markov process with a one-step transition matrix $\mathbf{P} = [P_{mn}]$, where P_{mn} is the probability that vintage n appears, given that the last vintage that appeared is m . The actual costs of acquiring the technology may depend on the firm's past experience and ability in implementing new technologies and this could be incorporated in the expected cost function $f_{mt}(\cdot)$.

A mathematical programming formulation of the problem can now be given as:

$$\underset{x_{prj}, I_{pt}, Y_{pt}, Z_{pt}, Z_{pt}^e}{\text{Minimize}} \sum_{t=1}^T \sum_{m(1), m(2), \dots, m(t)}^E \sum_{p=1}^{m(t)} \left(f_{pt}(Y_{pt}) + h_{pt}(I_{pt}) + c_{pt} \left(I_{p0} + \sum_{\tau \leq t} \sum_{j > t} x_{prj} - I_{pt} \right) \right. \\ \left. + g_{pmt}(Z_{pt}^e) + \bar{g}_{pmt}(Z_{pt} - Z_{pt}^e) \right)$$

subject to: (3.1), (3.2), (3.3), (3.4), and (3.5).

The expectation operator E in period t is defined over the joint distribution of the technology levels $m(i)$, $i = 1, \dots, t$, where $m(i)$ is the best technology level available in period i . The joint distribution, in turn, is a function of the time to discovery distribution $Q_i(\cdot)$ and the one-step transition matrix \mathbf{P} . The first term within the summation signs is the acquisition cost for capacity purchases of technology type p made in period t . The second term is the cost of carrying excess capacity. The third set of terms is the operation cost in period t for technology p . The fourth and fifth terms, respectively, are the salvage cost of disposing excess and used capacity in period t . In any period t , the only technologies that can be acquired, operated and disposed are those that have appeared until period t , i.e., $p = 1, \dots, m(t)$.

We now present our main result that is critical in reducing the state space.

Theorem 1 *There exists an optimal solution to the problem that has the following properties:*

(i) *One would never purchase capacity of a vintage in a period when there is excess capacity of the same or different vintage on hand, i.e.,*

$$I_{p,t-1} Y_{mt} = 0 \quad \text{for all } p, m, t \quad (3.6)$$

(ii) *One would never purchase capacity of more than one vintage in any period, i.e.,*

$$Y_{pt} Y_{mt} = 0 \quad \text{for all } p, m, t \quad (3.7)$$

(iii) *At any time, there would never be excess capacity of more than one vintage on hand, i.e.,*

$$I_{pt} I_{mt} = 0 \quad \text{for all } p, m, t \quad (3.8)$$

Proof First, for a *given sample realization*, note that the objective function is deterministic. It is also concave since the functions $f(\cdot)$ and $h(\cdot)$ are concave, $g_{pmt}(\cdot)$ and $\bar{g}_{pmt}(\cdot)$ are fixed charge functions, and operation costs are linear. Second, we have a set of linear constraints.

Given these two observations, for a particular realization of the stochastic evolution of technology, the optimal solution is at an extreme point. By definition, *an extreme point solution* dominates all non-extreme point solutions for any set of values of the objective function coefficients (i.e., for any sample realization). Observe that the only uncertainty in the formulation is in the objective function coefficients (costs and salvage values), since demand is known. The expected cost function is simply a weighted average of the total cost for all the sample realizations (i.e., a linear combination), and so an extreme point solution dominates non-extreme points in terms of expected costs too. Now, note that in equations (3.1–3.4), all the variables— x , I , Y , Z and Z^e appear exactly once with a positive (+1) coefficient; in all other occurrences they have a negative (–1) coefficient. This can be verified simply by transferring the right hand side terms to the left in constraints (3.2–3.4) and transferring the left hand side terms to the right in equation (3.1). Also, all the variables are non-negative from (3.5). As a result, the constraint matrix determined by equations (3.1–3.4) is *Leontief* (Veinott 1969). From the characterization of extreme points in Leontief matrices (Veinott 1969, also Theorem 6 in Veinott 1968), it follows that if more than one variable appears with a positive coefficient in the same constraint, then only one of these variables can be positive in the optimal solution. Hence, conditions (3.6–3.8) follow directly from constraint (3.4). \square

Regeneration results similar to those in Theorem 1 have been developed in earlier work (Veinott 1969, Zangwill 1968, Love 1973, Chand and Morton 1982) in a *deterministic context* and for models without replacement. Next, we state a number of key implications of Theorem 1 that will be useful in developing the recursive formulation in the next two sections. The conditions (3.6–3.8) together imply directly that,

Corollary 1 *There exists an optimal solution in which capacity purchases to meet future demand and excess capacity disposed correspond to demand increments for an integral number of periods.*

If this were not true, one would end up in a period with excess capacity on hand that is not enough to meet the demand increment for that period. Then capacity will have to be acquired in that period, violating condition (3.6).

Corollary 2 *There exists an optimal solution in which whenever a unit of used capacity is replaced, all units corresponding to that vintage are replaced.*

Consider a situation when the firm has an amount of excess capacity on hand that is just enough to meet the demand increment for the current period. The firm can either put this capacity to use or dispose it altogether and acquire another technology to meet this demand. It will not be optimal, according to Corollary 1, to dispose the excess capacity partially. Now, consider the replacement of used capacity of a technology. All past demand increments met by this capacity can be thought of as demand increment for the current period with an amount of available excess capacity that is just enough to meet this demand. The fact that this “excess” capacity has been used earlier has no relevance, except that cost function $\bar{g}(\cdot)$ instead of $g(\cdot)$ shall be used for deciding whether to dispose or not. As a result, the firm can either (continue to) put this “excess” capacity to use, or dispose it altogether and acquire another technology to meet this demand. In other words, it will not be optimal to dispose the used capacity of a technology *partially*.

Corollary 3 *There exists an optimal solution in which all capacity purchases and disposals correspond to an integral number of periods.*

According to Corollary 1, purchases and disposals of *excess* capacity correspond to an integral number of periods. We need to show that the same holds for the disposal of *used* capacity (and corresponding purchase) as well. Condition (3.7) together with Corollary 1 imply that demand increments for any period must have been satisfied entirely by capacity of one technology. That is, if a technology is in use, its capacity must be equal to demand increments for an integral number of periods. Now, according to Corollary 2, used technology can only be replaced in its entirety. As a result, any used capacity disposed (and correspondingly replaced) must be equal to demand increments for an integral number of periods.

Corollary 4 *There exists an optimal solution in which capacity meant to satisfy demand increments for earlier periods is purchased before purchasing capacity meant for later periods.*

If this were not true, we would have excess capacity meant for later periods on hand when acquiring capacity for earlier periods, violating condition (3.6).

For similar reasons, we have:

Corollary 5 *There exists an optimal solution in which excess capacity meant to satisfy demand increments for later periods is disposed before disposing capacity meant for later periods.*

Let a period ending with zero excess capacity on hand be referred to as a *regeneration period*. A period in which capacity is acquired to meet future demand increments is referred to as an *acquisition period*. From the above four results, it is clear that capacity of exactly one vintage is acquired in an *acquisition period* to meet the demand increments for all the periods until the *next acquisition period*. That is, all the excess capacity of an earlier vintage must be completely disposed (or consumed) before purchasing capacity of a later vintage.

Note that, if technological breakthroughs were predictable, i.e., if the sequence of discoveries and their timing were all known, it would never be optimal to acquire excess capacity and then dispose it unused at a future time. It is the uncertainty in the nature and timing of breakthroughs that makes disposal of unused capacity possible. However, not all excess capacity need be disposed and the new vintage adopted immediately. For instance, if further breakthroughs are anticipated soon with high probability, the firm may be unwilling to dispose all of the excess capacity and adopt the new vintage immediately. The firm may dispose part of the excess capacity and wait for even better vintages. Corollary 5 provides us some guidance in restricting the choices to be considered in deciding what part of the excess capacity to dispose. More generally, the conditions in Theorem 1 and the corollaries 1 to 5 drastically reduce the choices to be considered in deciding how much and what technology type capacity to purchase, dispose and replace. Finally we have,

Corollary 6 *There exists an optimal solution in which used capacity is never replaced while still holding excess capacity on hand at the same time. All excess capacity must either be disposed or consumed (to meet demand) before used capacity can be replaced.*

Again, this is because replacement involves (disposal followed by immediate) acquisition which cannot occur in a period concurrently with excess capacity from condition (3.6). An implication of Corollary 6 is that we need consider replacing used capacity only in periods where capacity is anyway going to be acquired to meet future demand increments. That is, *replacement of used capacity can occur only in acquisition periods* (note, however, that every acquisition need not be accompanied by replacement). Clearly, this avoids the need for evaluating replacement in periods other than the acquisition periods, leading to reduction in the computational effort.

Now assume that,

$$g_{mnt}(x) \leq h_{mt}(x) + g_{mn,t+1}(x) \quad \forall x, t \text{ and } n > m \quad (3.9)$$

which stipulates that the cost of salvaging an amount of capacity in a period shall not exceed the cost of carrying it to the next period and then salvaging it. Condition (3.9) affirms the lack of speculative motive for holding excess capacity, arising from a later than necessary disposal. Suppose the firm has some excess capacity on hand when a newer technology appears. Consider the choice of disposing a part of this excess capacity. The following proposition stipulates when such a disposal should be carried out.

Proposition 1 *It is optimal to consider the disposal of excess capacity as soon as a new technology become available; any excess capacity to be disposed shall be disposed immediately.*

From condition (3.9), it is clear that postponing the actual disposal of the excess capacity to a period later than the period of appearance of the new technology will only result in higher costs. When a technology appears, (i) the characteristics of the new technology that has appeared relative to the vintage of the excess capacity becomes known, (ii) the probabilistic information about the next innovation becomes available. Both of these pieces of information are not going to change until the next innovation occurs. Based on this information, the amount of excess capacity to be disposed can be determined and this decision is not going to be revised until the next innovation occurs. If we denote a period in which excess capacity is disposed as a *disposal period*, it follows from Proposition 1 that *the disposal period always coincides with the appearance of a new vintage.*

A similar result holds for replacement of used capacity.

Proposition 2 *Any capacity acquired for the sake of replacement is put into use immediately.*

Suppose this is not true, i.e., capacity of vintage m is acquired in period t for the sake of replacing vintage p and is carried as excess capacity until some period τ . Since capacity meant for replacement can not be used to satisfy future demand increments, it can only be used for one of the following two purposes: (i) the excess capacity of technology m is put into use in period τ replacing vintage p , or (ii) the excess capacity is disposed in τ because another vintage (say n) may appear

and it may not be worthwhile to use m at all. In case (i), it would have been better to put m into use immediately because we would have saved carrying costs for vintage m and the difference in operating costs between vintages m and p . Now consider the possibility (ii). It would never be profitable to purchase capacity in a period and then dispose it immediately. From (3.9), the cost of immediate disposal never exceeds the cost of carrying capacity to a future period and then disposing it. Together, the previous two statements imply that case (ii) would not be profitable either.

From the above results, it is clear that the acquisition and disposal periods constitute the only two decision epochs in the model, with replacement of used capacity occurring only in the acquisition periods. Of course, disposal and acquisition periods may coincide. This will happen if in an acquisition period, a new technology becomes available. In the next two sections, we use these observations to develop regeneration based dynamic programming recursions in terms of these two decision epochs.

4 Optimal Capacity Expansion without Replacement

In this section, we present a regeneration point based formulation for the case where capacity, once put into use, is not replaced. Many organizations, for example, let the older personal computer trickle down the organizational hierarchy, as newer models are acquired. Similar trends occur in many other industries where alternative usage for older vintages are identified instead of disposing them. This trickle-down effect may be a result of low salvage value of used capacity, or simply a matter of management policy to relegate the used equipment for training or other purposes. The model proposed in this section serves another purpose. It demonstrates how the properties of the optimal solution derived in the last section can be used to develop an extremely efficient solution procedure. It is much simpler than the model *with replacement of used capacity*, in terms of both exposition and solution efficiency, because a history of technologies and the amount of capacity in-use need not be maintained in this case. Hence, the model here provides a good foundation for the more complex model *with replacement of used capacity*, to be presented in Section 5.

Suppose all capacity related decisions are taken at the beginning of a period. The sequence of

events in any period is as follows. The information regarding current technology becomes available. Based on this information, capacity disposal, if any, takes place followed by any acquisition. The production is then carried out incurring operating expenses. At the end of a period, carrying cost is charged on unused capacity and this excess capacity is made available in its entirety at the beginning of the next period.

The problem can be thought of as a sequence of *acquisition* and *disposal* decisions. The acquisition decisions are made whenever there is no excess capacity (Theorem 1), and disposal of excess or unused capacity is considered only when a new technology appears (Proposition 1). We present a stochastic dynamic programming formulation of the problem in terms of these two decision points. We first consider the optimal acquisition decision followed by the optimal disposal decision.

4.1 Optimal Acquisition Decision

Consider a period i , with no excess capacity on hand. Suppose the current technology is m which was introduced in period k , $k \leq i$. Let $C(m, k, i)$ be the minimum expected total cost for period i onwards, given that current technology m was introduced in period k and the firm has no excess capacity on hand. Since capacity once put into use is never disposed, all future operating expenses are sunk as soon as a technology is put into use. Hence, whenever a capacity is put into use, its operating expenses over the rest of the horizon are immediately accounted. For this reason, we will exclude from our discussion any operating expenses resulting from capacity already in use at the end of period 0. Thus, $C(m, k, i)$ represents the total cost-to-go *excluding* operating expenses due to capacity already in use at the end of period $(i - 1)$, assuming that the firm makes the best acquisition decision now and continues to make decisions optimally in all future periods.

Now consider the immediate acquisition decision. Given the scale economies in capacity acquisition, the firm would like to buy capacity for current as well as a number of future periods (Corollary 2). An impediment to large acquisitions is the carrying charges that the firm must pay on any unused capacity. An even greater deterrent to large acquisition is the uncertainty surrounding possible introduction of a superior technology. Clearly, the firm must weigh the savings derived from buying larger amount of capacity against the carrying charges and potential obsolescence due to a newer, superior technology.

Consider the choice of buying capacity to satisfy demand increments upto period $(j - 1)$. All potential candidates for j ($i < j \leq T + 1$) need to be considered in computing the minimum expected total cost. For notational simplicity, let $d(i, j)$ represent the cumulative demand increments for period i through j i.e.,

$$d(i, j) = \sum_{t=i}^j d_t$$

with the convention that $d(i, j) = 0$ whenever $i > j$. The cost of acquiring capacity of type m which can meet the demand for periods i through $(j - 1)$ is $f_{mi}(d(i, j - 1))$.

Once the firm acquires capacity for periods i through $(j - 1)$, the next acquisition decision is scheduled for period j . The following mutually exclusive scenarios are possible depending upon the timing of the next innovation:

1. No new technological innovation appears until period j , in which case the next acquisition decision will take place as planned in period j , with technology m still the newest available technology. The probability of this event, given that the technology m has already been around for the last $(i - k)$ periods is

$$\frac{1 - Q_m(j - k)}{1 - Q_m(i - k)}$$

where $Q_m(\cdot)$ and $q_m(\cdot)$ denote the cumulative and probability distribution functions respectively of the amount of time elapsed between advent of technology m and the technology following m .

2. A new technology appears before or at the beginning of period j , say in period ν ($i < \nu \leq j$), in which case the firm will have to reconsider its future strategy. The probability that the new technology appears in period ν is

$$\frac{q_m(\nu - k)}{1 - Q_m(i - k)}$$

Scenario 2 represents a variety of possibilities which may lead to the next decision point anywhere on the spectrum between $\nu = (i + 1)$ to $\nu = j$. Also, period j can be the next decision point under scenario 1 or 2, however the newest available technology will be different in each case. The total cost for each scenario comprises of carrying charges and operating expenses until the

next decision point plus the minimum expected cost-to-go from there onwards. First, consider the carrying and operating costs. Let $F(m, i, \nu, j)$ represent the accumulated carrying costs and all sunk operating expenses incurred in periods i thru $(\nu - 1)$ by starting in period i with an amount of capacity of vintage m just enough to satisfy the demand for periods i thru $j - 1$. That is,

$$F(m, i, \nu, j) = \sum_{l=i}^{\nu-1} h_{ml}d(l+1, j-1) + \sum_{\xi=i}^{\nu-1} d_{\xi} \sum_{l=\xi}^T c_{ml} \quad i < \nu \leq j \quad (4.10)$$

with all future operating expenses accounted for as soon as a capacity is put into use. While $F(m, i, \nu, j)$ represents the carrying and operating costs for each possible ν value in scenario 2, these costs for scenario 1 are denoted by $F(m, i, j, j)$. The total cost for scenario 1 is therefore

$$(F(m, i, j, j) + C(m, k, j))$$

Now consider the cost-to-go for each possible decision epoch (ν) for scenario 2. What might a firm do if it has excess capacity on hand when a new technology is introduced? The firm may immediately dispose all the excess capacity and adopt the new technology. Or it may choose to delay the adoption until some future point in time by retaining and continuing to use some or all of the surplus capacity on hand. It is possible, that yet another innovation appears in the meantime and the firm skips a technology in the chain of technological innovations. These possibilities will be considered when we formulate the optimal disposal decision. For now, let $D(m, n, \nu, j)$ denote the minimum expected total cost from period ν onwards, given that a new technology n appeared in period ν and excess capacity of technology m is on hand to satisfy demand increments upto period $(j - 1)$. This again excludes operating expenses due to capacity already in use at the end of period $(\nu - 1)$. Considering all the technologies n that may immediately follow m and weighing the cost-to-go for each possible n with the respective probabilities, we get the cost-to-go for any ν as,

$$\sum_{n(>m)} P_{mn} D(m, n, \nu, j)$$

and the total cost for scenario 2, assuming that the next decision epoch is ν , is given by

$$F(m, i, \nu, j) + \sum_{n(>m)} P_{mn} D(m, n, \nu, j)$$

Weighing the costs for all possible events in both scenarios by their respective probabilities, we obtain the expression for $C(m, k, j)$ in recursive form as

$$C(m, k, i) = \underset{i < j \leq T+1}{\text{Minimize}} \left\{ \begin{aligned} & f_{mi}(d(i, j-1)) + \left(\frac{1-Q_m(j-k)}{1-Q_m(i-k)} \right) (F(m, i, j, j) + C(m, k, j)) \\ & + \sum_{\nu=i+1}^j \left(\frac{q_m(\nu-k)}{1-Q_m(i-k)} \right) \left(F(m, i, \nu, j) + \sum_{n(>m)} P_{mn} D(m, n, \nu, j) \right) \end{aligned} \right. \quad (4.11)$$

Observe that all possible values of j are considered in computing $C(m, k, i)$.

4.2 The Optimal Disposal Decision

We next consider the optimal disposal decision. Consider a period ν when a technology n has just become available. The firm still has excess capacity of technology m available which can satisfy the demand increments for upto period $(j-1)$. Suppose some of the excess capacity is disposed (Corollary 3) and after disposal, the firm is left with enough capacity of technology m to satisfy the demand for upto period $(\tau-1)$. All choices of τ will be considered, including $\tau = \nu$ (disposal of all surplus capacity) and $\tau = j$ (no disposal) and the value that minimizes the expected total cost will be chosen to compute $D(m, n, \nu, j)$. The disposal of unused capacity will incur a cost of $g_{mn\nu}(d(\tau, j-1))$ units, which may be negative.

The next acquisition decision is scheduled for period τ , when the firm will consider the choice of acquiring technology n . Depending upon the timing of the next innovation, the following mutually exclusive scenarios are possible:

1. No new technological innovation appears until period τ , in which case the next acquisition decision will take place as planned in period τ . The probability of this event, given that the technology n has just appeared, is $(1 - Q_n(\tau - \nu))$.
2. A newer technology appears before or in period τ , say in period σ ($\nu < \sigma \leq \tau$), in which case the firm will again reconsider its future strategy by possibly making another disposal decision in period σ (this, in turn, may lead to yet another disposal decision, and so on). The probability of this event is $q_n(\sigma - \nu)$. Note that technology n will be skipped altogether in this case.

Scenario 2 represents a variety of possibilities which may lead to the next decision point anywhere on the spectrum between $\sigma = (\nu + 1)$ to $\sigma = \tau$. Also, period τ can be the next decision point under scenario 1 or 2, however the newest available technology will be different in each case. The total cost for each scenario comprises of carrying charges and operating expenses until the next decision point plus the minimum expected costs-to-go from there onwards. Consider the carrying and operating expenses first. In both the scenarios, the firm starts in period ν with an amount of excess capacity m that can satisfy the demand for upto period $(\tau - 1)$. While in scenario 1 the firm utilizes all of the excess capacity and incurs a cost $F(m, \nu, \tau, \tau)$ in the process, in scenario 2 it stops at σ , with an accumulated cost $F(m, \nu, \sigma, \tau)$. The cost-to-go from period τ onwards in scenario 1 will be $C(n, \nu, \tau)$, making the total cost in this case

$$F(m, \nu, \tau, \tau) + C(n, \nu, \tau)$$

To compute the cost-to-go in scenario 2, suppose the technology that appears immediately following n is of type s (the probability of this happening is P_{ns}). The cost-to-go from period σ onwards is then $D(m, s, \sigma, \tau)$ since the firm still has enough capacity of type m to satisfy the demand for upto period $(\tau - 1)$. Considering all technologies that may immediately follow n and weighing the cost-to-go in each case with their respective probabilities, we obtain the cost-to-go for scenario 2

$$\sum_{s(>n)} P_{ns} D(m, s, \sigma, \tau)$$

and the total cost for scenario 2, assuming that the next decision epoch is σ , is given by

$$F(m, \nu, \sigma, \tau) + \sum_{s(>n)} P_{ns} D(m, s, \sigma, \tau)$$

Weighing the costs for all possible events in both scenarios by their respective probabilities, we obtain the expression for $D(m, n, \nu, j)$ in recursive form

$$D(m, n, \nu, j) = \underset{\nu \leq \tau \leq j}{\text{Minimize}} \begin{cases} g_{mn\nu}(d(\tau, j - 1)) + (1 - Q_n(\tau - \nu))(F(m, \nu, \tau, \tau) + C(n, \nu, \tau)) \\ + \sum_{\sigma=\nu+1}^{\tau} q_n(\sigma - \nu) \left(F(m, \nu, \sigma, \tau) + \sum_{s(>n)} P_{ns} D(m, s, \sigma, \tau) \right) \end{cases} \quad (4.12)$$

Observe that all possible choices of τ are considered in computing $D(\cdot)$. Equations (4.11) and (4.12) together define the optimization problem and can be used to solve the problem in the usual sequential backward fashion. Whenever a new technology, say n , becomes available in a period, say j , which is already scheduled as an acquisition period, the cost-to-go from there onwards can be given by either $C(n, j, j)$ or $D(\cdot, n, j, j)$, both yielding the same value. In other words, the cost obtained from the optimal disposal decision, with no excess capacity on hand to dispose, is the same as that obtained from the optimal acquisition decision, with the same new technology that triggered disposal, i.e.,

$$D(\cdot, n, j, j) = C(n, j, j) \quad \forall n, j \quad (4.13)$$

This relationship can also be verified from the expression (4.12).

5 Optimal Capacity Expansion with Replacement

When replacement of used capacity is allowed, one must consider all technologies in use as a candidate for replacement. For this reason, let us define the history of all past acquisitions by a sequence $\{u(1), u(2), \dots, u(p)\}$, where $u(\xi)$ represents the period in which ξ th acquisition was made and variable p represents the number of acquisitions made so far. A related string is technology-mix sequence $\{w(1), w(2), \dots, w(p)\}$, where $w(\xi)$ represents the technology type acquired in period $u(\xi)$. As time progresses and more capacity is acquired, both acquisition history and technology-mix sequences are updated by appending the acquisition period and technology type to the respective sequences. To simplify the notation, we will often use $\langle u \rangle_{\xi=1}^p$ to represent the sequence $\{u(1), u(2), \dots, u(p)\}$.

To keep the exposition uncluttered, we assume that initially (at the beginning of period 1), there is capacity of only one technology type, say type 1. This capacity consists of D_o units already in use and, possibly, excess capacity to satisfy the demand increments for an integer number of future periods. These assumptions are incorporated by initial conditions $u(1) = 1$, and $w(1) = 1$. More general initial conditions can be incorporated by simply redefining $u(1)$ and $w(1)$. Let \bar{d}_ξ be the sum of demand increments in periods $u(\xi)$ through $u(\xi + 1) - 1$, i.e.,

$$\bar{d}_\xi = d(u(\xi), u(\xi + 1) - 1)$$

except for \bar{d}_1 which also includes the capacity initially in use, i.e.,

$$\bar{d}_1 = d(u(1), u(2) - 1) + D_o$$

Clearly, the cumulative demand, \bar{d}_ξ , is satisfied by technology type $w(\xi)$ for all ξ .

An Example: We now illustrate through a sequence of hypothetical decisions how acquisition history and technology-mix sequences evolve. The decisions, though hypothetical, conform to the properties of the optimal solution specified in Section 3 and as such, are illustrative of the sequence of decisions resulting from the proposed model. Consider the acquisition of successive generations of microprocessors by a firm which initially starts with enough 8086 chips to satisfy the demand increments for up to period 4. The sequence of innovations in microprocessor chips — 8086, 286, 386, 486, etc. are labelled, respectively, as technology types 1, 2, 3, 4, etc. The sequence of decisions are summarized in Table 1. Recall that decisions are triggered either by the advent of a new technology or by lack of excess capacity. The next acquisition decision is thus scheduled for period 5. Suppose the 286 chip becomes available at the beginning of period 3 when the firm still has enough 8086 chips to satisfy demand until period 4. The firm must reconsider its decision due to availability of a new product. The firm, after reconsidering this new information, decides not to adopt the new technology at this time. No new vintage appears in the next two periods (4 and 5). At the beginning of period 5, the next scheduled acquisition point, the firm decides to buy enough 286 chips to satisfy the demand for the next six periods (until 10). In period 6, however, 386 chips become available, and the firm must again reconsider its future strategy. After considering all factors, it decides to dispose some excess capacity (those meant to satisfy the demand increment for period 10) and retain the rest. The next acquisition decision is thus scheduled for period 10 when the firm plans to acquire 386 chips. However, another discovery occurs in period 7 and a superior microprocessor, 486, becomes available. At this point the firm decides to dispose all excess capacity of 286 chips but those already in use are not replaced. A decision to buy 486 chips to satisfy the demand for periods 7 through 10 is also taken. No new product appears until the next acquisition point, period 11. At this point, the firm decides to get rid of all 286 chips already in use (Corollary 5), replace it with 486 chips, and buy extra 486 chips to meet the demand for periods 11 through 14. No new technology appears until period 15, which is again an acquisition period and the sequence of innovations, acquisitions, and replacements continues.

Table 1: A Summary of Decisions

t	Event that triggered the decision-making	Acquisition history, $\langle u \rangle$	Technology-mix, $\langle w \rangle$	Decision taken
3	Advent of 286 chip (technology type 2)	{1}	{1}	Do not dispose excess 8086 chips.
5	Acquisition period; no excess capacity	{1}	{1}	Buy enough 286 chips to satisfy demands for periods 5 through 10.
6	Advent of 386 chip (technology type 3)	{1,5}	{1,2}	Dispose excess capacity of 286 chip meant to satisfy demand for period 10.
7	Advent of 486 chip (technology type 4)	{1,5}	{1,2}	Dispose all excess capacity of 286 chip. Buy enough 486 chips to satisfy demand for periods 7 through 10.
11	Acquisition period; no excess capacity	{1,5,7}	{1,2,4}	Dispose all 286 chips in use. Buy enough 486 chips to replace them (demand for periods 5 and 6) as well as to meet demand for periods 11 through 14.
15	Acquisition period; no excess capacity	{1,5,7,11}	{1,4,4,4}	Under consideration.

The sequence of events and the decisions made in this example illustrate several interesting aspects of the model. First, the advent of a superior technology does not imply its immediate adoption if the firm has excess capacity of an older vintage. In the example, the firm continued to use 8086 chips, as planned, despite the availability of 286 chips in period 3. Similarly, the firm made only a minor revision in its plan to use 286 chips to meet demand increments until period 10 when 386 chips appeared on the market. On the other hand, the firm immediately embraced 486 chips when it first appeared in period 7. The choice of whether to dispose all the excess capacity and adopt the new technology immediately or, to continue using the excess capacity of the older vintage and thus delay the adoption, depends on a number of factors including salvage cost of the excess capacity, relative merit of the newest technology compared to the old one, and the likelihood of an even better forthcoming technology. An optimal technology path may not follow the chain of technological development link by link. For example, a decision to delay the adoption of 386 chips led to its exclusion altogether due to advent of an even better technology — 486 chips. Second, recall that used capacity may be replaced only in acquisition periods (Corollary 4). Not

every acquisition, however, is accompanied by a replacement. In the example, acquisition occurs in periods 5, 7, and 11 but replacement is done only in period 11. \square

The problem can be thought of as a sequence of acquisition, replacement, and disposal decisions. Disposal of excess capacity is considered only when a new technology appears (Proposition 1). On the other hand, acquisition and replacement are considered whenever the firm has no excess capacity (Corollary 4). We provide below a dynamic programming formulation in terms of these two decision epochs. We first consider the optimal acquisition and replacement decisions followed by the optimal disposal decision. It is possible that the two decision epochs coincide i.e., a disposal decision is followed by an acquisition (and possibly replacement) decision in the same period. This was the case in period 7 in the example.

5.1 Optimal Acquisition and Replacement Decision

Consider a period i with no excess capacity on hand. Let $\langle u \rangle_{\xi=1}^p$, and $\langle w \rangle_{\xi=1}^p$ represent the acquisition history and technology-mix sequences, respectively, until period i . Suppose the newest technology is m which was introduced in period k , $k \leq i$. Let $C(m, k, i, \langle u \rangle_{\xi=1}^p, \langle w \rangle_{\xi=1}^p)$ be the minimum expected total cost from period i onwards, given that the firm has no excess capacity on hand, the newest technology is m which was first introduced in period k and the acquisition history and technology-mix sequences until this point in time are $\langle u \rangle_{\xi=1}^p$, and $\langle w \rangle_{\xi=1}^p$ respectively. $C(\cdot)$ is the total cost incurred from now on, assuming that the firm makes the best acquisition and replacement decisions now and continues to make decisions optimally in all future periods.

Consider the issue of replacement first. The firm has the choice of replacing one or more technologies adopted in the past by the current best technology, m . Suppose $\langle \bar{w} \rangle_{\xi=1}^p$ is the technology-mix sequence after the replacement. All choices of $\langle \bar{w} \rangle_{\xi=1}^p$ will be considered such that

$$\bar{w}(\xi) = w(\xi) \text{ or } m; \xi = 1, 2, \dots, p. \quad (5.14)$$

Condition (5.14) stipulates that for each technology, we have the choice of either replacing it with the current best technology or continuing without replacement. Recall that if we choose to replace a particular technology, we replace all capacity corresponding to that vintage (Corollary 5). These

conditions, in effect, reduce the choice of $\langle \bar{w} \rangle_{\xi=1}^p$ significantly, making the dynamic programming solution extremely efficient. Let $\bar{\delta}_\xi$ be a *replacement indicator* defined as

$$\bar{\delta}_\xi = \delta(\bar{w}(\xi) - w(\xi))$$

where $\delta(x)$ is the Kronecker delta function given by

$$\delta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$$

The replacement indicator, $\bar{\delta}_\xi$, takes the value 1 or 0, depending upon whether technology $w(\xi)$, satisfying the cumulative demand \bar{d}_ξ , is replaced or not. Overall, the replacement will involve disposal of $\sum_{\xi=1}^p \bar{d}_\xi \bar{\delta}_\xi$ units of old capacities which will incur a cost $\sum_{\xi=1}^p \bar{g}_{w(\xi), m, i}(\bar{d}_\xi \bar{\delta}_\xi)$ (possibly negative). Recall that replacement always coincides with an acquisition period, wherein capacity is purchased to meet future demand increments. Suppose the firm decides to acquire $d(i, j-1)$ units of capacity to satisfy demand increments in periods i through $(j-1)$, where $(i < j \leq T+1)$. Then, the acquisition of capacity for future demand and replacement of used capacity together incur an immediate cost

$$f_{mi} \left(d(i, j-1) + \sum_{\xi=1}^p \bar{d}_\xi \bar{\delta}_\xi \right) + \sum_{\xi=1}^p \bar{g}_{w(\xi), m, i}(\bar{d}_\xi \bar{\delta}_\xi)$$

As a result of these decisions, the new acquisition history and technology mix sequences are now $\{\langle u \rangle_{\xi=1}^p, i\}$ and $\{\langle \bar{w} \rangle_{\xi=1}^p, m\}$ respectively (we will often indicate the last element of a sequence explicitly for clarity; $\{\langle u \rangle_{\xi=1}^p, i\}$ is equivalent to $\langle u \rangle_{\xi=1}^{p+1}$ with $u(p+1) = i$). The next acquisition decision is scheduled for period j . Depending upon the timing of the next innovation, as in Section 4, the following mutually exclusive scenarios are possible:

1. No new technological innovation appears until period j , in which case the next acquisition (and possibly replacement) decisions will take place as scheduled in period j . The probability of this event is

$$\frac{1 - Q_m(j-k)}{1 - Q_m(i-k)}$$

2. A new technology appears in or before the next scheduled acquisition period j , say in period ν ($i < \nu \leq j$), in which case the firm will have to reconsider its future strategy. The probability

that the new technology appears in period ν is

$$\frac{q_m(\nu - k)}{1 - Q_m(i - k)}$$

The total cost for each scenario comprises of carrying charges and operating expenses until the next decision point (which is j or ν , as the case maybe) plus the minimum expected cost-to-go from there onwards. First, we consider the carrying and operating costs. Let $G(m, i, \nu, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p)$ represent the accumulated carrying costs and operating expenses incurred over periods i through $(\nu - 1)$ by starting in period i with acquisition history $\langle u \rangle_{\xi=1}^p$, technology-mix $\langle \bar{w} \rangle_{\xi=1}^p$ and an amount of capacity m sufficient to exactly satisfy the demand increments in periods i thru $(j - 1)$. That is,

$$G(m, i, \nu, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p) = \sum_{l=i}^{\nu-1} h_{ml} d(l+1, j-1) + \sum_{\xi=i}^{\nu-1} d_{\xi} \sum_{l=\xi}^{\nu-1} c_{ml} + \sum_{\xi=1}^p \bar{d}_{\xi} \sum_{l=i}^{\nu-1} c_{\bar{w}(\xi), l} \quad i < \nu \leq j \quad (5.15)$$

The cost function $G(\cdot)$ is similar to function $F(\cdot)$ defined in the last section, except for the following two aspects. First, in contrast to $F(\cdot)$, there is an additional term (the last term in 5.15) in $G(\cdot)$ corresponding to the operating expenses for used capacity satisfying all demand increments prior to period i . Second, $F(\cdot)$ accounted for all future operating expenses the moment a capacity was put into use, since replacement of used capacity was not an issue. On the other hand, $G(\cdot)$ accounts for operating expenses only until the next decision point (ν). Hence, while it was unnecessary to maintain history of past adoptions for computing $F(\cdot)$, the possibility of replacement here makes it necessary to maintain the acquisition history in computing $G(\cdot)$. The carrying and operating expenses for scenario 1 can be computed simply by letting $\nu = j$ in the expression (5.15), which otherwise gives these costs for each possible event in scenario 2.

The minimum expected cost-to-go in scenario 1 is $C(m, k, j, \{\langle u \rangle_{\xi=1}^p, i\}, \{\langle \bar{w} \rangle_{\xi=1}^p, m\})$. The minimum expected cost-to-go in scenario 2 can be computed by conditioning on the next technological breakthrough following m . Say, n is the new technology that appears in period ν when the firm still has excess capacity of technology m . Suppose $D(n, \nu, j, \{\langle u \rangle_{\xi=1}^p, i\}, \{\langle \bar{w} \rangle_{\xi=1}^p, m\})$ is the minimum expected total cost for period ν onwards, given that a new technology n was introduced this period, the acquisition history and technology mix sequences until this point in time are $\{\langle u \rangle_{\xi=1}^p, i\}$, and $\{\langle \bar{w} \rangle_{\xi=1}^p, m\}$ respectively, and excess capacity from the last acquisition (of

technology m) can satisfy demand increments until period $(j-1)$. The computation of $D(\cdot)$ involves optimal disposal decision which we will consider soon. For any ν , the cost-to-go in scenario 2 is given by

$$\sum_{n(>m)} P_{mn} D\left(n, \nu, j, \left\{ \langle u \rangle_{\xi=1}^p, i \right\}, \left\{ \langle \bar{w} \rangle_{\xi=1}^p, m \right\}\right)$$

Then, the total cost for scenario 1 is

$$G\left(m, i, j, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p\right) + C\left(m, k, j, \left\{ \langle u \rangle_{\xi=1}^p, i \right\}, \left\{ \langle \bar{w} \rangle_{\xi=1}^p, m \right\}\right)$$

and the total cost for each possible ν value in scenario 2 is

$$G\left(m, i, \nu, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p\right) + \sum_{n(>m)} P_{mn} D\left(n, \nu, j, \left\{ \langle u \rangle_{\xi=1}^p, i \right\}, \left\{ \langle \bar{w} \rangle_{\xi=1}^p, m \right\}\right)$$

Weighing the cost for all possible events in both scenarios with the respective probabilities, we obtain the expression for $C(\cdot)$ in recursive form,

$$C\left(m, k, i, \langle u \rangle_{\xi=1}^p, \langle w \rangle_{\xi=1}^p\right) = \begin{cases} f_{mi} \left(d(i, j-1) + \sum_{\xi=1}^p \bar{d}_{\xi} \bar{\delta}_{\xi} \right) + \sum_{\xi=1}^p g_{w(\xi), m, i} (\bar{d}_{\xi} \bar{\delta}_{\xi}) \\ + \left(\frac{1-Q_m(j-k)}{1-Q_m(i-k)} \right) \left(G\left(m, i, j, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p\right) \right. \\ \left. + C\left(m, k, j, \left\{ \langle u \rangle_{\xi=1}^p, i \right\}, \left\{ \langle \bar{w} \rangle_{\xi=1}^p, m \right\}\right) \right) \\ + \sum_{\nu=i+1}^j \left(\frac{q_m(\nu-k)}{1-Q_m(i-k)} \right) \left(G\left(m, i, \nu, j, \langle u \rangle_{\xi=1}^p, \langle \bar{w} \rangle_{\xi=1}^p\right) \right. \\ \left. + \sum_{n(>m)} P_{mn} D\left(n, \nu, j, \left\{ \langle u \rangle_{\xi=1}^p, i \right\}, \left\{ \langle \bar{w} \rangle_{\xi=1}^p, m \right\}\right) \right) \end{cases} \quad (5.16)$$

Minimize
 $j, \langle \bar{w} \rangle$
 $i < j \leq T+1$
 $\bar{w}(\xi) = w(\xi) \text{ or } m; \forall \xi$

Note that all possible choices of j ($i < j \leq T+1$) and $\langle \bar{w} \rangle_{\xi=1}^p$ have been considered in computing $C(\cdot)$, the minimum expected total cost.

5.2 Optimal Disposal Decision

Disposal is triggered by the advent of a new technology. Consider a period ν when a new technology n has just become available. The acquisition history and technology-mix until this point in time

are given by sequences $\langle u \rangle_{\xi=1}^{p+1}$ and $\{\langle w \rangle_{\xi=1}^p, m\}$ respectively. The firm still has enough excess capacity of the last acquired technology, m , to satisfy demand increments until period $(j - 1)$. We are interested in evaluating the cost-to-go $D(n, \nu, j, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\})$. Suppose, after the disposal, the firm is left with enough capacity of technology m to satisfy the demand increments until period $(\tau - 1)$. All choices of τ will be considered including $\tau = \nu$ (disposal of all surplus capacity) and $\tau = j$ (no disposal) and the value that minimizes the expected total cost will be chosen to compute $D(\cdot)$.

First, the disposal of unused capacity will immediately incur a cost $g_{m\nu}(d(\tau, j - 1))$ (possibly negative). The next acquisition and replacement decisions are scheduled for period τ . As before, depending on the timing of the next innovation, the following mutually exclusive scenarios are possible:

1. No new technological innovation appears until period τ , in which case, the next acquisition and replacement decisions will take place, as planned, in period τ with technology n still the newest available technology. The probability of this event, given that the technology n has just appeared, is $(1 - Q_n(\tau - \nu))$.
2. A newer technology appears in or prior to τ , say in period σ ($\nu < \sigma \leq \tau$), in which case the firm will have to reconsider its future strategy in period σ (This, in turn, may lead to yet another disposal decision, and so on). The probability of this event is $q_n(\sigma - \nu)$.

The total cost for each scenario comprises of carrying charges and operating expenses until the next decision point (τ or σ) plus the minimum expected cost-to-go from there onwards. Consider the carrying and operating expenses first. The firm starts in period ν with an amount of capacity of vintage m just enough to satisfy the demand for periods ν through $(\tau - 1)$ and with an acquisition history sequence $\langle u \rangle_{\xi=1}^{p+1}$ and technology mix $\langle w \rangle_{\xi=1}^{p+1}$. For scenario 2, carrying and operating expenses until the next decision point, σ , are to be considered, and are therefore given by $G(m, \nu, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1})$. For scenario 1, the next decision point is τ , and so the carrying and operating expenses are given by $G(m, \nu, \tau, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1})$.

We consider the cost-to-go for each scenario next. In case of scenario 1, the firm will find itself in period τ with no excess capacity on hand; technology n , which first appeared in period ν , will still be the newest available technology; and the acquisition history and technology-mix sequences

will still be the same, $\langle u \rangle_{\xi=1}^{p+1}$ and $\langle w \rangle_{\xi=1}^{p+1}$ respectively. As a result, the cost-to-go from period τ onwards will be $C(n, \nu, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1})$. To evaluate the cost-to-go for scenario 2, suppose that the new technology that appears in period σ is of type s (the probability of this happening is only P_{ns}). The firm will find itself in period σ with enough excess capacity of the last acquired technology, m , to satisfy demand for upto period $(\tau - 1)$. The acquisition history and technology mix sequences will still be $\langle u \rangle_{\xi=1}^{p+1}$ and $\{\langle w \rangle_{\xi=1}^p, m\}$ respectively. The cost-to-go in this case will be $D(s, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\})$. Considering all technologies that may follow n and weighing the cost-to-go in each case with their respective probabilities, we get the cost-to-go for scenario 2,

$$\sum_{s(>n)} P_{ns} D(s, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\})$$

The total cost for scenario 1 is

$$G(m, \nu, \tau, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1}) + C(n, \nu, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1})$$

while that for scenario 2 is

$$G(m, \nu, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1}) + \sum_{s(>n)} P_{ns} D(s, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\})$$

Weighing the costs for all possible events in each scenario with their respective probabilities, we obtain the expression for $D(\cdot)$ in recursive form,

$$D(n, \nu, j, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\}) = \underset{\nu \leq \tau \leq j}{\text{Minimize}} \begin{cases} g_{m,n,\nu}(d(\tau, j-1)) \\ + (1 - Q_n(\tau - \nu)) \left(G(m, \nu, \tau, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1}) \right. \\ \left. + C(n, \nu, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1}) \right) \\ + \sum_{\sigma=\nu+1}^{\tau} q_n(\sigma - \nu) \left(G(m, \nu, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \langle w \rangle_{\xi=1}^{p+1}) \right. \\ \left. + \sum_{s(>n)} P_{ns} D(s, \sigma, \tau, \langle u \rangle_{\xi=1}^{p+1}, \{\langle w \rangle_{\xi=1}^p, m\}) \right) \end{cases} \quad (5.17)$$

Observe that all possible choices of τ are considered in computing $D(\cdot)$. Equations (5.16) and (5.17) together define the optimization problem and can be used to solve the problem in the usual sequential backward fashion. Note that the length of sequences $\langle u \rangle$ and $\langle w \rangle$ grow only due to acquisitions; disposals have no effect on them.

A relationship analogous to (4.13) in this case would be

$$D(n, j, j, \langle u \rangle, \langle w \rangle) = C(n, j, j, \langle u \rangle, \langle w \rangle) \quad \forall n, j, \langle u \rangle, \langle w \rangle \quad (5.18)$$

which can be verified from the expression (5.17).

The regeneration point based formulation proposed in this section can be made even more efficient by further reduction in state space. For example, whenever two consecutive elements of technology-mix sequence are the same, one can simply delete the second element from both the sequences $\langle u \rangle$, and $\langle w \rangle$. In Table 1, the technology-mix and acquisition history sequences for period 15 can thus be represented simply by $\{1, 4\}$, and $\{1, 5\}$ respectively. In fact, since capacity in use cannot be distinguished from one another except by their technology type, one needs to keep track of only what kind of technologies are currently in use and in what amount. The detailed acquisition history and technology-mix sequences are unnecessary and were used in this section only for ease of exposition. The number of technologies in use at any point will be much smaller than the potential number of acquisitions until that point, leading to substantial reduction in state space.

6 Computational Experience

In the past two sections, we presented stochastic dynamic programming recursions for the models with and without replacement of used capacity. We were able to obtain efficient regeneration point formulations for an otherwise intractable problem. For the model *with* replacement of used capacity, without using any of the results of Section 3, the computational effort would have been approximately of the order 2^{TM^D} , where D is the total capacity requirement at the end of period T and M is the number of possible technology levels that may appear over the next T periods. Clearly the computational effort is super-exponential. However, using the results, the computational effort is drastically reduced and is of the order $T^{M+4}M^52^M$. The computational effort is exponential even after using the results from Section 3, but it is reasonable for say 4 or 5 possible technology levels. For the model *without* replacement of used capacity (Section 4), the computational effort is only of the order M^2T^5 , which is moderate.

To assess the computational time, we implemented the recursions for both the models without any special data structures or coding schemes. The computations were run on an IBM-compatible personal computer based on a 486 chip running at 33MHz clock speed (almost the *current best technology* in personal computers at the time of writing). For the model without replacement of used capacity, the computation time for problems with 5 technology levels and 12 periods averaged 2 seconds while the computation time for 20 periods averaged 13 seconds. For the model with replacement of used capacity, the computation time for problems with 12 periods and 4 technology levels averaged 460 seconds (less than 8 minutes). With 4 technology levels, the computational time for problems with 6, 8, and 10 periods, respectively, were 7, 35 and 135 seconds. These computation times appear to be quite reasonable, particularly since capacity related decisions are not taken in real time.

7 Conclusions

In this paper we have studied a fairly comprehensive model of capacity and technology acquisition and replacement in environments where successive technological breakthroughs take place stochastically, leading to frequent and uncertain obsolescence of production technologies. We modeled this problem and presented some key properties of the model that are used to reduce the state space and computational effort to solve this otherwise intractable problem. In particular, we showed that it is optimal to purchase, dispose and replace capacity in amounts equal to the demand increments for an integral number of periods. Further, we showed that it is optimal to: (i) dispose excess capacity only in periods when a new technology appears, and (ii) replace used capacity only in *acquisition periods* i.e., periods when capacity is anyway going to be acquired to meet future demand increments. Using these results, we developed regeneration point based dynamic programming recursions for the models *with* and *without* replacement of used capacity. Finally, we showed that it is possible to solve moderate size problems on a personal computer in a reasonable amount of time.

We are currently in the process of developing a decision support system for a firm that leases MRI and other medical equipments to hospitals. The firm faces capacity and technology acquisition and replacement decisions on an ongoing basis. Our immediate plan is to assess as to what degree the model developed here captures all the key aspects of the problem faced by this firm and consider

the appropriate model changes or extensions. Second, we have not considered revenue effects of technology acquisition in this paper. This should also provide much scope for further research.

Acknowledgements

We would like to thank Mr. Ron Van Horssen at Mobile Technology Inc. for motivating us to research this problem and Professor Thomas Morton at Carnegie-Mellon University for his many helpful comments and suggestions.

References

- [1] Balcer, Y. and S. A. Lippman. 1984. Technological Expectations and Adoption of Improved Technology, *Journal of Economic Theory*, **34**, 292–318.
- [2] *Business Week*, December 10, 1990, Science and Technology, 185–186.
- [3] Chand, S., and S. Sethi. 1982. Planning Horizon Procedures for Machine Replacement Models with Several Possible Replacement Alternatives. *Naval Research Logistics Quarterly*, **29**, 483–493.
- [4] Chand, S., and T. E. Morton. 1982. A Perfect Planning Horizon Procedure for a Deterministic Cash Balance Problem. *Management Science*, **6**, 652–669.
- [5] Cohen, M. A., and R. M. Halperin. 1986. Optimal Technology Choice in a Dynamic Stochastic Environment. *Journal of Operations Management*, **6**, 317–331.
- [6] Dantzig, G. 1955. Optimal Solution to a Dynamic Leontief Model with Substitution. *Econometrica*, **23**, 295–302.
- [7] Erlenkotter, D. 1969. Pre-investment Planning for Capacity Expansion: A Multi-Location Dynamic Model. *Ph.D. Dissertation*, Grad. School of Business, Stanford University, Stanford, California.
- [8] Freidenfelds, J. 1981. *Capacity Expansion: Analysis of Simple Models with Applications*. Elsevier, New York.
- [9] Gaimon, C. 1989. Dynamic Game Results of the Acquisition of New Technology, *Operations Research*, **37**, 3, 410–425.
- [10] Goldstein, T., Ladany, S. P. and A. Mehrez. 1988. A Discounted Machine Replacement Model with an Expected Future Technological Breakthrough, *Naval Research Logistics*, **35**, 209–220.
- [11] Hinomoto, H. 1965. Capacity Expansion with Facilities under Technological Improvement. *Management Science*, **11**, 581–592.
- [12] Jones, P. C., Zydiak, J. L., and W. J. Hopp. 1991. Parallel Machine Replacement. *Naval Research Logistics*, **38**, 351–365.

- [13] Klineciewicz, J. G. and H. Luss. 1985. Optimal Timing Decisions for the Introduction of New Technologies, *European Journal of Operations Research*, **20**, 211–220.
- [14] Li, S., and D. Tirupati. 1990. Dynamic Capacity Expansion Problem with Multiple Products: Technology Selection and Timing of Capacity Additions. *Working Paper 89/90-4-12*, Graduate School of Business, The University of Texas at Austin, Austin, TX.
- [15] Love, S. F. 1973. Bounded Production and Inventory Models with Piecewise Concave Costs. *Management Science*, **20**, 313–318.
- [16] Luss, H. 1982. Operations Research and Capacity Expansion Problems: A Survey. *Operations Research*, **30**, 907–947.
- [17] Lee, S. B., and Luss, H. 1986. Multi-facility type Capacity Expansion Planning: Algorithms and Complexities. *Operations Research*, **35**, 249–253.
- [18] Monahan, G. E. and T. L. Smunt. 1989. Optimal Acquisition of Automated Flexible Manufacturing Process, *Operations Research*, **37**, 2, 288–300.
- [19] Morton, T. E. 1981. Forward Algorithms for Forward-thinking Managers. Applications of *Management Science*, **1**, 1–55.
- [20] Nair, S., and W. J. Hopp. 1989. A Model for Equipment Replacement due to Technological Obsolescence. *Working Paper*, Dept. of IE and Management Sciences, Northwestern University, Chicago, IL.
- [21] Pierskalla, W., and J. Voelker. 1976. A Survey of Maintenance Model: Control and Surveillance of Deteriorating Systems. *Naval Research Logistics Quarterly*, **23**, 353–388.
- [22] Ross, S. M. 1983. *Introduction to Stochastic Dynamic Programming*. Academic Press, New York.
- [23] Sethi, S., and S. Chand. 1979. Planning Horizon Procedures for Machine Replacement Models. *Management Science*, **25**, 140–151.
- [24] Standard and Poor's Industry Surveys, *Electronics*, (May 24, 1990), E 15–17, E 34–36.
- [25] Van Horssen, R., CEO, Mobile Technology Inc., Los Angeles, CA, *Personal communication*, 1991.
- [26] Veinott, Jr., A. F. 1968. Extreme Points of Leontief Substitution Systems. *Linear Algebra and Its Applications*, **1**, 181–194.
- [27] Veinott, A. F., Jr. 1969. Minimum Concave Cost Solution of Leontief Substitution Models of Multi-facility Inventory Systems. *Operations Research*, **17**, 262–291.
- [28] Wagner, H. M., and T. Whitin. 1958. Dynamic Version of the Economic Lot Size Model. *Management Science*, **5**, 89–96.
- [29] Zangwill, W. I. 1968. Minimum Concave Cost Flows in Certain Networks. *Management Science*, **14**, 429–450.