THE DYNAMICS OF SCALE, TECHNOLOGICAL SUBSTITUTION AND PROCESS MIX

John A. Buzacott
Kiichiro Tsuji

July 1980
WP-80-117

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

At the workshop on "Size and Productive Efficiency--The Wider Implications" held at IIASA in June 1979 there was a great deal of discussion on the dynamics of scale, with particular focus on scale, technology and the learning curve, scale and innovation and the effect of uncertainty about the future on scale decisions.

This paper reports the results of research on using formal models of the decision on process and scale in order to understand the dynamics of change in scale and process mix.

CONTENTS

The Dynamics of Scale, Technological Substitution and Process Mix

John A. Buzacott
Kiichiro Tsuji

INTRODUCTION

Characteristic of the dynamic behavior of many industries is the way plant size increases with time. There are many examples of this. Simmonds investigated the increase in maximum size of plant in various processes since the time of its original development. Figure 1 illustrates his data. Other examples are the increase in the maximum size of blast furnace (Figure 2), the increase in maximum size of vessel in the basic oxygen process (Figure 3) and the increase in the size of nuclear generating units (Figure 4).

Another characteristic of the dynamic behavior is the way in which, as new production processes appear, the mix of production from the various process changes over time. For exsmple, Figure 5 shows the amount of steel produced by various processes in the Ruhr district of Germany. Figure 6 shows the proportion of steel produced by various processes in Japan. Figure 7 shows the mix of electricity generation in the U.S. according to type of fuel used.

This dynamic behaivor of industry is of interest for a variety of reasons. The policy analyst concerned with the future development of industry would like to discover whether the behavior demonstrates regular patterns. If so, the identification of the underlying pattern would enable him to make meaningful projections about the future. Next, because of the wider implications of the behavior it might be considered desirable to change it or modify it in some way. An understanding of the factors determining the behavior might provide a means of doing so. Then there are more specific questions like: what would happen if maximum plant size is limited to reduce environmental impact? or what would happen if a new process is invented and developed?

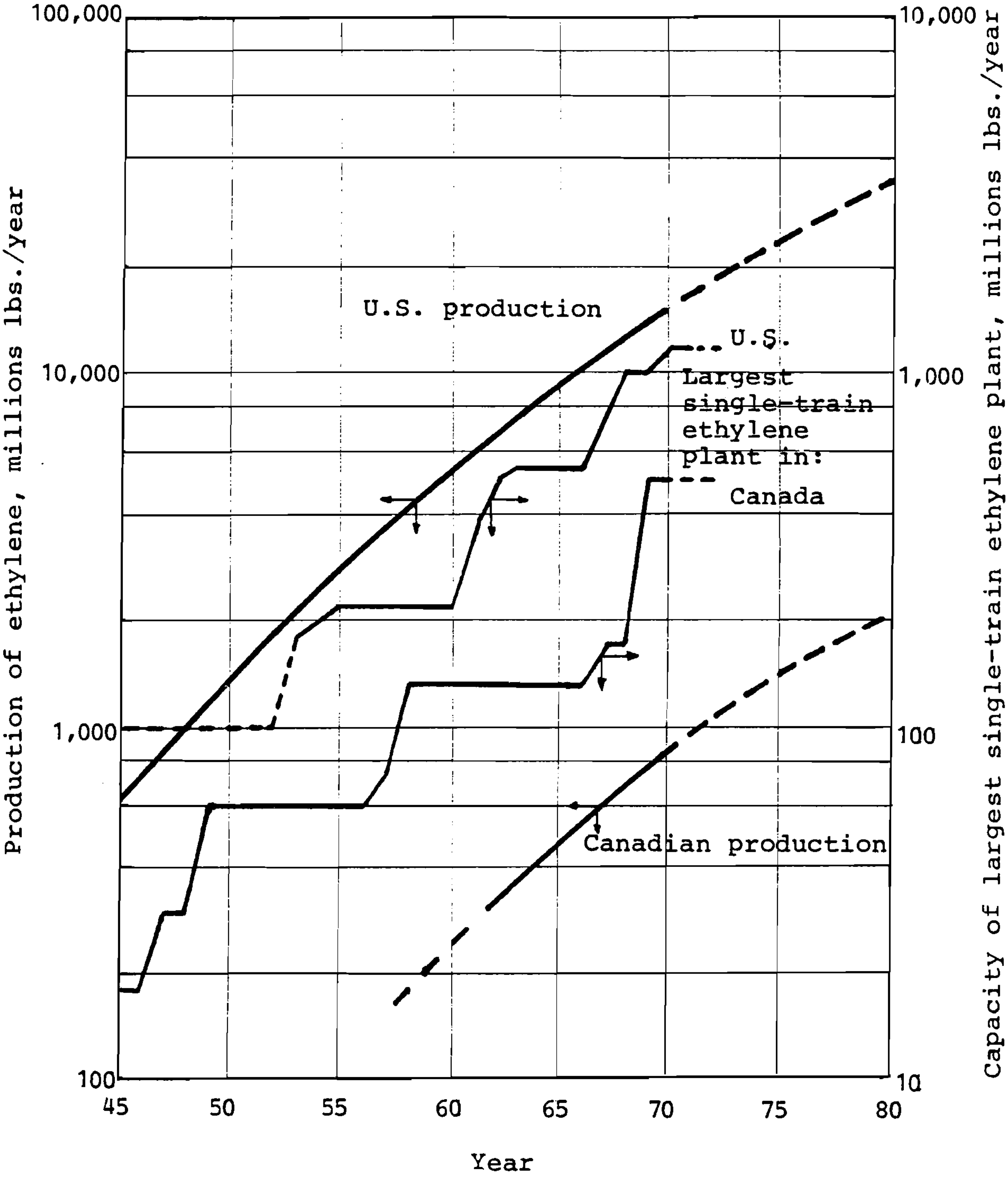


Figure 1. Relation Between Largest Plant Size and Production in Canada and the United States--Ethylene (source: Simmonds 1972)

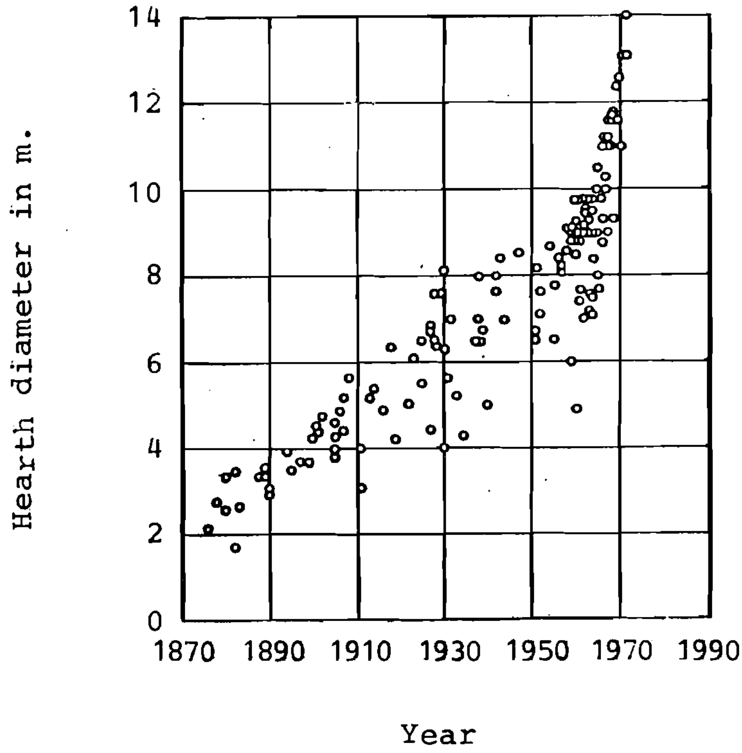


Figure 2. Development of the Hearth Diameter of Blast Furnaces (source: Harders 1971)

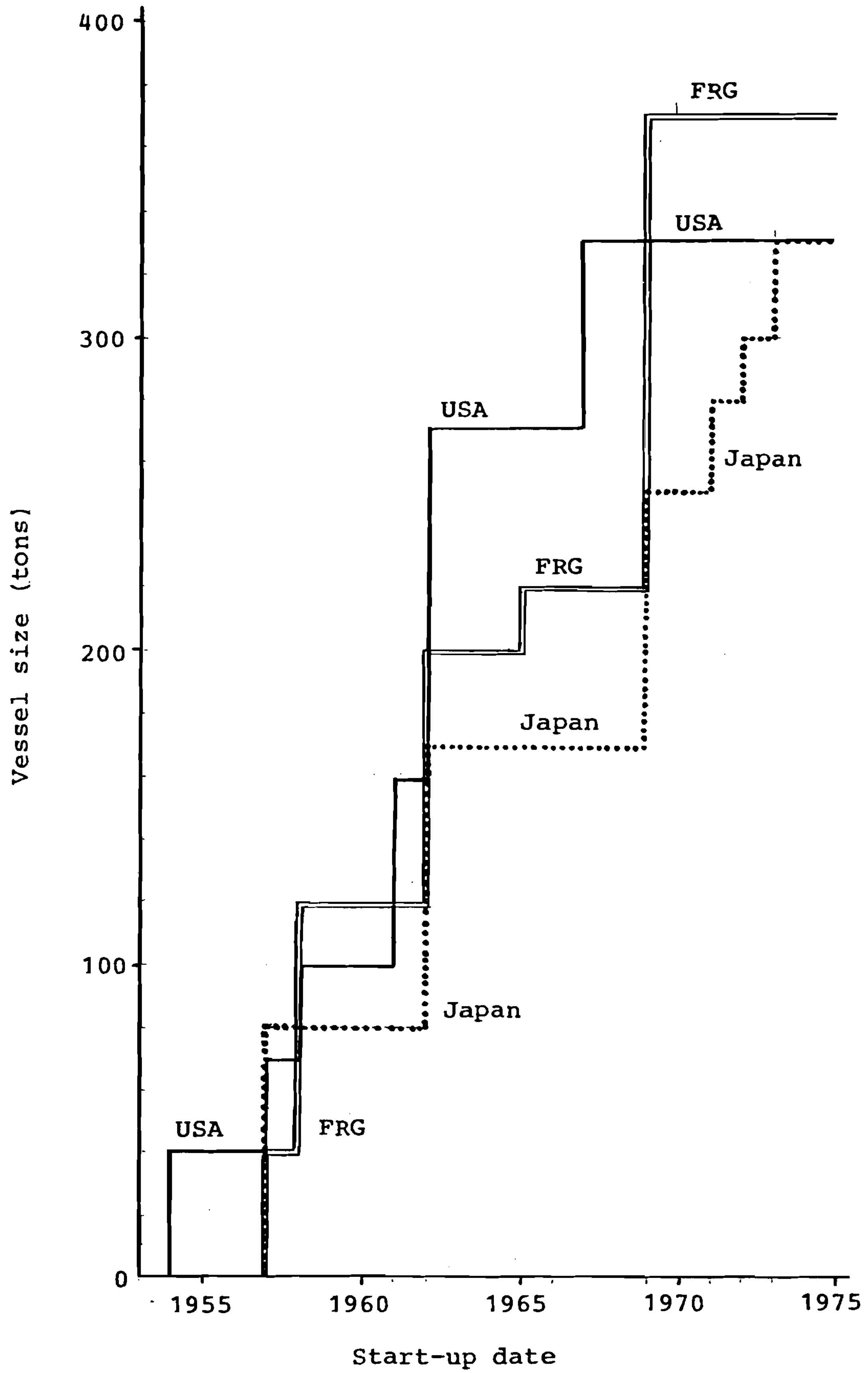


Figure 3. Growth of Largest Size of BOP Vessel for USA, Japan and FRG (Source: Buzacott 1980 based on Resch 1973)

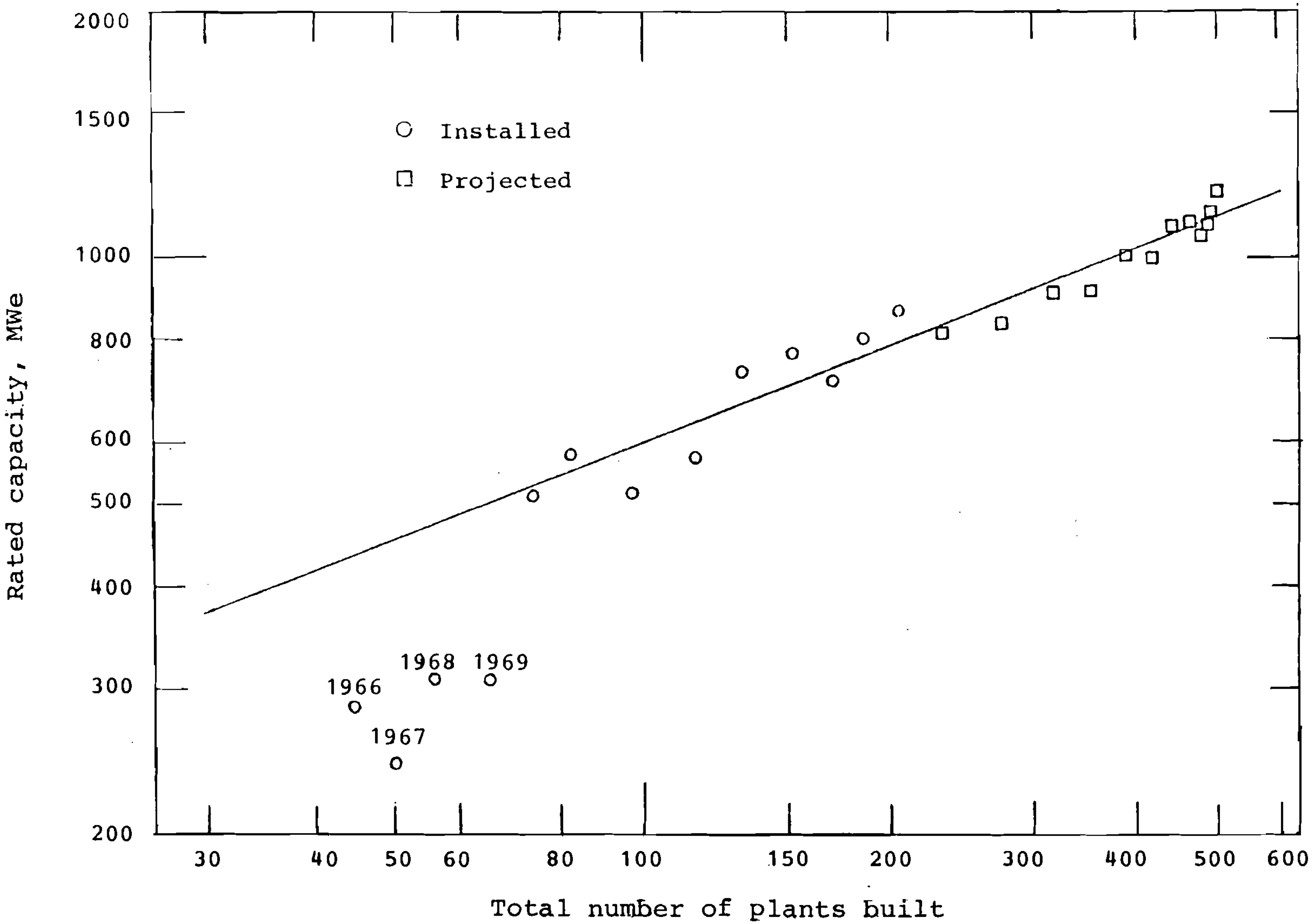
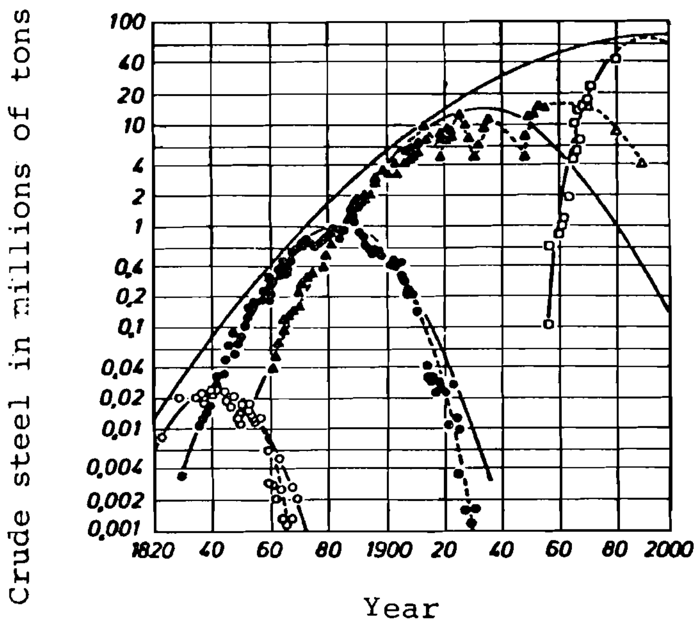


Figure 4. Relationship between Average Rated Capacity of Nuclear Reactors Completed in the Year and Cumulative Number of Reactors Built (Source: Spinrad 1980)



○ charcoal hearth refining

● puddling process

▲ liquid steel processes (open hearth and Bessemer)

□ LD process

Figure 5. The Life of Various Steel Production Processes in Rheinland-Westfalia (Source: Kootz et al. 1973) (Note: solid lines are log normal shaped curves fitted to the data)

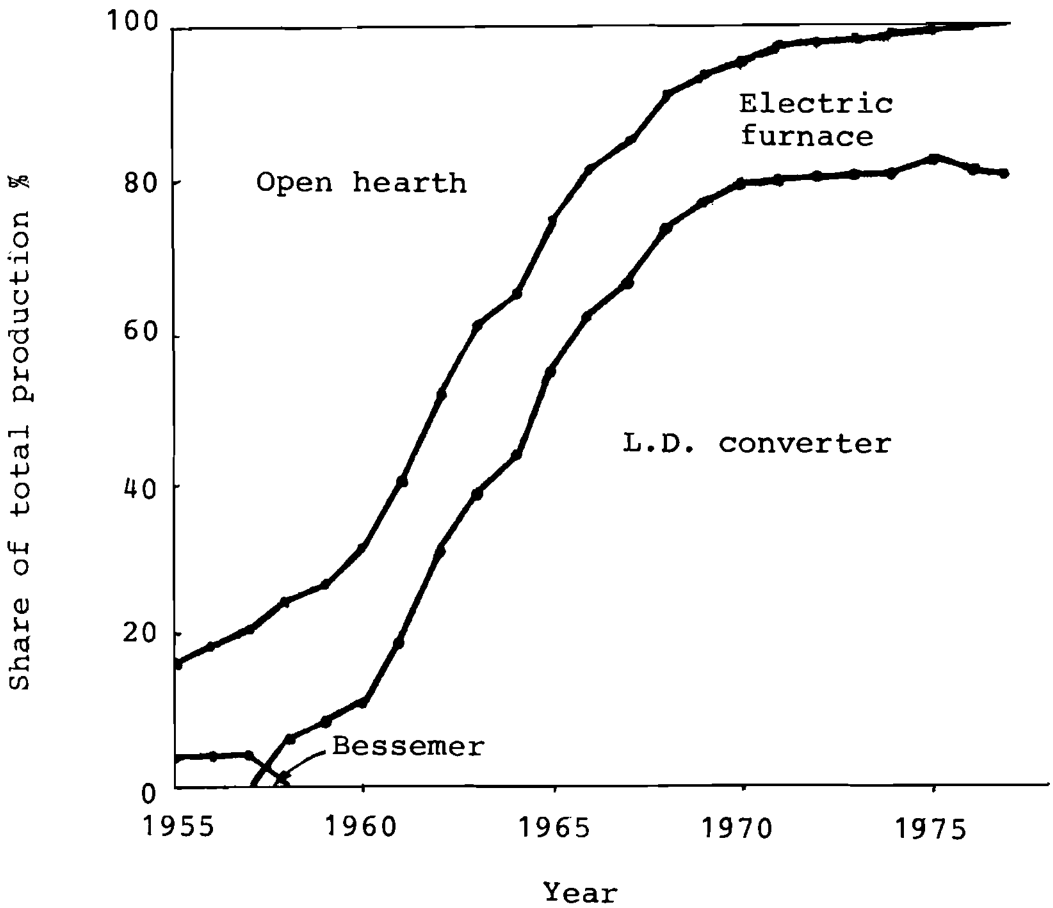


Figure 6. Development of the Proportion of Crude Steel Production Methods in Japan (Source: Resch 1973)

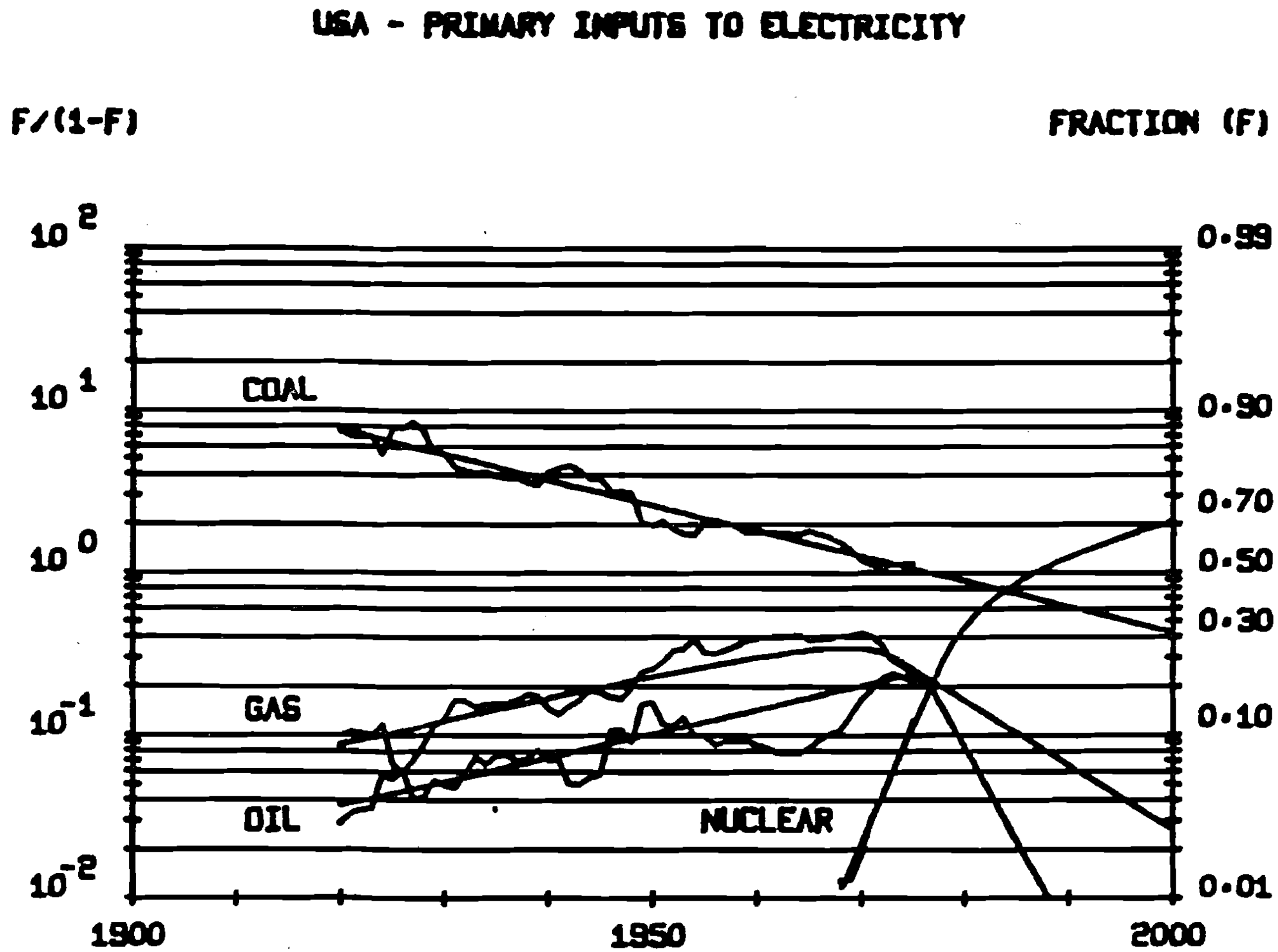


Figure 7. Share of Primary Inputs to Electricity Production in U.S.A. (Source: Marchetti and Nakicenovic 1978) (Note: The curves fitted to the data are based on the Peterka model.)

Since the observed behavior is the result of decisions by individual firms on size of plant and choice of process an understanding of the dynamics is of interest to those firms which design or manufacture process plant. To guide research and development it is desirable to determine the attributes of the "plant after next" or see what should be the characteristics of new processes in order to be accepted and adopted. Since such decisions will involve substantial commitment of resources the firm would like to have means by which they can evaluate alternatives and choose the most appropriate one.

The purpose of this paper is to review the available models for understanding the choice of process and size of plant. While there is literature on the qualitative aspects of technological progress the emphasis is on the extent to which formal models can be used to answer some of the questions posed above.

APPROACHES TO MODELLING

There are basically two general appraoches in developing models of the dynamics of choice of process and size of plant.

One is based on the aggregate description of past behavior. A particular mathematical relationship is suggested, its parameters estimated by standard statistical methods and, provided it is a good enough fit to the past it is assumed that it will continue to apply in the future. The mathematical relationship can either be a simple functional form (e.g., plant size increases exponentially with time) or it can be developed by considering the analogy between change in plant size or process mix with biological or psychological situations, in particular learning and growth. The advantage of this approach is that the mathematical relationships are usually quite simple and easy to comprehend, however, the disadvantage is that the future is assumed to be a projection of the past and thus it is not usually possible to answer questions concerning the influence of policy variables or the occurrence of unique events.

The other approach is based on the analysis of the sequence of decisions on plant size and mix. It is assumed that each decision is made by a decision maker who behaves rationally in the light of his perceptions about the future. Thus the emphasis in this approach is on developing models of rational behavior which enable the key parameters to be identified. Such models should allow for the existence of uncertainty about the future so it is necessary to consider the way in which the decision makers' perceptions about the future are related to his past experience. Generally this approach gives a relatively complex mathematical model but on the other hand it makes it possible to identify the effect of changes in the key parameters.

It is sometimes possible to extend these models of individual decisions to enable conclusions about the aggregate behavior to be drawn, that is, develop aggregate models which are derived from the actual decision making situation. In many ways these are the most useful models as, like the learning of grwoth models, they are sufficiently simple to have a comprehensible structure but, in contrast to other approaches, they are based on the actual decision making situation.

This paper consists of five sections. In the first the models of learning and growth are described. The next three consider rational decision maker oriented models of plant size, of technological substitution and of process mix. The last section considers the effect of uncertainty about the future.

MODELS OF LEARNING AND GROWTH

Learning

The increase in the size of plant over time is possible because of the ability of plant designers and opeartors to learn from their experience and incorporate experience gained from one plant into the next plant. If this is the dominant factor in determining the increase in plant size then it should be described by a learning model:

maximum plant size at t = f (cumulative experience at t)

where cumulative experience could be measured by

-- number of plants built prior to t
-- time since the first plant was constructed
-- total accumulated production up to time t.

In the psychological literature on learning a variety of functional forms have been proposed. The simplest is

$$y_t = kx_t^{\alpha}$$

where x_t is the measure of cumulative experience to t and y_t is the performance measure at t (e.g., maximum plant size).

This model has been applied by a number of authors to plant size data. Spinrad (1980) applied it to the growth in size of nuclear generating units, setting x_t equal to the number of units built up to time t. He found that the fit of the model was good.

Sahal (1979b) applied it to the growth in size of electrical generating units in Canada, setting x_t equal to the time since the first unit was built. He found a reasonably good fit to data series on both hydro electric and steam electric units.

The limitation in the above functional form of the learning curve is that it assumes no upper limit. A variety of models have been proposed, in particular the replacement model

$$y_t = k(1 - e^{-x_t/R})$$

the accumulation model

$$y_t = kx_t/(x_t + R)$$

where k is the upper limit on y_t and R determines the initial rate of increase of y_t.

Mazar and Hastie (1979) on the basis of an extensive review of the data on human performance on repetitive tasks, considered that the accumulation model fits the data better.

No attempt has been reported in the literature on fitting either of the above models to plant size data. However, it would appear that the accumulation model would be a good fit to Spinrad's data on light water reactors, with an asymptotic size of 1870 MW and R = 270.

Growth

The change in size of plant or the change in production using the different available processes is assumed to have the same characteristics as growth in biological systems.

That is, the basic mathematical relationship is (von Bertalanffy 1968)

$$\frac{dy_t}{dt} = f(y_t) \quad .$$

A variety of different models have been proposed for $f(y_t)$:

<u>exponential growth</u>: $f(y_t) = gy_t$

where g is constant whence

$$y_t = y_0 e^{gt} \quad .$$

<u>Gompertz growth</u>: $f(y_t) = g_t y_t$

where $g_t = be^{a-bt}$, i.e., $\frac{dg_t}{dt} = -b\, g_t$ whence

$$y_t = y_\infty \, e^{-\exp(a-bt)} \quad .$$

Note that $\lim_{t \to \infty} y_t = y_\infty$.

<u>logistic growth</u>: $f(y_t) = ay_t - by_t^{\,2}$

whence

$$y_t = \frac{ake^{at}}{1 + bke^{at}}$$

with $\lim_{t \to \infty} y_t = a/b$.

Next, consider a system consisting of n components, where the growth of the components is described by the equations

$$\frac{dy_{jt}}{dt} = g_{jt} \, y_{jt} \qquad (j = 1, 2, \ldots, n)$$

where y_{jt} is the size of component j at time t.

Suppose, however, that the g_{jt} are not known but

$$\frac{g_{jt}}{g_{nt}} = \beta_j$$

a constant for all t.

The solution to the growth equations is then characterized by

$$y_{jt} = k(y_{nt})^{\beta_j} \quad .$$

This is known as <u>allometric growth</u> (von Bertalanffy 1968: 64).

This model has been used by Sahal (1979b) to describe the relationship between the growth of maximum size of plant and the total size of the system. He found that it fitted the data in Canadian electric generating units quite well.

He also developed a more complex model of growth. He proposed that the maximum size of plant grew in accordance with the Gompertz model

$$y_t = y_\infty(t) \, e^{-\exp(a-bt)}$$

and also the value of the asymptotic plant size changed with time, either according to allometric growth

$$y_\infty(t) = k'Y_t^{\beta}$$

where Y_t was the total installed capacity, or the simple learning model

$$y_\infty(t) = k''x_t^{\alpha}$$

where x_t was taken as the time since the first unit was installed. He found that these models fitted the data very well although it must be noted that there are now three parameters instead of the two parameters for the allometric growth or learning models, i.e., $y_t = cY_t^{\beta}$ or $y_t = c'x_t^{\alpha}$.

Rather than the ratio g_{jt}/g_{nt} being constant an alternative hypothesis is that the difference is constant, i.e.,

$$g_{jt} - g_{nt} = c_{jn} \quad .$$

When n = 2 the equations then simplify to

$$\frac{d}{dt} \ln(f_t/(1 - f_t)) = c_{jn}$$

where $f_t = y_{1t}/(y_{1t} + y_{2t})$.

This is equivalent to the Fisher-Pry model of technological substitution of the old process or product 2 by the new process or product 1.

Sahal (1979a) compared this model with the allometric growth model for a variety of innovations. Both described the data quite well.

Peterka (1978) developed a solution for this model of constant difference in growth rate for n components. He fitted it to a variety of data series on the adoption of innovations or on the change in relative shares of different energy sources and also found that it fitted the data quite well.

Peterka also extended the above approaches for eliminating the unknown g_{jt} by assuming that

$$g_{jt} = (P_t - c_j)/\alpha_j$$

where P_t is unknown. Eliminating P_t gives the n - 1 equations

$$\alpha_j \frac{d \ln Y_{jt}}{dt} + c_j = \alpha_n \frac{d \ln Y_{nt}}{} + c_n$$

to which he added the equation

$$\frac{d}{dt} \ln Y_t = \rho \text{ a constant}$$

where $Y_t = \sum_{j=1}^{n} Y_{jt}$.

The solution to this set of equations also fitted the data very well but it must be noted that it has 2(n - 1) parameters as compared to the n - 1 parameters of the allometric growth or constant growth rate differential models.

Peterka justified the assumption on the form of g_{jt} by arguing that in the resulting growth equation

$$\alpha_j \frac{d Y_{jt}}{dt} = (P_t - c_j) Y_{jt}$$

the left hand side denotes the cost of increasing production in a period and the right hand side denotes the net revenue from sales in the period. That is, the equation describes the operation of a single product firm which invests a constant multiple of its net earnings.

The limitation of learning and growth models is that they appear to imply that the processes of increase in maximum size or technological substitution are totally determined and that there is no opprotunity for policy intervention to modify them. On the other hand it is remarkable how well they seem to fit the data.

DECISION MAKER ORIENTED MODELS OF GROWTH IN MAXIMUM SIZE

In deciding on the appropriate size of plant the decision maker balances the economies of scale in building larger plants with the penalties of having surplus capacity.

Srinivasan (1967) showed that, if demand has an exponential growth, characteristic of the optimum solution is that plants will be built a constant time interval T* apart. The size of plant built at time t will be

$$Y_t = d_t (e^{gT^*} - 1)$$

where d_t is the demand at time t and T* is that value of T minimizing

$$\frac{(e^{gT} - 1)^m}{1 - e^{-(r-mg)^T}}$$

where m is the economy of scale parameter and r the discount rate. The cost of building a plant of size y is ky^m.

Thus, given the parameters g, m and r the model specifies the size of plant which a rational decision maker would build.

In order to analyze the aggregate behavior of an industry it is necessary to consider how the parameters would be estimated. While r and m are likely to be reasonably constant, the decision maker would revise his estimate of g in accordance with experience.

One simple estimation method he might use is to estimate the growth rate at time t by

$$\hat{g}_t = \frac{\ln d_t - \ln d_0}{t}$$

Now, since a new plant would only be built if total capacity C_t is fully utilized one can set $d_t = C_t$ and set

$$\hat{g}_t \approx \ln C_t/t \quad .$$

It is possible to develop a variety of models which could then describe aggregate behavior.

While T* is dependent on g, for values of m and r which would be characteristic of electric generating units Peck (1974) showed that T* is quite insensitive to g.

Thus one can write

$$\ln Y_t = \ln d_t + \ln(e^{gT^*} - 1)$$

or setting $e^{gT^*} - 1 \approx gT^*$

$$\ln Y_t = \ln C_t + \ln \ln C_t - \ln t + \ln T^* \quad .$$

An equation of this form can also be obtained from the hypothesis of linear growth in demand. Manne's (1967) model shows that in that case

$$Y_t = gT^{**}$$

where T** is independent of g.

g can be estimated by

$$\hat{g}_t = C_t/t$$

so $\ln y_t = \ln (C_t/t) + \ln T^{**}$.

This model seems to fit Sahal's data on Canadian electrical generating units quite well.

Of course, alternative methods of estimating $\hat{g}_t$ can be assumed and the resulting aggregate behavioral model derived. Peck (1974) found that the Srinivasan model explained data on the size of electric generating units installed by a group of U.S. utilities.

The Distribution of Plant Size

Hjalmarsson (1974) considered the effect of firms following the Srinivasan model on the distribution of plant sizes.

He showed that if there are N plants the share of capacity due to the i largest plants is given by

$$F(i) = 1 - \frac{e^{(N-i)gT^*} - 1}{e^{NgT^*} - 1} \approx 1 - e^{-igT^*} \qquad 0 \leqslant i \leqslant N$$

While the proportion of plants of size x or smaller

$$G(x) = \frac{1}{N+1}\left(1 + \frac{1}{gT^*} \ln \left(\frac{x}{x_0'}\right)\right)$$

for $x_0' \leqslant x \leqslant x_0' e^{NgT^*}$ and x_0' is the size of the smallest plant (i.e., the first plant built).

If it is assumed that plants have a fixed life L it can be shown that the share of capacity due to the i largest plants is given by

$$F(i) = \frac{1 - e^{-igT^*}}{1 - e^{-gL}} \qquad i \leqslant L/T^*$$

Another distribution which can be derived is the share of capacity accounted for by plants larger than some value x.

Let

$$t = \frac{1}{g} \ln \left(\frac{x}{x_0'}\right)$$

Then the share of capacity at time t' accounted for by plants larger than x is given by

$$F(x,t') = 1 - e^{-g(t'-t)} \qquad t' > t$$

or

$$= 1 - \frac{x}{x_0'} e^{-gt'} \quad .$$

DECISION MAKER ORIENTED MODELS OF TECHNOLOGICAL SUBSTITUTION

If a new process (process 1) is developed then the rational decision maker will adopt the process in preference to the old process (process 2) if

$$PW_t(1) < PW_t(2)$$

where $PW_t(j)$ is the present worth of costs associated with process j at time t, the time when the new process becomes available.

Assuming exponential growth in demand and that an optimal policy of capacity expansion will be followed

$$PW_t(j) = \frac{k_j \, d_0^{m_j} e^{m_j gt} \left(e^{gT_j} - 1\right)^{m_j}}{1 - e^{-(r-m_j g)T_j}} + \frac{d_0 e^{gt} \gamma_j'}{r - g}$$

where d_0 is the demand at time 0, m_j is the economy of scale parameter, T_j is the optimum time between plant additions and γ_j' is the variable production cost per unit produced by process j.

The time dependency of $PW_t(j)$ can be emphasized by writing

$$PW_t(j) = h^{m_j t} A(m_j) + h^t \gamma_j$$

where $h = e^g > 1$.

It follows that the decision maker will choose the new process if

$$h^t\gamma_1 + h^{m_1 t} A(m_1) < h^t\gamma_2 + h^{m_2 t} A(m_2)$$

or

$$\gamma_2 - \gamma_1 > f(m_1,m_2,t)$$

where $f(m_1,m_2,t) = h^{-(1-m_1)t} A(m_1) - h^{-(1-m_2)t} A(m_2)$.

Figures 8A, 8B show the form of $f(m_1,m_2,t)$ depending on whether (A) $m_1 < m_2$ or (B) $m_1 > m_2$. From the figures it is possible to make some general conclusions about the pattern of adoption of the new process and how this is influenced by the relative values of γ_1, γ_2, m_1 and m_2.

(1) $\gamma_1 \leqslant \gamma_2$, $m_1 < m_2$. In this case the new process will be adopted once the demand reaches the level such that

$$\gamma_2 - \gamma_1 = h^{-(1-m_1)t} A(m_1) - h^{-(1-m_2)t} A(m_2) \quad .$$

From Figure 8A it can be seen that there will always be some value of t such that this equation is satisfied.

(2) $\gamma_1 > \gamma_2$, $m_1 < m_2$. In this case it can be seen from Figure 8A that there are several possibilities:

(i) the new process will never be adopted
(ii) the new process will only be adopted while demand is in a certain range. Once demand increases sufficiently the decision maker will revert to the old process. Also, at low demand levels the new process may not be appropriate.

(3) $\gamma_1 < \gamma_2$, $m_1 > m_2$. This is the opposite to case (2). That is, the alternatives are

(i) the process will be adopted irrespective of demand
(ii) the process will be used only at high levels of demand
(iii) the process will be used only at high levels of demand or at low levels of demand. There is an intermediate range of demand in which the old process is preferable.

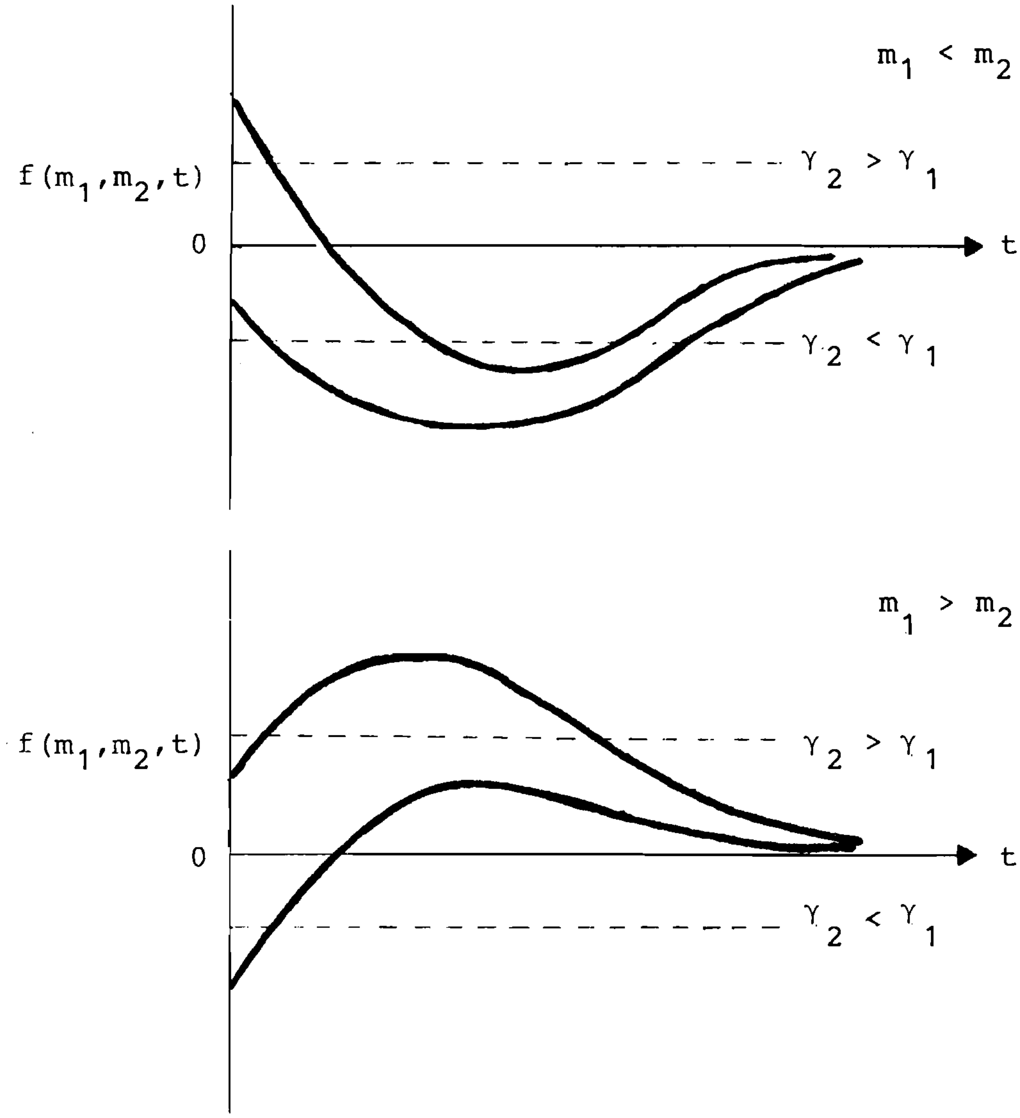


Figure 8. Form of $f(m_1, m_2 t)$

(4) $\gamma_1 \geqslant \gamma_2$, $m_1 > m_2$. In this case the new process will, if it is used at all, only be appropriate at low levels of demand.

Examples of each situation can be found. For example, case (1) is characteristic of the choice of the basic oxygen furnace in steel making and, conversely, case (4) is characteristic of the choice of the electric furnace for steel making. Case (3) would seem to characterize the use of solar collectors for energy conversion--present schemes seem to be either small scale or large scale. Case (2) seems to characterize the present role of oil fired electric generating units.

The implications for the aggregate behavior of adoption of the innovation are

Case (1). The innovation will be adopted once demand reaches a critical level. The share of production due to the new process will increase asymptotically to 100%.

Cases (2) and (4). The share of production due to the new process will initially increase to some maximum value but then decline.

Case (3). The behavior is more complex. In alternative (iii) it will increase, then decrease, but eventually increase again.

In case (1) it is possible to develop more specific models describing the extent of adoption of the innovation. Case (1) is characterized by the existence of a critical demand level above which the innovation should be used, or alternatively, there is a minimum feasible plant size x for the new process.

Hence, if there is a single firm the share of production at time t' accounted for by the new process will be given by

$$F_a(x,t') = 1 - e^{-g(t'-t)} \qquad t' > t$$

where $gt = \ln (x/x_0')$.

Figure 9 shows $\ln \{F_a/(1-F)\}$ as a function of $t' - t$.

With a finite plant life the form of F will be

$$F_b = \frac{1 - e^{-g(t'-t)}}{1 - e^{-gL}} \qquad t' > t$$

$\ln F_b/(1 - F_b)$ is shown for this case on Figure 9b with gL set at 1.5.

Next, suppose there are a number of firms in the industry and it is desired to model the overall behavior. Then it is necessary to assume some distribution of firm size. Suppose the firm sizes have a Pareto distribution. That is, the total production of the i-th largest firm will be such that (Ijiri and Simon 1977:196)

$$d_{it} = d_{1t} i^{-(1+\rho)}.$$

Now, if the new process can only be used at a minimum plant size of x it follows that the time at which the i-th largest firm will adopt the process is given by

$$t_i = t_1 + \frac{(1 + \rho) \ln i}{g} \quad .$$

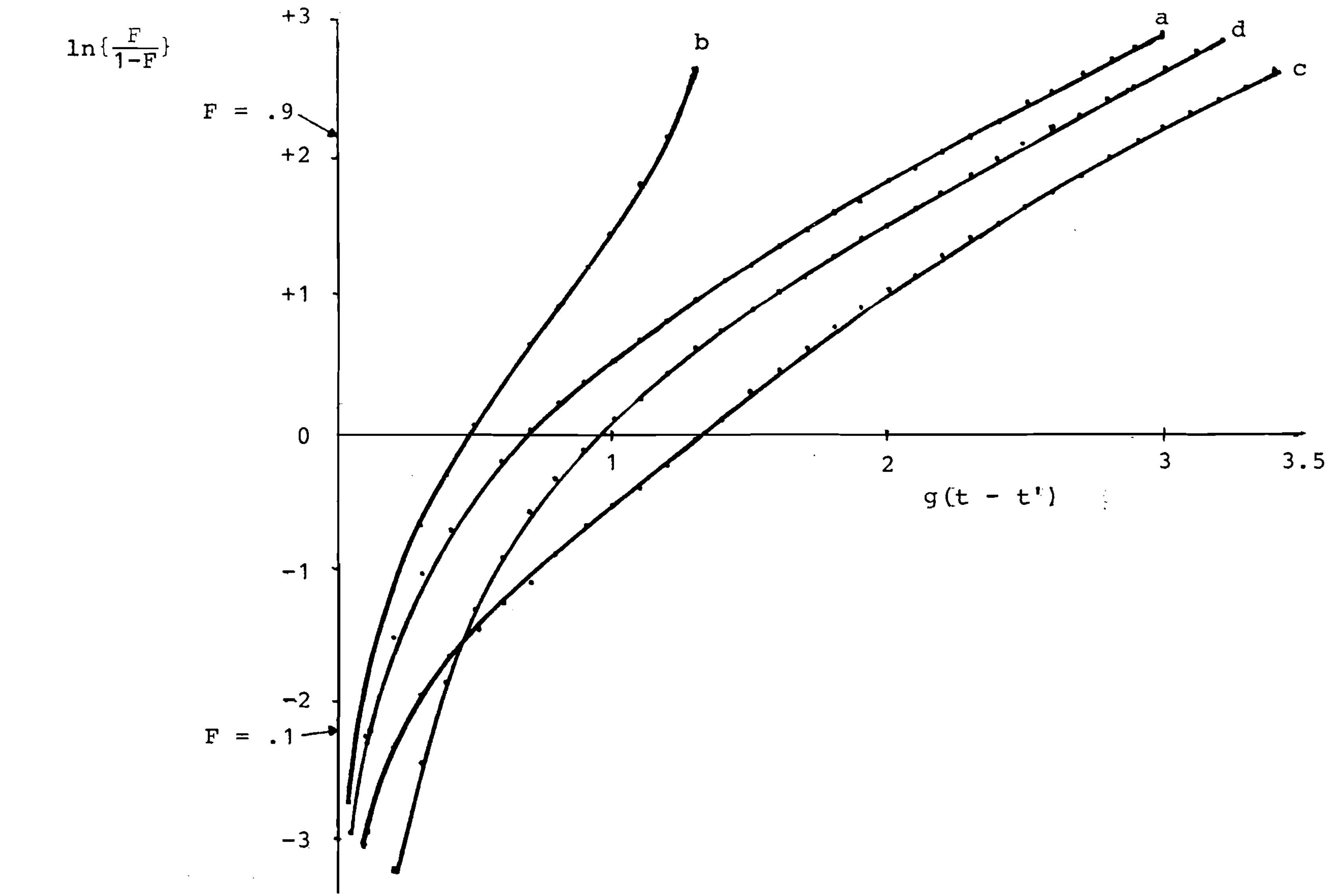


Figure 9. ln $F/(1 - F)$ as a Function of $g(t - t')$ for Various Substitution Models

Hence the fraction of production at time t accounted for by the new process is given by

$$F_C(t) = \frac{\sum_{i=1}^{N} f_i (1 - e^{-g(t-t_i)})\,\delta(t,t_i)}{\sum_{i=1}^{N} f_i}$$

where $f_i = i^{-(1+\rho)}$ and $\delta(t,t_i) = 1$ if $t > t_i$

$= 0$ otherwise

whence

$$F_C = \frac{\sum_{i=1}^{N} f_i\,\delta(t,t_i) - e^{-g(t-t_1)} \sum_{i=1}^{N} \delta(t,t_j)}{\sum_{i=1}^{N} f_i} .$$

Figure 9C shows a plot of $\ln \{F_C/(1 - F_C)\}$ against $t - t_1$ for $\rho = 0$ and $N = 4$.

Effect of Initial Size Limit on New Process

Even though the new process may be such that all firms would adopt it, there could initially be technical constraints on the maximum size of plant for the new process. As a result it may only be appropriate to firms in a particular size range. However, as experience in the use of the process is obtained the maximum technically feasible plant size will increase and it will be appropriate for an increasing proportion of firms. This appears to have been the situation when the basic oxygen process for steelmaking was introduced (Buzacott 1980).

The way in which the applicability of the new process changes with time will depend on two factors, (i) the way in which the technical limit on plant size increases with time, (ii) the distribution of size of firm and hence size of plant appropriate to their requirements.

As a tentative model of the combined effect of the two factors, let f_t be the proportion of the total demand increment which can be met with plants of the new process at time t.

One possible form of f_t is that

$$f_t = f_0 + bt \qquad 0 < t < T = (1 - f_0)/b$$

$$= 1 \qquad t > T$$

Then the proportion of total capacity which will consist of plants of the new capacity will be

$$F_d(t) = f_0(1 - e^{-g(t-t')} + \int_0^T (1 - e^{-g(t-t'-u)})df_u$$

$$t > T + t'$$

$$= f_0(1 - e^{-g(t-t')}) + \int_0^{t-t'} (1 - e^{-g(t-t'-u)})df_u$$

$$t' < t < t' + T$$

If it is assumed that $f_0 = 0$ it follows that

$$F_d(t) = 1 - be^{-g(t-t')}(e^{gT} - 1)/g \qquad t > T + t'$$

$$= b(t-t') - be^{-g(t-t')}(e^{g(t-t')} - 1)/g$$

$$t' < t < t' + T$$

Figure 9d shows $\ln \{F_d/(1 - F_d)\}$ for $b/g = 2$ and $gT = .5$.

It can be seen from Figure 9 that the models which give $\ln \{F/(1 - F)\}$ closest to a straight line over the range $F = .1$ to $F = .9$ are models b and c. A combination of b and c, i.e., a model which allows for both a finite life of plant and a distribution of firm size, would give a curve which is even closer to a straight line and thus be consistent with the Fisher-Pry model.

DECISION MAKER ORIENTED MODELS OF PROCESS MIX

The purpose of this section is to review models which explain why a firm will consider using a mixture of different processes in order to meet the total demand.

One reason is that the firm supplies geographically distinct markets and the nature of transport costs is such that the demand in each market can best be met from a local plant. The differences in size of the geographically distinct markets may mean that different processes are appropriate to different locations. There is a considerable literature on the question of the optimal size and location of plants so this aspect of process mix will not be considered. Erlenkotter (1967) has considered the dynamics of the interaction between market growth, economies of scale in plant construction costs and the transportation costs. He has shown that a constant cycle time between capacity expansions is not optimal.

The explanation for the existence of a mix of processes based on the spatial distribution of markets and raw materials combined with transportation costs is well known and, at least qualitatively, its implications are understood. So in this section we will focus on why a firm would consider using a mixture of processes in situations where transport costs are not significant.

Although apparently different, a situation which has been shown to be formally equivalent to geographically distinct markets is that in which the total market can be segmented into different market sectors. For example, one such segmentation might be based on quality requirements or it could be based on the physical dimensions of the product. Suppose also that some of the available processes can only supply some of the sectors, that is there could be general purpose processes and specialized processes.

Consider the specific case of a general purpose process, process 1, and a specialized process, process 2. Then divide the total market into that portion, market 2, which can be met using process 2 and that portion, market 1, which can only be met using process 1. Then Erlenkotter (1974) showed that this is equivalent to the case of geographically distinct markets 1 and 2 in which the cost of shipment from a plant in market 2 to market 1 is zero but no shipment is possible from a plant in market 1 to market 2. Unfortunately the fact that optimum cycle time is not optimal means that it is difficult to derive aggregate models of behavior. Kalotay (1973) derived some results concerning whether specialized plant should be used in the case where both specialized and general purpose plants had the same scale characteristics but the practically more interesting case is that where the specialized plant has an economy of scale parameter m_2 which is greater than m_1.

Yet, even when any of the available processes can meet all the requirements of the market and transport costs are not significant, it is still possible that a mixture of processes will be appropriate.

Graphical Determination of Generation Mix

We consider here a well-known graphical procedure for determining optimal mix of electricity generation types.

In electricity generation the fact that there are no effective storage device implies that demand must be met instantaneously. The time varying characteristic of demand is often described by the load duration curve in which the amount of time over which a certain level of demand has occurred is plotted against the demand level.

Demand variation over time is not necessarily particular to electricity generation. In other industries demand fluctuates but normally it is absorbed either by proper inventory scheme or delay in delivery of products. Thus for those industries where inventories are not appropriate or where delivery delay implies reduced service the situation is similar to electricity industry. Also there are industries such as steel making where sporadic demand occur over the normal fluctuation of demand. Thus the concept of load duration curve may very well be relevant to some industries other than electricity generation and the following discussion can be applied.

Now let C_i (\$/MW) be the annualized plant construction cost (fixed cost) and γ_i (\$/MWH) be the plant operating cost (variable cost) where i indicates the plant type. For the time being we assume that there are no economies of scale in the above costs. Also we assume that the maximum demand d and the load duration curve $\phi(x)$ are given (see Figure 10) and that there are no initial set of plants.

Let x_i be the capacity of plant type i. For the given set of C_i, γ_i $\phi(x)$ and d the graphical determination of optimal mix* is demonstrated in Figure 10 for i = 1, 2, 3. The thick solid line in the upper graph of Figure 10 represents a minimum cost polygon. The dotted lines drawn from the two intersections of the cost polygon and reflected on the demand axis by the load duration curve determine the optimal capacity for each plant.

This graphical procedure is verified for the illustrated case as follows:

Objective function to be minimized is the total costs (Capital and Operating costs), which is expressed by

$$C_1x_1 + C_2x_2 + C_3x_3 + \gamma_1\int_0^{x_1}\phi(x)dx + \gamma_2\int_{x_1}^{x_1+x_2}\phi(x)dx + \gamma_3\int_{x_1+x_2}^{x_1+x_2+x_3}\phi(x)dx \qquad (1)$$

subject to

$$x_1 \geqslant 0 \quad , x_2 \geqslant 0 \quad , x_3 \geqslant 0 \qquad (2)$$

$$x_1 + x_2 + x_3 = d \qquad (3)$$

Note that $\phi(x)$ is nonlinear and hence the model is nonlinear.

---

*For a complete treatment see Philips et al. (1969).

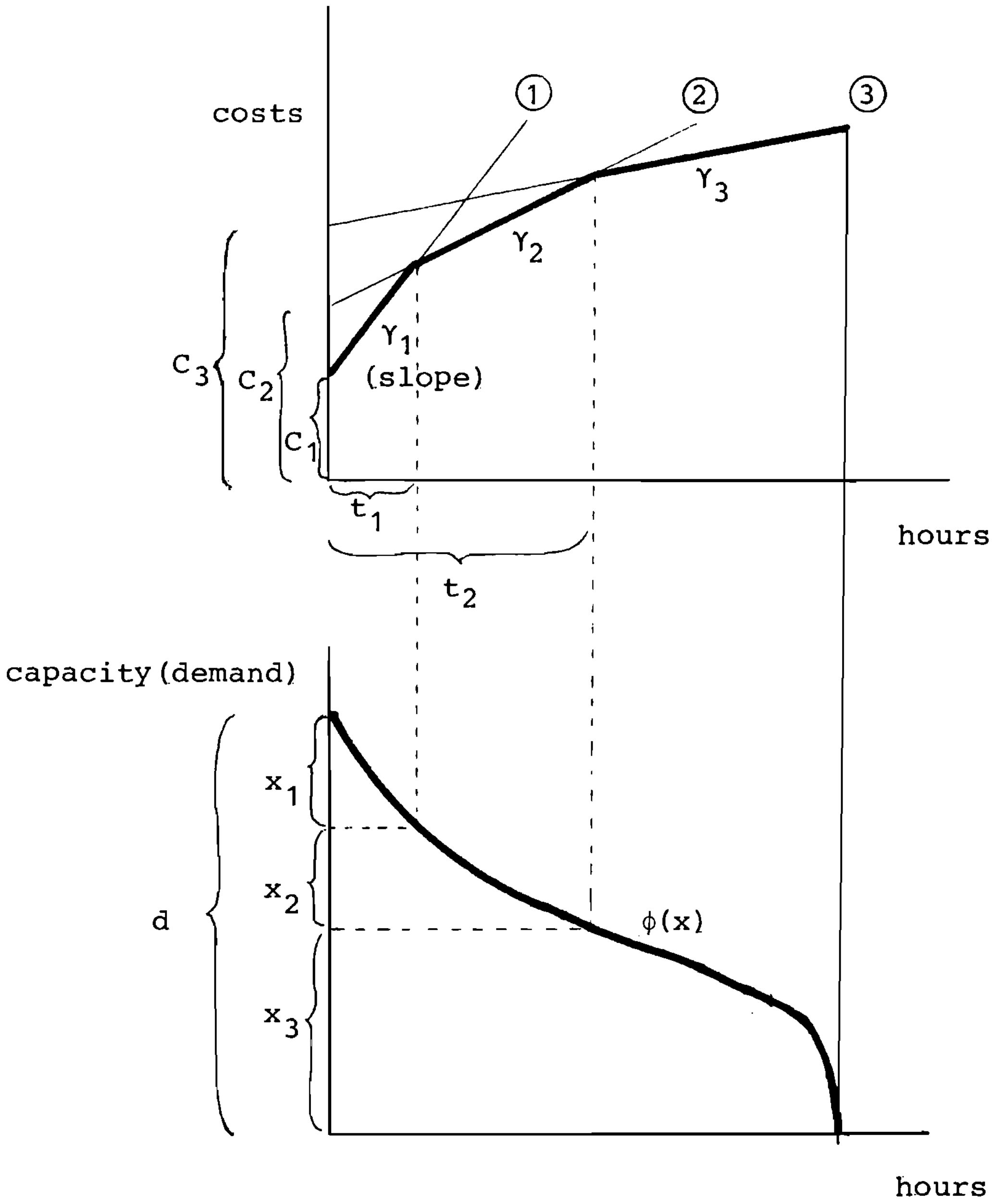


Figure 10. Graphical Determination of Optimal Mix

Now define

$$\Phi(x) = \int_0^x \phi(x')dx'.$$

$\Phi(x_1 + x_2 + x_3)$ is constant over all x_1, x_2, and x_3 and by eliminating the variable x_3 using (3), (1) becomes

$$\Gamma(x_1,x_2) = (C_1 - C_3)x_1 + (C_2 - C_3)x_2 + (\gamma_1 - \gamma_2)\Phi(x_1)$$

$$+ (\gamma_2 + \gamma_3)\Phi(x_1 + x_2) \tag{4}$$

$\frac{\delta\Gamma}{\delta x_1} = 0$, $\frac{\delta\Gamma}{\delta x_2} = 0$ yield the necessary condition for optimality, which can be written as

$$\phi(x_1) = \frac{C_2 - C_1}{\gamma_1 - \gamma_2} \tag{5}$$

$$\phi(x_1 + x_2) = \frac{C_3 - C_2}{\gamma_2 - \gamma_3} \tag{6}$$

On the other hand, the cost curve can be represented by

$$C_i + \gamma_i t$$

where t is the duration over which plant i is operated.

Thus the intersection of the curves 1 and 2 is obtained by setting

$$C_1 + \gamma_1 t_1 = C_2 + \gamma_2 t_1 \quad .$$

Hence

$$t_1 = \frac{C_2 - C_1}{\gamma_1 - \gamma_2} \tag{7}$$

similarly

$$t_2 = \frac{C_3 - C_2}{\gamma_2 - \gamma_3} \qquad (8)$$

(5), (6) and (7), (8) are the basis of the graphical procedure.

The graphical determination of an optimal mix is for only one term (often called one-year problem). But it can be regarded as an expression of some steady state in the dynamic context.

Assume that

-- maximum demand increases term by term.
-- the shape of the load duration curve is invariant.
-- incremental addition of each plant type is allowed.
-- no retirement of plants occur.
-- no changes in C_i's and γ_i's over time.

Then there will be no change in the proportion of the mix over time. This is because under the above assumptions the difference between the two load duration curves has the same shape as these two load duration curves and hence the same proportion for each plant type will result again.

Thus the optimal mix obtained from the graphical procedure can be interpreted as the steady state development pattern of an optimal expansion plan. Using this interpretation we consider some questions presented in the following.

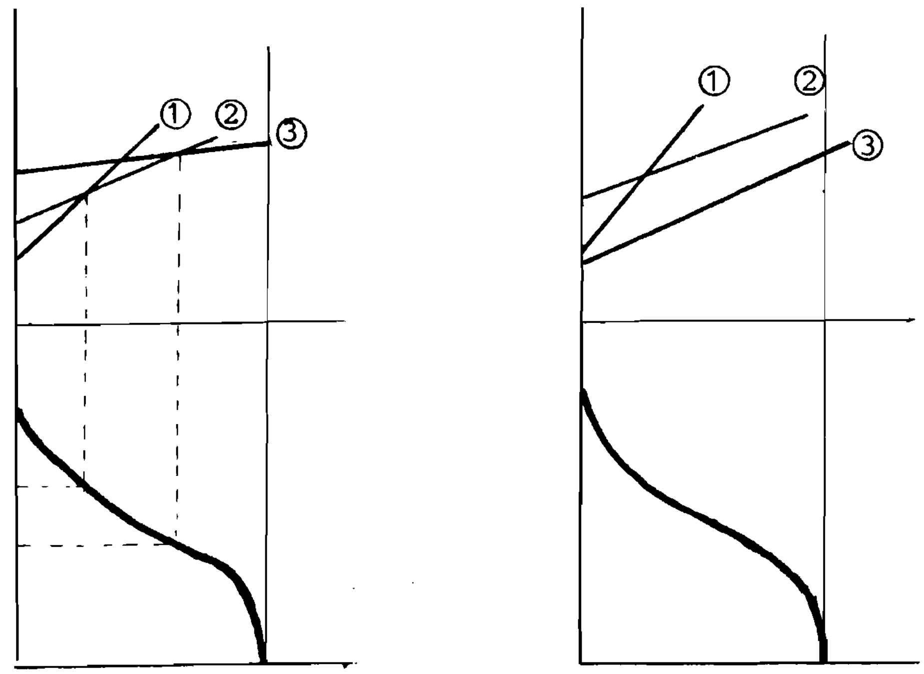
Why Does Mix of Different Generation Types Exist?

It is clear from the graphical procedure why a mix of generation types results in the cheaper total costs. That is, the monotone decreasing characteristics of the load duration curve $\phi(x)$ together with the cost curve characteristics (C_i, γ_i) determine the necessity of generation mix.

Intuitively it is clear that the condition for a mix to exist is that the cost curves of two different generation types intesect. Depending on how these cost curves intersect, all or a part of generation types will participate in the generation mix (of course in the sense of minimizing total present worth costs also see Figure 11).

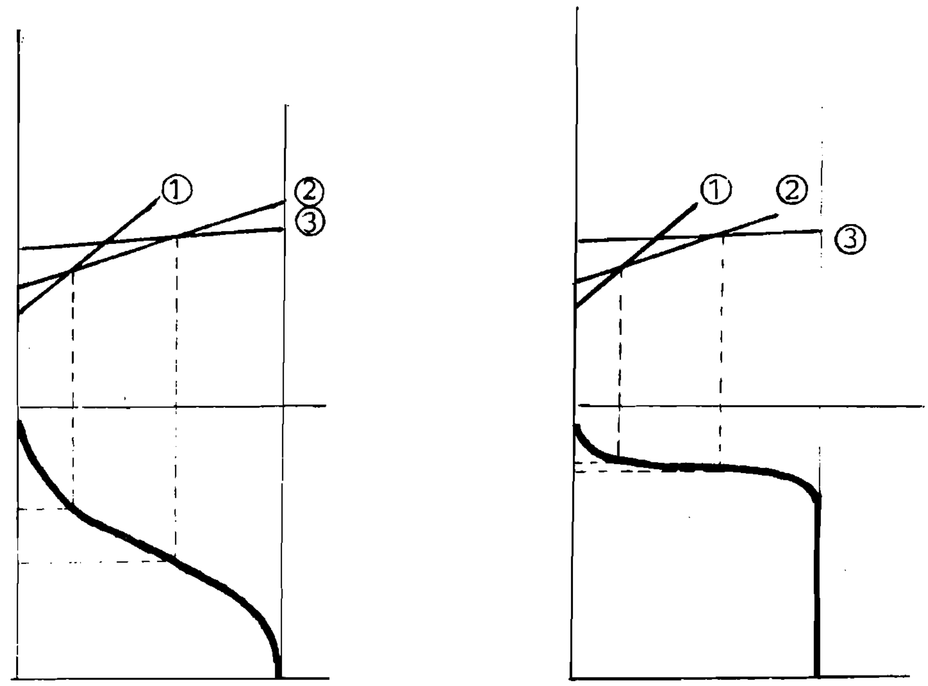
On the other hand if the shape of a load duration curve become more like a square (this corresponds to the case where one has perfect storage which absorbs any fluctuation of demand over time), then advantages of having a mix of generation types diminish (see Figure 12).

Obviously there are some other reasons why a mix exists. In electricity generation the demand could not be covered by a single method (e.g., hydro) and therefore some other means for electricity generation had to be introduced. The operating characteristics of nuclear power plants are such that load



a) all types participate b) a single type 3 dominates

Figure 11. Effects of Cost Characteristics on the Optimal Mix



a) all types equally participate b) optimum mix dominated by type 3

Figure 12. Effects of the Shape of Load Duration Curve

following ability has not been established and therefore storage hydro or gas turbine plants are necessary. Other reason might be that if we depend on a single technology then it has less flexibility. In this context, it is desirable to have mixture of different plant types. However, it should be noted that having a mix of different generation types may result in a lower total cost as was indicated by the graphical procedure.

How Generation Mix Change Over Time?

Since an optimal generation mix is dependent on the shape of load duration curve and the costs characteristics (C_i's and γ_i's), it follows that any significant change with respect to these factors would change the optimal generation mix.

Now assume that the demand is increasing without any change in $\phi(x)$, C_i's, and γ_i's. Suppose that the present mix is different from the optimal mix for the given $\phi(x)$, C_i, and γ_i. Then an optimal expansion plan will bring the mix eventually to this optimal mix. This kind of arguments can be put forward more clearly if we formulate the expansion planning problem as an optimal control problem (Schlaepfer 1978, see Figure 13). In Figure 13, optimal trajectories are drawn for different initial conditions. The relatively slow convergence to the steady state expansion plan is due to the restriction on the capacity which can be added at any instant of time.

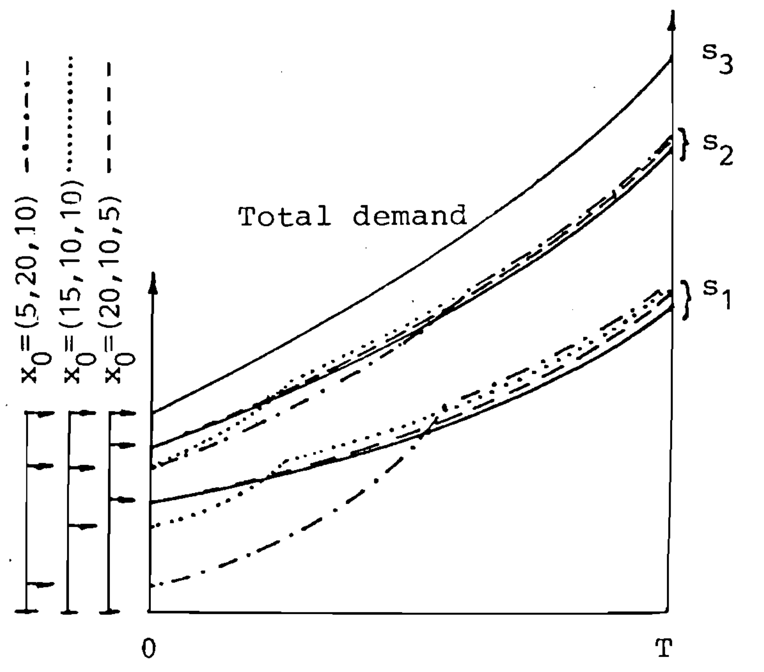


Figure 13. Trajectories of Optimal Expansion Plan for Different Initial Generation Mix (Source: Schlaepfer 1978)

How a New Technology Comes In?

In principle, a new technology will participate in an optimal generation mix if its cost curve appears in the minimum cost polygon. Whether the new technology will occupy the base load range, the middle load range or the peak load range depends on its cost characteristics.

For example the use of solar energy is characterized by a relatively high cost for construction ($/MW) and a very low operating cost ($/MWH). Thus this technology would participate in the base load range. While combined heat and power station will come in either the middle or the base load range by the similar argument. It should be mentioned, however, that the above arguments disregards the fact that the generated power itself can vary heavily over time. (An effective storage device might help solar energy to come in.)

Effect of Economies of Scale

Economy of scale suggests the cost C is not constant for each MW capacity to be installed, but rather C is a function of size, i.e., C(x). Typical relationship between x and C is shown in Figure 14.

Similar arguments can be made for γ also.

In this case the graphical procedure shown in Figure 10 is no longer applicable. An expression for the objective function when γ_i's the assumed to be constant can be expressed as

$$\hat{\Gamma}(x_1,x_2) = C_1(x_1)x_1 + C_2(x_2)x_2 + C_3(x_3)x_3 + (\gamma_1 - \gamma_2)\Phi(x_1) + (\gamma_2 - \gamma_3)\Phi(x_1 + x_2) \quad (9)$$

Thus (9) is to be minimized with respect to x_1, x_2, and x_3, where $x_1 + x_2 + x_3 = d$ is the constraint.

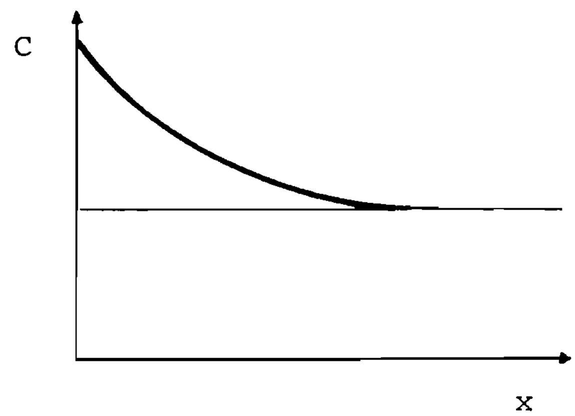


Figure 14. Economy of Scale

Again by setting $\frac{\delta\hat{\Gamma}}{\delta x_1} = 0$ and $\frac{\delta\hat{\Gamma}}{\delta x_2} = 0$ and rearranging we get the following necessary conditions for optimality.

$$\frac{\delta C_1}{\delta x_1}x_1 + C_1(x_1) - \frac{\delta C_2}{\delta x_2}x_2 - C_1(x_2) = -(\gamma_1 - \gamma_2)\phi(x_1) \qquad (10)$$

$$\frac{\delta C_2}{\delta x_2}x_2 + C_2(x_2) - \frac{\delta C_3}{\delta x_3}x_3 - C_3(x_3) = -(\gamma_2 - \gamma_3)\phi(x_1 + x_2) \qquad (11)$$

The simplest expression for the economy of scale would be of the following form:

$$C(x) = a + \frac{b}{x} \qquad \text{or} \qquad C(x)\cdot x = ax + b \qquad (12)$$

so the necessary conditions become*

$$a_1 - a_2 = -(\gamma_1 - \gamma_2)\phi(x_1) \qquad (13)$$

$$a_2 - a_3 = -(\gamma_2 - \gamma_3)\phi(x_1 + x_2) \qquad (14)$$

(13) and (14) are exactly the same forms as (5) and (6), respectively, in which C_i's are replaced by a_i's. Thus if $a_i \neq C_i$ then the resulting optimum mix will be, in general, different from when there are no economies of scale.

The effect of economy of scale in the form of (12) can be illustrated as in Figure 15. Let us assume that the plant type 3 has significant scale economy compared with the other types of plants. Then it is reasonable to assume that $a_3 < C_3$. This brings the cost curve 3 in Figure 15 down to 3'. The new intercepts with the curve 2 will give the new set of optimal mix in which the optimum size for plant type 3 becomes larger than no economy of scale is assumed.

It should be noted that the significance of the effect of economy of scale depends very much on such parameters as γ_i, a_i and b_i. For example, consider the case where there are two types of plants; one is coal and the other is nuclear. Usually a nuclear plant is characterized by higher capital costs per MW and lower operating costs per MWH than the corresponding costs for coal, i.e., $C_n > C_c$ and $\gamma_n < \gamma_c$. Thus the resulting minimum cost ploygon becomes as shown in Figure 16a. Now when the economy of scale is expressed by the straight lines (as in (12))

*This condition applies only for $x_i \varepsilon (0,d)$, $i = 1, 2$, and moreover a solution to (13) or (14) may or may not give a local minimum.

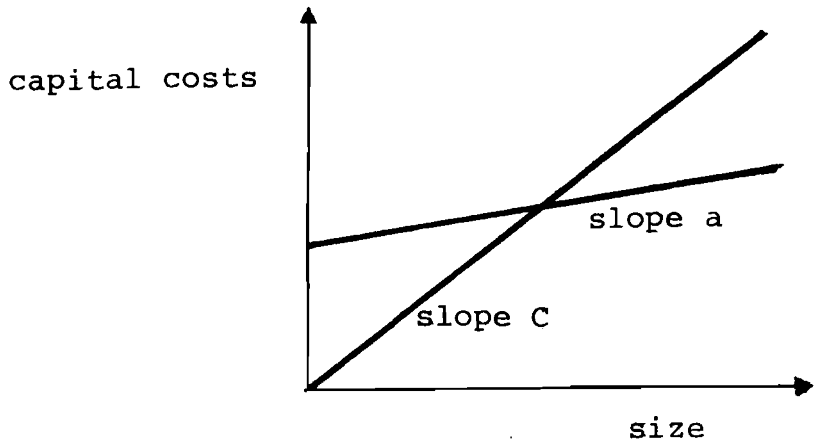


Figure 15a. A Simple Expression of Economy of Scale

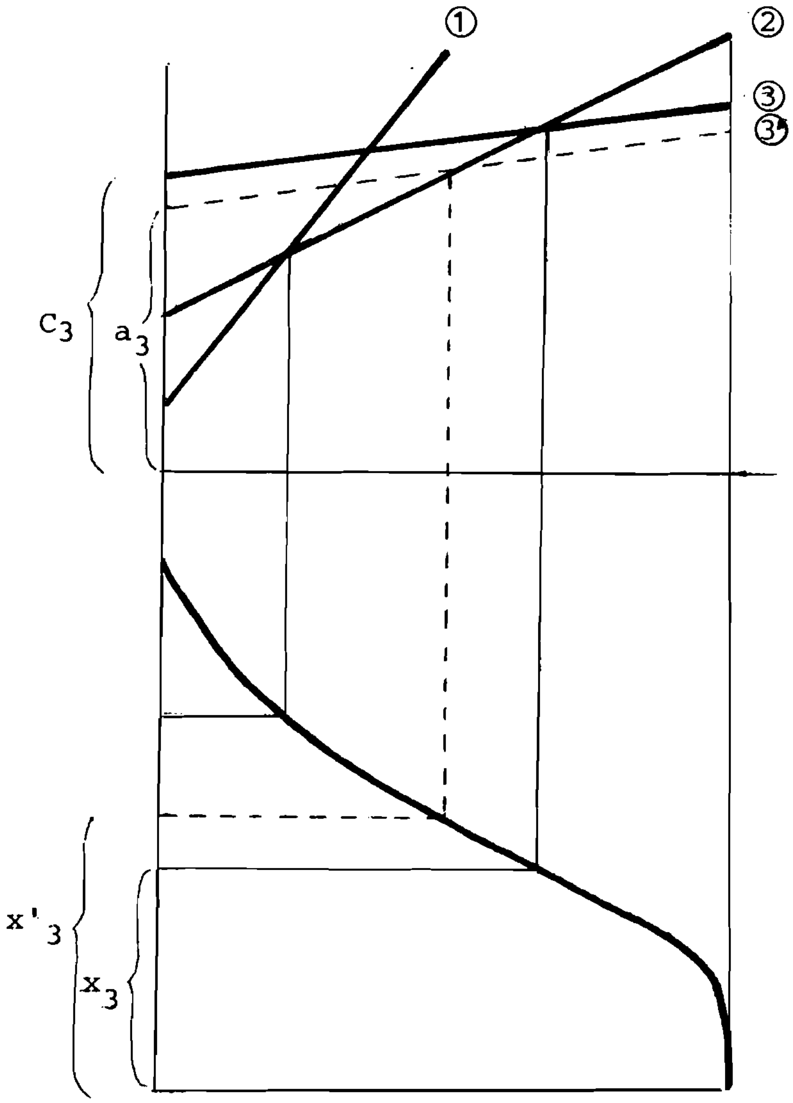


Figure 15b. Effect of Economy of Scale

as shown in Figure 16b such that $a_n < a_c$, then by replacing C_n and C_c by a_n and a_c we obtain the new minimum cost polygon shown in Figure 16c. So in this case the effect of economies of scale is very significant that the necessity of having generation mix diminishes. Whether nuclear or coal plants dominate will depend on the values of b_n and b_c. If b_n is so high that the economic advantage indicated by Figure 16c is offset, then coal plants will dominate. On the other hand if the characteristics of the two plant types were such that $a_n > a_c$ a new mixture of these two plants may exist.

Now consider the case where economy of scale is represented by "power law." We express

$$C(x) = Cx^{m-1} \qquad 0 < m < 1$$

or

$$C(x) \cdot x = Cx^m \tag{14}$$

If m = 1, then C(x) = C implying the case where there is no economy of scale (see Figure 17). The function to be minimized is readily written by using (9), i.e.,

$$\tilde{\Gamma}(x_1,x_2) = C_1 x_1^{m_1} + C_2 x_2^{m_2} + C_3 x_3^{m_3} + (\gamma_1 - \gamma_2)\Phi(x_1) + (\gamma_2 - \gamma_3)\Phi(x_1 + x_2) \tag{15}$$

The necessary condition for optimality is

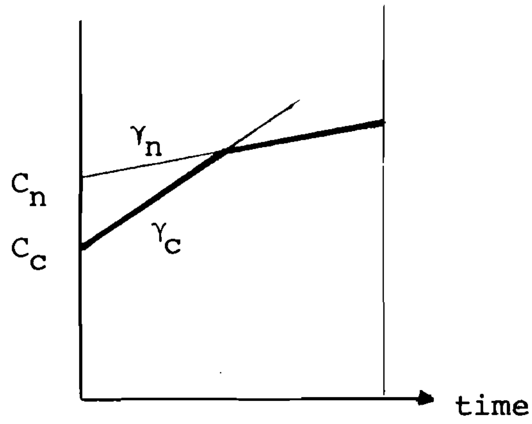
$$C_1 m_1 x_1^{m_1-1} - C_2 m_2 x_2^{m_2-1} = -(\gamma_1 - \gamma_2)\,\phi(x_1) \tag{16}$$

$$C_2 m_2 x_2^{m_2-1} - C_3 m_3 x_3^{m_3-1} = -(\gamma_2 - \gamma_3)\,\phi(x_1 + x_2) \tag{17}$$

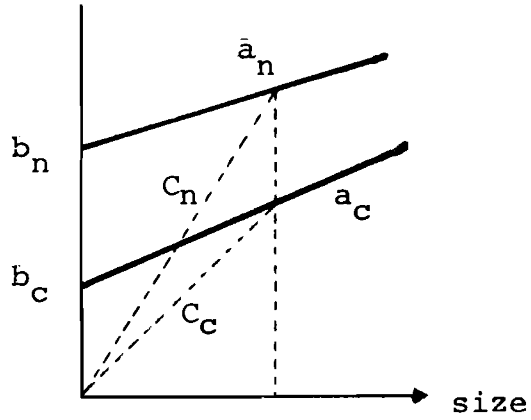
Note again that if $m_i = 1$ then (16) and (17) are exactly the same as (5) and (6), respectively.

When $0 < m_i < 1$, it is not as straightforward as the previous case where economy of scale was represented by a fixed part plus a variable part (e.g., (12)).

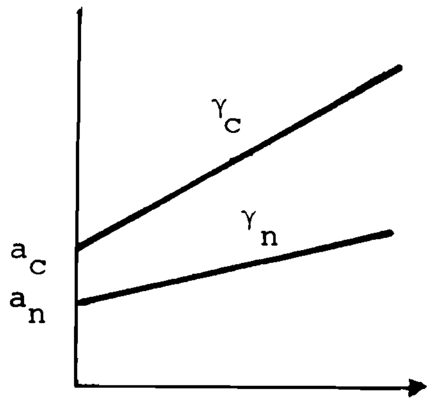
To see the economy of scale in more concerte way, let us consider again the case where we have only two generation types;



a) Minimum Cost Polygon (no economies of scale)



b) An Expression for Economies of Scale in Relation to C_n and C_c



c) Minimum Cost Polygon with Economies of Scale

Figure 16.

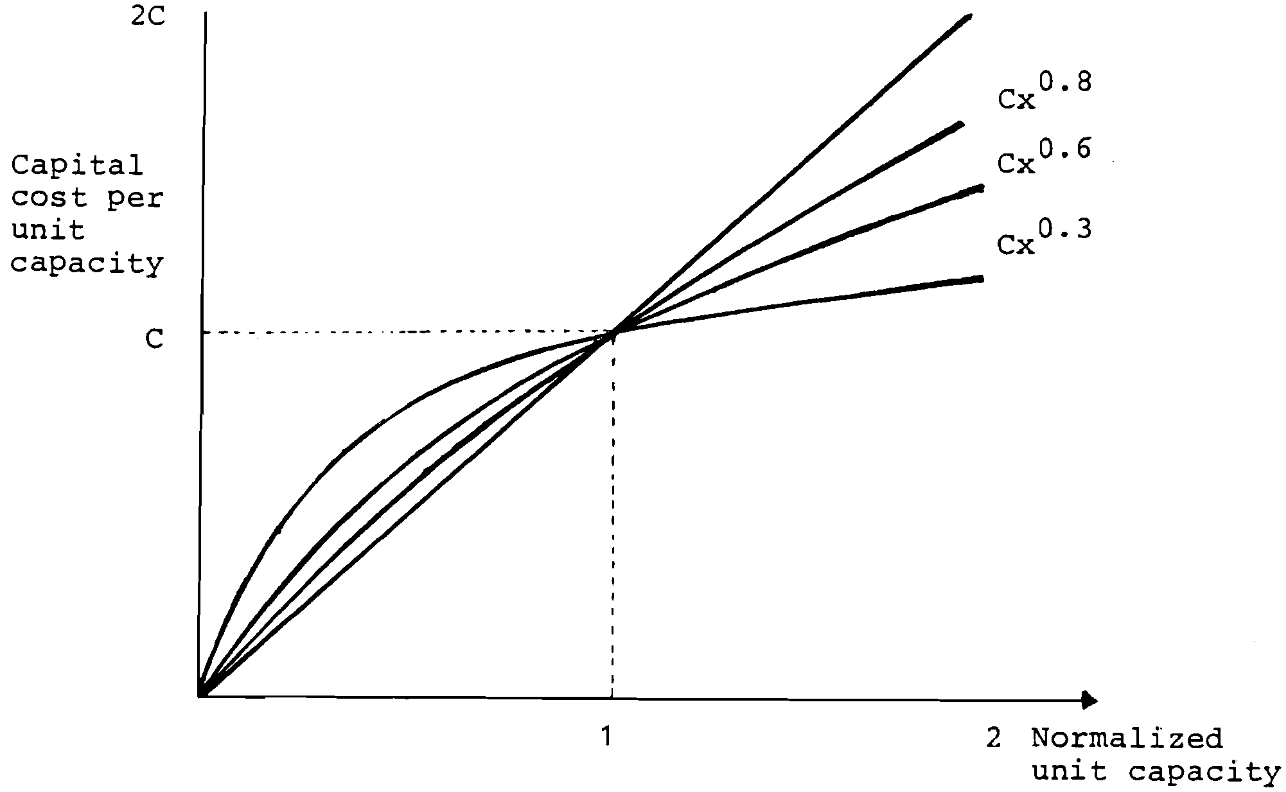


Figure 17. Representation of Economy of Scale

one is coal and the other is nuclear. So let us suppose that $x_c + x_n = d$, where c and n corresponds to coal and nuclear, respectively.

The function to be minimized in this case is, from (15),

$$\Gamma(x_c) = C_c x_c^{m_c} + C_n (d - x_c)^{m_n} + (\gamma_c - \gamma_n)\Phi(x_c) \tag{18}$$

Now

$$\frac{d\Gamma}{dx_c} = C_c m_c x_c^{m_c - 1} - C_n m_n (d - x_c)^{m_n - 1} + (\gamma_c - \gamma_n)\Phi(x_c) \tag{19}$$

and

$$\frac{d^2\Gamma}{dx_c^2} = C_c m_c (m_c - 1) x_c^{m_c - 2} + C_n m_n (m_n - 1)(d - x_c)^{m_n - 2} + (\gamma_c - \gamma_n)\left.\frac{d\phi}{dx}\right|_{x_c} \tag{20}$$

A numerical example is shown in Figure 18 and Figure 19. Figure 18 illustrates the case where no economies of scale is taken into account. The numbers for C_c, γ_c, C_n, and γ_n are arbitrarily chosen by referring to, for example, Huettner (1975). For the values of m_c and m_n, various authors suggest different values. Table 1 shows these numbers from which $m_c = 0.8$ and $m_n = 0.5$ are arbitrarily chosen for illustrating purpose.

In Figure 19, the cost curve (representing equation (18)) is shown together with the cost curve for the case where no economies of scale are considered. The effect of taking scale economy into account is significant; in this case the effect is to change the convexity of the cost curve (with no economy of scale) into a concave function and thus a generation mix no longer exist. The minimum point is given by $x_c = 0$ (no coal plants).

Although in Figure 19 no coal ($x_c = 0$) is the optimal, coal may dominate if the cost for nuclear is much higher. In fact it can be checked that when $C_n = 145$, either coal only or nuclear only is optimal whereas when $C_n = 170$ coal plants will dominate. Also if the parameters were such that $m_n > m_c$ and $C_n > C_c$, it is possible that a mixture of plant types is optimal.

The effect of economy of scale is significant within the scope of the static model presented here. It can be perceived that as demand increases the advantage due to economy of scale is also increased and eventually the generation type which has more significant scale economy will completely dominate the whole capacity. Also it is conceivable that when a new technology is to be introduced the scale economy plays the key role.

Table 1. Economy of Scale Factors

	m_{fossil}	$m_{nuclear}$
Crowley (1978)	--	0.45
Comtois* (1977)	0.81	0.86
Spinrad (1980)	0.67	0.5
Lee (1978)	0.73	0.51
Fisher (1979)	0.84	--
Lucas (1979)	--	0.7
Abdulkarim & Lucas (1977)	0.8	--

* escalation included.

Note: This table is not for the purpose of making any comparisons.

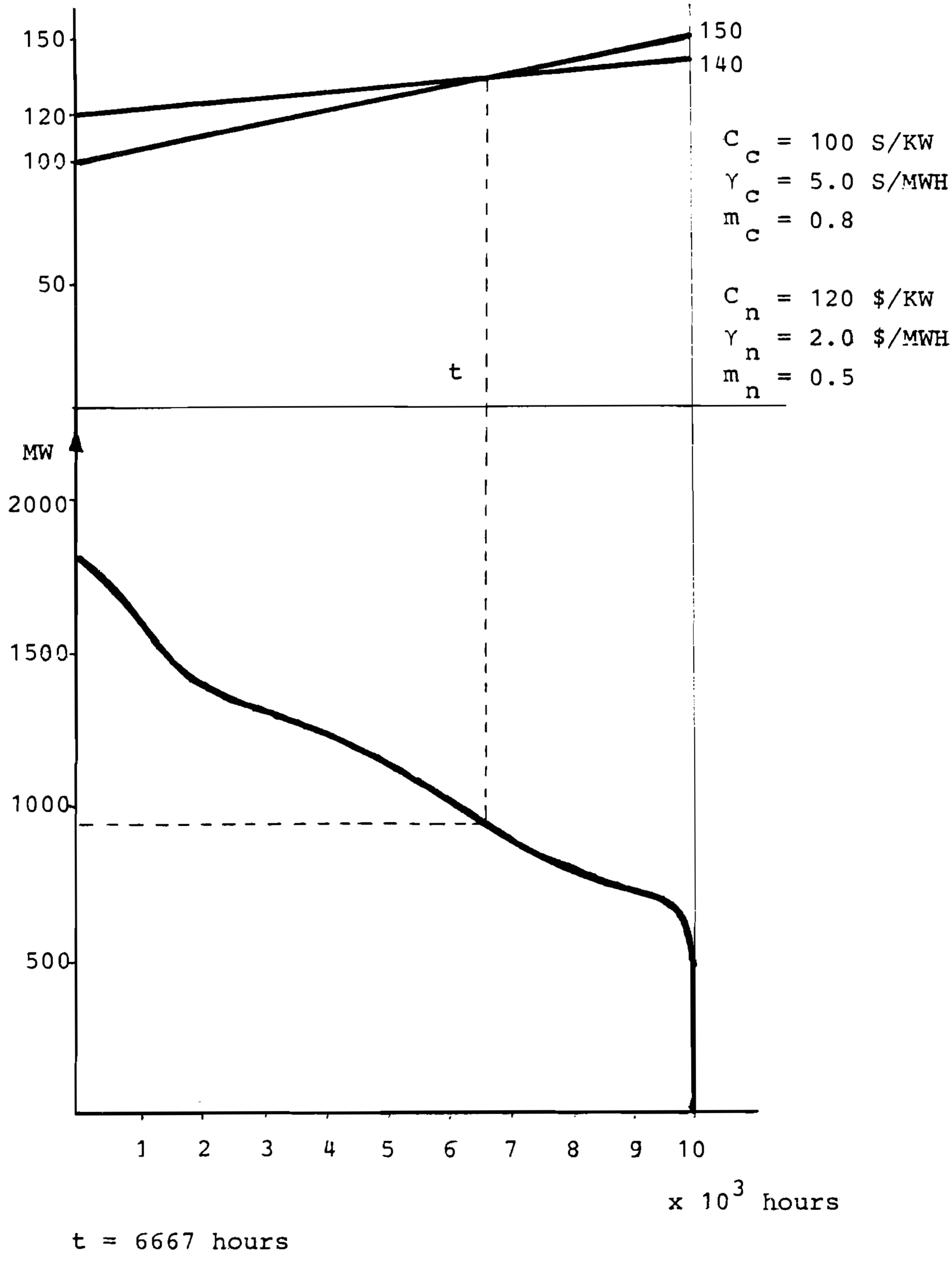
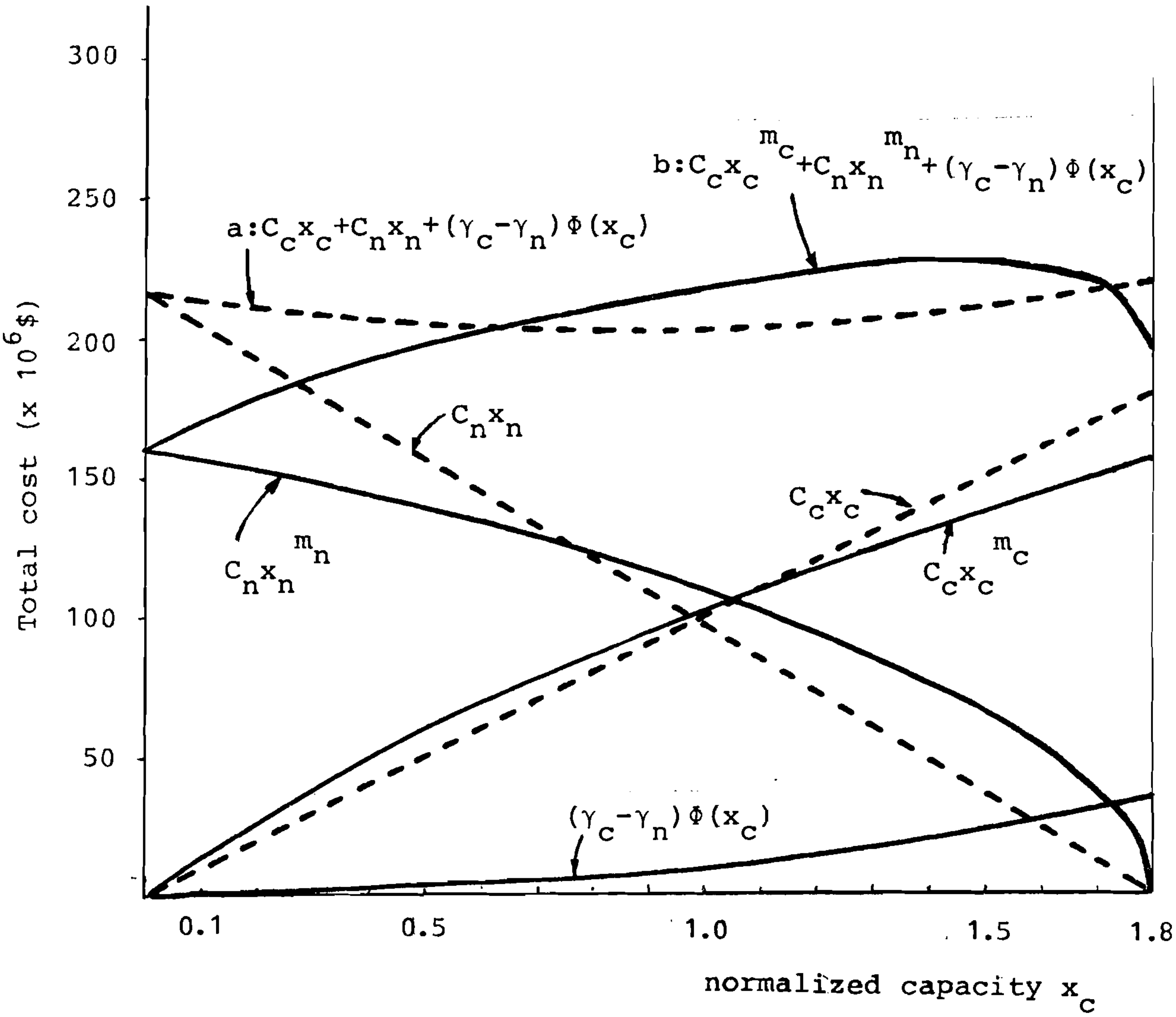


Figure 18. A Numerical Example



a: total cost curve with no economy of scale

b: total cost curve with economy of scale

Figure 19. Effect of Economy of Scale

THE EFFECT OF UNCERTAINTY ABOUT THE FUTURE ON SCALE DECISIONS

The question "Does uncertainty about the future result in decision makers' choosing a plant size which is different to that appropriate when the future is assumed known?" seems to be one which can give a wide variety of answers. Some people feel that the plant should be smaller, others feel that the plant should be larger.

The question and the outcome of the decision are sufficiently well understood that most managers and their advisors can give an opinion, yet the variety of answers indicates that, in fact, deciding on plant size when there is uncertainty about the future requires systematic and careful analysis.

Before developing any model it is necessary to clarify

(i) the sources and nature of uncertainty about the future
(ii) how this particular decision on size relates to future decisions
(iii) the criterion for choosing plant size when there is uncertainty about the future.

Sources and Nature of Uncertainty

In deciding on the size of plant there can be uncertainty about a wide range of factors, for example

(a) the accuracy of the cost estimates and the nature of the scale economies
(b) the time required to construct the plant and bring it into operation
(c) the operating cost and other performance indexes
(d) the arrangements and costs of financing plant construction
(e) the future markets for the product and the price at which it can be sold.

For each of these factors it is necessary to clarify the nature of the uncertainty. If quantitative models for deciding on plant size are to be used it will be necessary to estimate the probabilities of the uncertain events. If the events are repetitive past experience can be used to estimate the probabilities but if the event is unique a subjective approach must be used.

As a specific example consider some alternative approaches to describing uncertainty in future demand. Suppose it is known that demand is characterized by arithmetic growth. However, this could mean

(I) at the time the decision on plant size is made the rate of increase of demand is not known. There are several alternatives: consider the case of two alternatives labelled H (high) and L (low). However, it is known that after the plant is built the rate of increase can be observed and it will remain at the observed rate in all future time periods.

(II) the demand increase in each future time period is uncertain. It can either be H or L. The probability of each value are known and they are independent of the increase in the previous time period.

(III) again demand increase in each future time period is uncertain but probabilities in one period are dependent on the actual realized increases in previous time periods.

One can represent each of these alternatives by an event tree (Figure 20). It is important for the decision maker to recognize which of these alternatives describes his view of the uncertainty of future demand as there is often confusion between (I) and (II).

Relation Between This Decision and Future Decisions

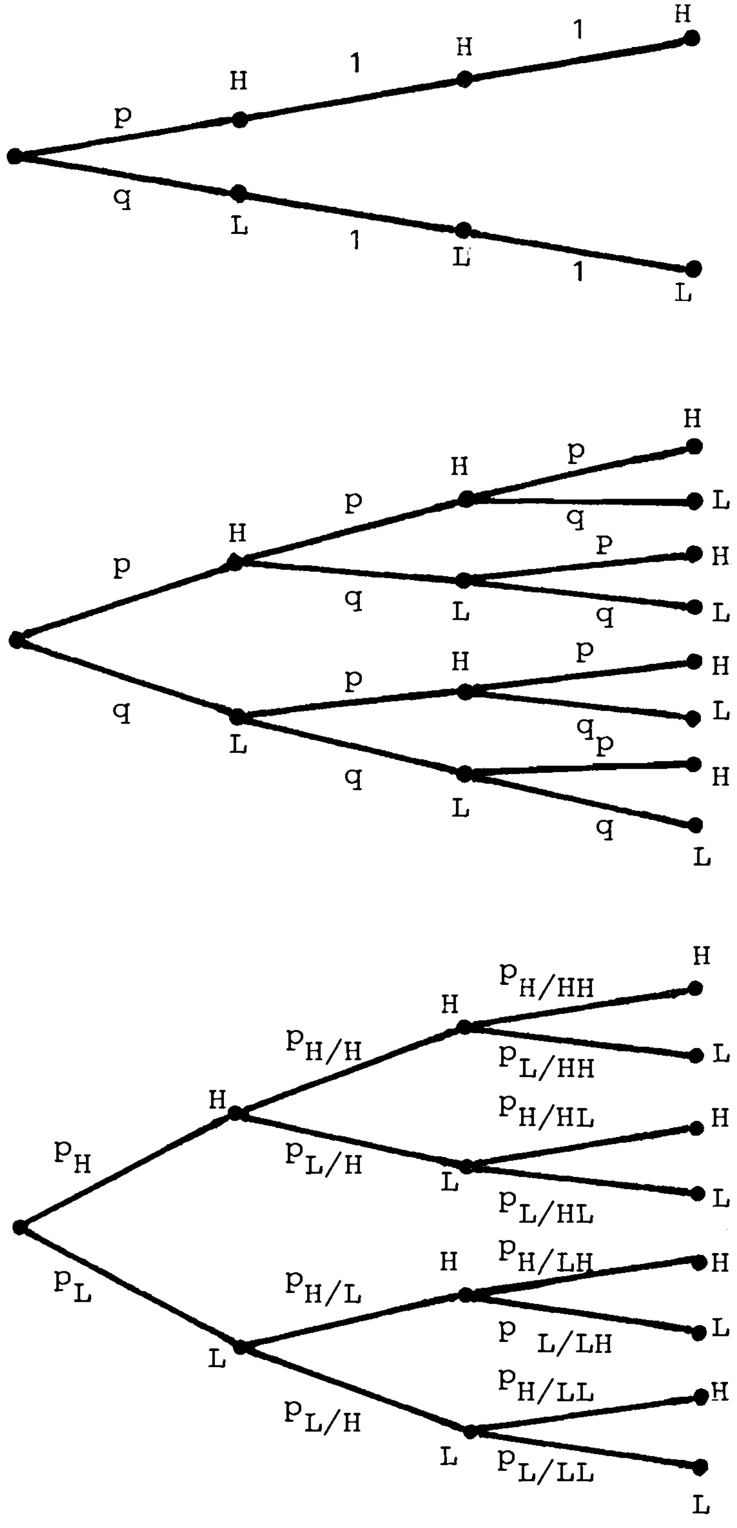
In some cases a decision is unique, in other cases it is part of a time sequence of decisions.

For example, the present decision on plant scale might imply that the same size of plant will be used for all future plant additions. This corresponds to the situation where the decision is actually to standardize on a particular plant design and only make minor changes in future plants. It is then not possible to change the size if the demand turns out to grow at a different rate than expected. We call this an open loop decision (cf. Bellman & Dreyfus 1965).

Alternatively, every time a plant addition is required in the future it will be possible to modify the size of plant in accordance with the demand pattern which has been experienced and the revised expectations about the future. We call this a closed loop decision.

The distinction is apparent if we consider alternative (I) Figure 20. The size resulting from the open loop decision made at time zero will be repeated irrespective of whether demand has turned out to be H or L. However, in the closed loop case the size of plant chosen at subsequent times will depend on whether demand has turned out to be H or L and, in either case, it will not be the same size as that selected at time zero.

In alternative (II) in Figure 20 there is a repetitive structure in which the uncertainty about the future always looks the same no matter what the past experience has been. So in this situation there is no difference between the open loop and the closed loop decision.



I

uncertainty about value of demand increase

II

demand increase in each period independent

III

demand increase correlated

Figure 20. Event Trees Describing Uncertainty About Future Demand

Criterion for Decision Making Under Uncertainty

Elementary decision theory usually recommends that the decision maker choose the appropriate course of action by finding the action which has minimum expected cost. (This is the criterion which has been used in almost all the literature on electricity generation expansion planning under undertainty (Tsuji 1980).) However, it is well known that with uncertainty there are a variety of other plausible criteria that can be used (Sage 1977).

Example

The difference between the different criteria and between open loop and closed loop decisions can best be illustrated by a simple example.

Suppose it is known that demand is increasing linearly but there is uncertainty about the rate of increase at the time when the decision on plant size is made. All other relevant parameters are known, e.g., the discount rate r and the economy of scale parameter m.

Now if the demand increase were known to be D* per period the optimal plant size would be that value of V minimizing

$$C(V;D^*) = \frac{kV^m}{1 - e^{-rV/D^*}}$$

which is known to be

$$V^* = X^*D^*$$

where X* is the solution of

$$X^* = \frac{m}{r}(e^{mX^*} - 1)$$

If m = .7, r = .1 the optimal size is then V* = 6.75D*.

Now consider the case where D is uncertain. Suppose that the nature of the uncertainty is best described by a situation analogous to Figure 20 (I), i.e., D can take on one of a finite number of values but once the particular value is realized it will remain at that value.

Our aim is to investigate what would be our best choice under the presence of uncertainty for different criteria.

1) Mathematical Expectation

In this case we assign subjective probability p_i that the demand rate be D_i, $i \varepsilon I$ and take expectation of the cost function over D. The forms of the expected cost function are given as follows:

For an open-loop decision structure,

$$\bar{C}_1(V) = \sum_{i \varepsilon E} C(V;D_i)p_i$$

For a closed-loop decision structure,

$$\bar{C}_2(V) = V^m + \sum_{i \varepsilon I} e^{-\frac{rV}{D_i}} C(V_i;D_i)p_i$$

where V_i is the value of V which minimizes $C(V;D_i)$, i.e.,

$$C(V_i;D_i) = \min_V C(V;D_i)$$

thus

$$V_i = X^*D_i \quad .$$

The optimal size can be obtained by minimizing either $\bar{C}_1(V)$ or $\bar{C}_2(V)$ with respect to V.

2) Min-Max Criterion (Optimist)

An optimist will assume that the demand goes higher. Thus he will be minimizing the cost function corresponding to the higher demand rate.

For the open-loop decision structure,

$$\min_V \max_i C(V;D_i) = \min_V C(V;D_{max})$$

For the closed-loop decision structure,

$$\min_V \max_i \{V^m + e^{-\frac{rV}{D_i}} C(V_i;D_i)\} = \min_V C(V;D_{max})$$

Note that in both cases he will choose the plant size appropriate to the maximum demand.

3) Min-Min Criterion (Pessimist)

This case is the reverse of 2) above.

For the open-loop decision structure,

$$\min_V \min_i C(V;D_i) = \min_V C(V;D_{min})$$

For the closed-loop decision structure,

$$\min_V \min_i \{V^m + e^{-\frac{rV}{D_i}} C(V_i;D_i)\} = \min_V C(V;D_{min})$$

Note that in this case he will choose the plant size appropriate to the minimum demand.

4) Min-Max Regret Criterion

We can define the deviation (loss) if the demand rate turned out to be different from what a decision maker had assumed to be. This criterion tries to choose the decision which minimizes maximum loss incurred by taking certain decision.

For the open-loop decision structure,

$$\min_V \max_i \{C(V;D_i) - C(V_i;D_i)\}$$

For the closed-loop decision structure,

$$\min_V \max_i \{V^m + e^{-\frac{rV}{D_i}} C(V_i;D_i) - C(V_i;D_i)\}$$

5) Raplace Criterion

This is formally the same as in 1), but here p_i are chosen to be all equal.

6) Expectation plus Variance

In this case not only mathematical expectation but also some measure which expresses variance of the cost function is to be minimized.

For the open-loop decision structure,

$$\min_V \{\bar{C}_1(V) + \beta \sum_{i\varepsilon I} |C(V;D_i) - \bar{C}_1(V)|^2 p_i\}$$

For the closed-loop decision structure,

$$\min_V \{\bar{C}_2(V) + \beta \sum_{i\epsilon I} \left| V^m + e^{-\frac{rV}{D_i}} C(V_i;D_i) - \bar{C}_2(V)\right|^2 p_i\}$$

where β is a weighting factor.

Table 2 presents a numerical example for which m = 0.7 and r = 0.1, and D ranges from 0.8-1.2. In the case of taking expectation, we assumed that D takes on 0.8, 1.0 or 1.2 with probability 0.25, 0.5 and 0.25 respectively. Note that D* = 1.0.

It is apparent that the closed-loop decisions always give lower values, for example,

$$\begin{aligned} \bar{C}_1(V) &= \sum_{i\epsilon I} p_i \; C(V;D_i) \\ &= \sum_{i\epsilon I} p_i\{V^m + e^{-\frac{rV}{D_i}} C(V;D_i)\} \\ &\geqslant V^m + \sum_{i\epsilon I} e^{-\frac{rV}{D_i}} C(V_i;D_i) p_i \\ &= \bar{C}_2(V) \quad . \end{aligned}$$

So in this example the closed-loop decision is always the better choice.

In this example, the optimal size when there is no uncertainty is 6.75. Table 2 demonstrates that the optimal size becomes either larger or smaller depending on the criterion to be used.

Table 2. Effect of Uncertainty on Optimal Size

Criteria	Open-loop decision	Closed-loop decision
Expectation	6.73	6.6
Min-max (optimist)	8.1	8.1
Min-min (pessimist)	5.4	5.4
Min-max regret	6.7	6.6
Laplace	6.73	6.5
Expectation plus weighted variance (β = 0.2)	7.0	6.8
Deterministic case	6.75	

Comment

It can be seen that the answer to the question of how uncertainty about the future effects the scale decision is quite complex even though the decision is well understood.

A quote (Betts, private communication) will perhaps give a flavor of the approach used by actual decision makers in coping with uncertainty.

> Other than in the short-term, forecasting in the present state of world economic and political turbulence, is largely a speculative art. The response of preparing alternative scenarios may look impressive on paper, but the commitment of finance to large-scale manufacturing hardware requires a positive decision. There are, however, the classic and obvious responses to uncertainty in the bulk chemicals and other capital intensive industries. These include delaying and phasing the degree of investment commitment as long as possible, sharing the advantages of plant scale by asset sharing, the vertical and horizontal integration between raw materials and products, and building as much flexibility as practicable in respect of feedstock variations and future plant extensions. Plant flexibility costs money, which may or may not in the event be fully utilized. The ability of the industry to pay for this ideal is influenced by the price competitiveness of the market for the particular product.
>
> I believe that each such decision is a unique exercise, and is very much conditioned by past experience (the learning process) and the internal and external circumstances applying at the time.

CONCLUDING REMARKS

The mathematical models required to guide decision makers in understanding the significant factors determining size and process mix are quite well developed. As has been shown in the last section they can also be extended to allow for uncertainty about the future once the appropriate scenario has been defined and the criterion for selecting the best course of action determined.

However, there needs to be more work done on assessing the extent to which these models of rational behavior describe the actual decisions by firms. Furthermore, greater understanding is required of the implications of the use of these models of rational decision making on the aggregate behavior of the industry. Our discussion on plant size, technological substitution and process mix represents a first step in this direction.

It would be desirable to take a specific situation, such as the dynamics of scale and process mix in the electricity generating industry, and see whether the evolution over time can be explained by the growth rate in demand, the economies of scale in capital costs of the different technologies the relative operating costs and the appropriate distribution of firm size. If such a model can provide an adequate explanation of past changes in the industry then it has the great advantage over aggregate models of learning and growth in that it suggests how policy intervention can modify behavior and how new technological developments will affect the industry.

In spite of our preference for explaining aggregate behavior by rational decision making at the level of the firm, it is remarkable how well models of learning and growth describe aggregate behavior. It may be that this is due to the logarithmic transformations sufficiently smoothing the data that it can be fitted using simple functions. Obviously more work needs to be done to develop an understanding of why these models work so well. It would seem that learning models would be even more appropriate in describing the way the technological characteristics of processes change, for example, the growth in maximum feasible plant size or the reduction in cost of a particular plant size with increasing experience.

Finally, there is a need for careful case studies of the way actual plant sizes are determined. It is necessary to develop a better understanding of how decision makers view the uncertainty about the future and on what criteria they base their decisions. This could be a useful study for IIASA, in particular to compare decision makers in different countries and in market and planned economies.

REFERENCES

Abdulkarim, A.J., and N.J.D. Lucas. 1977. Economies of Scale in Electricity Generation in the United Kingdom. Energy Research. 1:223-231.

Bellman, R. and S.E. Dreyfus. 1965. Applied Dynamic Programming. Princeton University Press.

Buzacott, J.A. 1980. Scale and Process Innovation: The Adoption of the Basic Oxygen Process by Canadian Steel Firms. WP-80-83. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Comtois, W.H. 1977. Economy of Scale in Power Plants. Power Engineering. August:51-53.

Crowley, J.H. 1978. Power Plant Cost Estimates Put to the Test. Nuclear Engineering International. July. pp.39-43.

Erlenkotter, D. 1967. Two Producing Areas--Dynamic Programming Solutions. Chapter 13 in Manne 1967.

Erlenkotter, D. 1974. A Dynamic Programming Approach to Capacity Expansion with Specialization. Management Science. 20:360-362.

Fisher, J.C. 1979. Optimum Size of Subcritical Fossil Fueled Electric Generation Units. Paper presented at the Workshop "Size and Productive Efficiency--The Wider Implications" June 1979. To be published in Buzacott, J.A. et al. ed. Scale in Production Systems. Oxford: Pergamon Press.

Harders, F. 1971. Rationalisierungsmöglichkeiten in der Stahlindustie unter gross wirtschaftlichen Gesichtspunkten und dem Aspekt der kommender 10 bis 20 Jahre (Possibilities of Rationalization in the Steel Industry from an Overall Economic Perspective and Prospects of the next 10 to 20 Years.) (in German) Stahl und Eisen. 91(5):241-249.

Hjalmarsson, L. 1974. The Size Distribution of Establishments and Finns Derived from an Optimal Process of Capacity Expansion. European Economic Review. 5:123-140.

Huettner, D.A. 1975. Scale, Costs and Environmental Pressures. Chapter 3 of B. Gold. ed. Technological Change: Economics, Management and Environment. Oxford: Pergamon Press.

Ijiri, Y. and H.A. Simon. 1977. Skew Distributions and the Sizes of Business Firms. Amsterdam: North-Holland.

Kalotay, A.J. 1973. Capacity Expansion and Specialization. Management Science. 20(1):56-64.

Kootz, Th., K. Lenzmann, A. Altheld. 1973. Gegenüberstellung der Stahlgewinnungs Verfahren aus technischer und wirtschaftlicher Sicht mit Betrachtungen über die weitere Zukunft (Comparison of Steel Making Processes from Technical and Economic Perspectives with Thoughts on the Long-Term Future). (in German). Thyssenforschung 5(2):59-77.

Lee, T.H. 1978. Optimization of Size in Power Generation. Talk given at International Institute for Applied Systems Analysis, Laxenburg, Austria, December 1978.

Lucas, N.J.D. 1979. Economies of Scale in Nuclear Power Plant. Energy Research. 3:297-300.

Manne, A.S. 1967. Investments for Capacity Expansion: Size, Location and Time-Phasing. London: George Allen and Unwin.

Marchetti, C. and N. Nakicenovic. 1978. Volume 1: Phenomenological Part. in C. Marchetti, N. Nakicenovic, V. Peterka and F. Fleck. The Dynamics of Energy Systems and the Logistic Substitution Model. AR-78-1. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mazur, J.E. and R. Hastie. 1978. Learning as Accumulation: A Reexamination of the Learning Curve. Psychological Bulletin. 85(6):1256-1274.

Peck, S.C. 1974. Alternative Investment Models for Firms in the Electric Utilities Industry. Bell Journal of Economics and Management Science. 5(2):420-458.

Peterka, V. 1978. Macrodynamics of Technological Change: Market Penetration by New Technologies. in C. Marchetti, N. Nakicenovic, V. Peterka and F. Fleck, The Dynamics of Energy Systems and the Logistic Substitution Model. AR-78-1. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Philips, D., F.P. Jenkin, J.A.T. Pritchard and K. Rybicki. 1969. A Mathematical Model for Determining Generating Plant Mix. Power System Computation Conference, Rome, 1969.

Resch, W. 1973. Stahlerzeugung in Japan--Technische Stand und Betriebsergebnisse (Steel Production in Japan--Technological Statics and Economic Performance) (in German).

Sage, A.P. 1977. Methodology for Large Scale Systems. New York: McGraw Hill.

Sahal, D. 1979a. The Temporal and Spatial Aspects of Diffusion of Technology. IEEE Transaction on Systems, Man and Cybernetics. Vol. SMC-9, 12, pp.829-839.

Sahal, D. 1979b. Laws which Govern the Maximum Capacity of Technology. Unpublished paper. Berlin: International Institute of Management.

Schlaepfer, F.M. 1978. Singular Optimal Control in Generation Expansion Planning. IEEE 1978 PES Winter Meeting, Jan-Feb. 1978. Paper A-78 262-8.

Simmonds, W.H.C. 1972. The Analysis of Industry Behavior and its use in Forecasting. Technological Forecasting and Social Change. 3:205-224.

Spinrad, B.I. 1980. Scaling and Learning in Nuclear Energy. CP-80-xx. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Srinivasan, T.N. 1967. Geometric Rate of Growth of Demand. Chapter 9 in Manne, 1967.

Tsuji, K. 1980. The Effects of Uncertainty in Generation Expansion Planning--A Review of Methods and Experiences. WP-80-xx. Laxenburg, Austria: International Institute for Applied Systems Analysis.

von Bertalanffy, L. 1968. General Systems Theory. New York: Braziller.