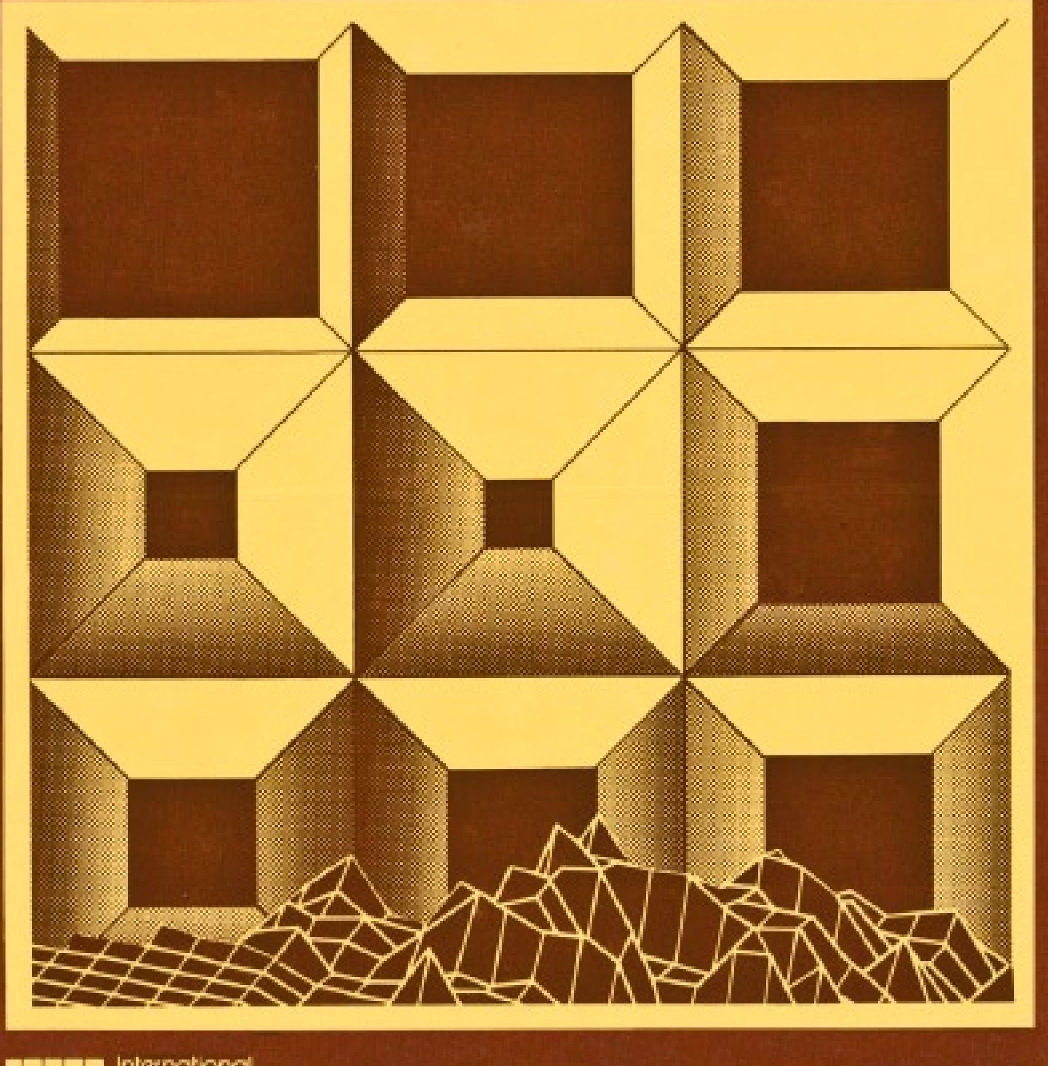
IIASA PROCEEDINGS SERIES

Scale in Production Systems

John A. Buzacott
Mark F. Cantley
Vladimir N. Glagolev
Rolfe C. Tomlinson, Editors



International
Institute for
Applied
Systems
Analysis

IIASA PROCEEDINGS SERIES

Volume 15

Scale in Production Systems

IIASA PROCEEDINGS SERIES

1 CARBON DIOXIDE, CLIMATE AND SOCIETY
Proceedings of an IIASA Workshop Cosponsored by WMO, UNEP, and SCOPE,
February 21–24, 1978
Jill Williams, Editor

2 SARUM AND MRI: DESCRIPTION AND COMPARISON OF A WORLD MODEL AND A NATIONAL MODEL
Proceedings of the Fourth IIASA Symposium on Global Modeling,
September 20–23, 1976
Gerhart Bruckmann, Editor

3 NONSMOOTH OPTIMIZATION
Proceedings of an IIASA Workshop,
March 28–April 8, 1977
Claude Lemarechal and Robert Mifflin, Editors

4 PEST MANAGEMENT
Proceedings of an International Conference,
October 25–29, 1976
G.A. Norton and C.S. Holling, Editors

5 METHODS AND MODELS FOR ASSESSING ENERGY RESOURCES
First IIASA Conference on Energy Resources,
May 20–21, 1975
Michel Grenon, Editor

6 FUTURE COAL SUPPLY FOR THE WORLD ENERGY BALANCE
Third IIASA Conference on Energy Resources,
November 28–December 2, 1977
Michel Grenon, Editor

7 THE SHINKANSEN HIGH-SPEED RAIL NETWORK OF JAPAN
Proceedings of an IIASA Conference,
June 27–30, 1977
A. Straszak and R. Tuch, Editors

8 REAL-TIME FORECASTING/CONTROL OF WATER RESOURCE SYSTEMS
Selected Papers from an IIASA Workshop,
October 18–20, 1976
Eric F. Wood, Editor, with the Assistance of Andràs Szöllösi-Nagy

9 INPUT–OUTPUT APPROACHES IN GLOBAL MODELING
Proceedings of the Fifth IIASA Symposium on Global Modeling,
September 26–29, 1977
Gerhart Bruckmann, Editor

10 CLIMATIC CONSTRAINTS AND HUMAN ACTIVITIES
Selected Papers from an IIASA Task Force Meeting,
February 4–6, 1980
Jesse Ausubel and Asit K. Biswas, Editors

11 DECISION SUPPORT SYSTEMS: ISSUES AND CHALLENGES
Proceedings of an International Task Force Meeting,
June 23–25, 1980
Göran Fick and Ralph H. Sprague, Jr., Editors

12 MODELING OF LARGE-SCALE ENERGY SYSTEMS
Proceedings of the IIASA/IFAC Symposium on Modeling of Large-Scale Energy Systems,
February 25–29, 1980
W. Häfele, Editor, and L.K. Kirchmayer, Associate Editor

13 LOGISTICS AND BENEFITS OF USING MATHEMATICAL MODELS OF HYDROLOGIC AND WATER RESOURCE SYSTEMS
Selected Papers from an International Symposium,
October 24–26, 1978
A.J. Askew, F. Greco, and J. Kindler, Editors

14 PLANNING FOR RARE EVENTS: NUCLEAR ACCIDENT PREPAREDNESS AND MANAGEMENT
Proceedings of an International Workshop
January 28–31, 1980
John W. Lathrop, Editor

15 SCALE IN PRODUCTION SYSTEMS
Based on an IIASA Workshop
June 26–29, 1979
John A. Buzacott, Mark F. Cantley, Vladimir N. Glagolev, and Rolfe C. Tomlinson, Editors

SCALE IN PRODUCTION SYSTEMS

JOHN A. BUZACOTT, MARK F. CANTLEY
VLADIMIR N. GLAGOLEV, ROLFE C. TOMLINSON

Editors

PERGAMON PRESS

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon Press Canada Ltd., Suite 104, 150 Consumers Rd., Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, 6242 Kronberg-Taunus, Hammerweg 6, Federal Republic of Germany

First edition 1982

Library of Congress Cataloging in Publication Data
Main entry under title:
Scale in production systems.
(IIASA proceedings series ; v. 15)
"Based on a workshop held at the International Institute for Applied Systems Analysis in June 1979"—Pref.
1. Economies of scale—Congresses. I. Buzacott, John A. II. International Institute for Applied Systems Analysis. III. Series.
HD69.S5S27 1982 658.4'02 81-22666
ISBN 0-08-028725-5 AACR2

Printed in Great Britain by A. Wheaton & Co. Ltd., Exeter

PREFACE

This book is concerned with the problem of scale, and is based on a workshop held at the International Institute for Applied Systems Analysis in June 1979 as part of the research program of the Management and Technology Area. It was about 15 years ago that I first seriously encountered the question of scale as an analytical problem. This was in the course of a consultancy assignment that my research team had been asked to undertake for the UK Health Service. This work had started because the chief hospital architect asked us to look at the relationship between design criteria and the effectiveness of various service functions in hospitals. "Scale" was clearly a matter of some importance here, and in due course we were led on to consider the point at which a hospital should have its own centralized services functions, rather than become part of a larger network. Subsequently, the Department of Health developed its own OR/systems team, and we were drawn into a much larger study designed to prepare policy recommendations with regard to the size of district hospitals. These guidelines went some way to halt the trend of continually increasing hospital size. Nevertheless, a very wide margin was allowed for local judgment, and there was still no adequate analytical base for making those local decisions. Shortly after this, my interest received an even stronger impetus. The oil crisis at last became recognized in the West and the National Coal Board, for whom I was working, was – after a 25-year hiatus – able to start planning new mines again. In the planning of a coal mine, a decision on size has to be made early; and yet, as I questioned my team, I found that we could not then propose a satisfactory analytical basis for such a decision. After some years of involvement in the planning process, I can say that the situation is now somewhat better. Even so, there is no generally available methodology for studying problems of scale, even though apparently similar problems appear in many industries; it is widely recognized that there have been mistakes of scale.

It was then this common problem, this widespread concern with regard to the general trend towards "giantism," and the lack of an accepted methodological approach, that led us to consider a research project in the Area. The first step was to decide whether this really was a problem worthy of study at *IIASA*. Not everybody agreed to this. When we first started to develop the work, we were asked to look at a particular case of electricity generation. We found ourselves under the criticism that this was a technological problem which had been adequately studied and solved by the experts in the field. For a whole variety of reasons we were advised that economies of scale in generating efficiency could be taken for granted. We had not thought to question this at the level of the generating unit, but had thought that more research would be justified on the impact of large units on the distribution system and the economy as a whole. We soon found however that it was unnecessary to be defensive, since we quickly came across evidence (Chapter 4) that we are already building generating units larger than the maximum economical size. The criteria of scale even in such a narrow technical field were still a problem.

The preliminary work was undertaken by Mark Cantley and Vladimir Glagolev and their first task was to structure the subject and identify the state of the art from the literature. The results of this work were summarized in a Research Memorandum, '"Problems of Scale" – the Case for *IIASA* Research' (RM-78-47). Much of the Material from that document is included in the discussion sections of this book.

A number of difficulties arose in moving forward from this point. The first, as we soon discovered, was that in studying the problem of scale we were faced with a problem of scale of our own. The problem could be looked at on many levels, from many specializations, and with a very wide variety of applications. It was for this reason that we decided in the first place to study the problem at the lowest production level and thus to look at size and productive efficiency. In order to discuss the research ideas developed at *IIASA*, and to identify present patterns of research, we decided to hold a workshop entitled "Scale and Productive Efficiency – The Wider Implications;" this book is a consequence of that workshop. It should be pointed out, incidentally, that the discussion went far beyond a consideration of production units. We did not regret narrowing the topic, however, since the decision gave a necessary focus for our work discussions.

Secondly, we found it increasingly difficult to isolate questions of size from other questions; for example, one cannot separate questions of organizational size from those of organization design and control. Such a separation can lead to misunderstanding. There is a tendency for people to assume that because we want to be able to determine the correct or appropriate size, we are "anti-big" or "pro-small." Some felt therefore that the work was directed against multinationals, others that it was against central planning. It is against neither. It is *for* the best solution in all political structures. In the same way, we found that a discussion on size was liable to become a discussion on innovation or learning or some other associated issue. Size is a pervasive concern, and more worthy of study because of it.

Thirdly, we found ourselves in danger of forgetting the purpose of the research. The need to understand and structure such a complicated subject led us inevitably into reductionism; we subdivided the topic into many discrete elements and investigated each element separately. We were thus able to examine a number of quite different problems related to scale, explore the state of the art, undertake research, and drive through to some satisfactory conclusion. Unfortunately, this traditional scientific approach will not in itself take the decision maker forward. He is faced with the decision problem: How large? Even if all the individual scientific problems are solved, those solutions will not in general add up to an answer to that one single problem. The problem of synthesis remains if the dicision maker is to be helped, and this is the special function of the systems analyst.

Finally, we found that this was a topic that cuts across many academic disciplines, notably economics, engineering and the social sciences. Yet most of the previous research had been done on a disciplinary basis. There was no established international, interdisciplinary community of people who were concerned with this problem. Our workshop, in fact, established such a community for the first time. An incidental consequence of the lack of such a community was the difficulty of identifying, through traditional scientific channels, the appropriate experts for our workshop. Many of our most valued contributors heard about it by chance. This meant that the workshop brought together, for the first time, representatives from industry, government, and the universities, and further, from different industries, disciplines, and nationalities. Discussion and controversy began after the first presentation and continued unabated throughout. We introduced an additional discussion session on the last evening and the final adjournment had to be made with discussion in full swing. There seemed to be a need for a new community.

This book is in part the proceedings of this workshop, but we have tried to create a more coherent document than most proceedings. Out of the 20 chapters, 12 are based on presentations given at the workshop and the remaining 8 have been written by the Management and Technology *MMT* team, incorporating both their own research and the discussion within the workshop so as to provide an overview of the field as we see it. Nevertheless, this is not a handbook on scale, i.e., it is not something that can be handed to a manager or decision maker to tell him what to do, though it should be of interest to them as well as to analysts and research workers. Thus, it both structures the subject and provides a good deal of practical case material analyzed by leading workers in the field. We hope that it will be of practical value as well as the starting point for much future research.

Some mention must be made of those responsible for the work. As has been indicated, the work in the 15 months leading up to the workshop was undertaken by Mark Cantley (UK) and Vladimir Glagolev (USSR). Immediately after the workshop, the leadership of the task was taken up by John Buzacott (Canada), who is the chief editor of this volume, though reference should also be made to the contribution of Kiichiro Tsuji (Japan) who joined the team at the same time and who has been deeply involved in the discussions that determined the final shape of the book. The main credit must, of course, go to those who came to the workshop and participated in such a lively and constructive fashion. Many more papers were given than we have been able to include in this volume and more people contributed to the discussion than we have been able to give individual reference to. The agenda and full list of attendees are given in the Appendixes.

Finally, thanks are due to Miyoko Yamada who, as secretary to the team responsible for the study, undertook many of the administrative chores associated with the workshop as well as the typing for this volume. She has been a marvel of friendly competence and devotion to duty.

Rolfe Tomlinson
Chairman
Management and Technology Area

CONTENTS

Part 1

INTRODUCTION

CHAPTER 1 PROBLEMS OF SCALE

1.1 THE BASIC QUESTIONS

Scale is more than size: it is size with proportions and consequences. When proportions are no longer in harmony, or consequences are unanticipated, we have "problems of scale."

The diversity of such problems was reflected in the interests and disciplines of the participants in the workshop that gave rise to this book. The workshop therefore started with some discussion of terminology and classification, in order to facilitate debate and mutual understanding. A similar purpose informs this opening chapter.

The range of material presented is wide, yet the applicability of some of the concepts is much wider than could be presented here. The scope was nominally limited to production systems, and therefore the problems and purposes considered are mainly based on the artifacts and organizations of manufacturing. But the boundaries are fluid; for organizations are groups of human beings, and our artifacts are derived from and set in a natural environment which displays size, form, and function, from subatomic particles to the limits of the universe.

1.1.1 Scale in Production Systems – What are the Questions?

There are many well-documented examples of the performance improvements achievable by increasing the scale of operation. Table 1.1, from the Soviet Union, provides a good example. Similar phenomena have been reported in many other branches of process industry, particularly those characterized by continuous production, and in mining. Given high utilization levels of capacity in such industries, the capital investment per unit of output and the material and fuel costs are lower with large-scale equipment.

In industries where capacity increase is achievable only by adding more units of capacity of the same size and technology, the ratio of resource inputs to outputs may be similar in small or large plants. Performance increase may depend more on improvement of production methods, organization, and management than on scale.

TABLE 1.1 Index of performance characteristics of thermal power stations (200 MW = 100).

	Thermal power station capacity (MW)				
	200	300	600	1,200	2,400
Capital investment per MW	100	86	75	66	60
Main building volume per MW	100	88	84	58	51
Construction and installation (as % of total capital cost)	66.5	64.0	60.0	50.5	45
Operating personnel per MW	100	84	60	32	24
Electrical energy production cost	100	91	87	78	70

SOURCE: Bolotryi and Itin (1976).

Both Eastern (Kosygin 1976) and Western sources (Gold 1955) show the systematic efforts made to achieve greater production efficiency through economies of scale, by specialization and concentration of production, and by rationalization and standardization of products and processes. Glagolev (1973, 1975), in studies of labor mechanization and automation in Lithuanian industires, concluded that the optimal scale of homogeneous production, together with the development of specialization and standardization, form the basis for increasing efficiency in the auxiliary production sectors of industry in the Soviet Union.

In the West, many industries have undergone a similar process of restructuring for greater efficiency, and have concentrated production in highly automated factories. Yet the advent of low-cost information technology (microprocessor controlled equipment, machine tools, automatic handling equipment, and transfer lines) may alter traditional economies of scale. These and other environmental changes pose continually varying questions about the management of scale in technological systems.

We study problems of scale to improve management decision making. We seek to learn as much as possible from past mistakes, from one another, and from decisions and disciplines which, however remote from our own, may nonetheless contain relevant experience. Investigating problems of scale, be they of organizations or of ethylene plants, is not a routine activity for production managers, but is typically delegated to specialist engineers and designers. However, the larger the scale envisaged, the less relevant experience is likely to be available; and the more experienced the specialist designer, the more aware he is likely to be of his need to understand how the system he is planning will operate in practice.

Considering the variety of political, social, economic, and environmental factors encountered in different countries and regions, and the special technologies, organizational peculiarities, objectives, managerial culture, and customer behavior in different industries, one returns repeatedly to two basic sets of questions that underlie research on scale.

- To what extent can scale be perceived as a general problem, or must problems of scale be peculiar to each activity (mining, manufacturing, social services), or each subactivity (coal mining, machine tools, health care)? Alternatively, is the problem of scale unique to the specific situation of the decision maker?
- Could a completely general methodology be developed for determining the size of hospitals, supertankers, agricultural farms, industrial plants, research and development

organizations, and so on? Alternatively, is there a general methodology for each of the "levels" of scale introduced in section 1.2 below? Are there methodologies appropriate for general families of scale problems?

This book is an attempt to answer these questions.

The process of mutual learning requires a common language. The exchange of experience demands common terms for its description, measurement, and classification. In the remainder of this chapter, a common language is presented from traditional disciplines, from what was offered at the start of the workshop, and from the contributions and discussions that followed. Section 2 introduces concepts of functional significance in the consideration of scale: the distinct levels, the methods of measurement, the role of environment, and the diverse factors of scale. Section 3 reviews briefly the various disciplinary approaches that have been used in the study of scale problems.

This book seeks to integrate these conceptual approaches with the ideas and the empirical material presented at the workshop; section 4 of this chapter explains how the material in the following chapters has been organized.

1.2 FUNCTIONAL CLASSIFICATIONS

1.2.1 Levels of Scale

A useful subdivision of problems of scale is provided by the hierarchy of the following levels; the terms in italics will be used in this sense in the remainder of the book.

- Level 1 (a): the scale of a single unit of physical equipment – the *engineering level* or *unit level*. (b): the scale of a single product line (which might be produced by several separate units of equipment)
- Level 2: the scale of a single plant or factory (i.e., on, or based on, one site but possibly containing several engineering units or product lines) – the *plant level*

Levels 1 and 2 coincide in the case of a single-unit (or "single-train") plant, which typically depends on a single major component.

- Level 3: the scale of a single organization – the *corporate level* or *organization level*

Level 3 is less clearly definable in operationally unambiguous ways, and in terms capable of clear and standard interpretation in different countries. For instance, it may coincide with level 2 in a single-factory company. In a company comprising several plants engaged in similar activities, the plants might collectively be viewed as a single organization; but this company might itself be a subsidiary of a larger company. This membership in a larger unit could be relevant to financial and negotiating strength, but might be irrelevant to the company's technical efficiency. "Organizational level" thus requires careful definition, particcularly where comparisons are being made: a "big" organization could be "small" on the scale of its activities in a specific field.

- Level 4: the scale of a total industry or an industrial complex – the *industry level*

During recent decades, new organization forms of large-scale coordinated programs (for example, The Tennessee Valley Authority in the United States) and territorial/industrial complexes (for example, the Bratsk-Ilimsk territorial production complex in the Soviet Union) have come into being in both East and West. The distinctive feature of level 4 is the cooperation among organizations that in other or smaller-scale situations operate more independently. Even in a competitive environment, common pressures and the perception of common interests foster cooperative behavior at the industry level. The separate actors coordinate their efforts in order to obtain the benefits of the large-scale program or the organized group: independence is exchanged for the benefits of interdependence, albeit at a cost in the problems of coordination.

A national or societal level may be viewed as level 5, although in some situations it is overlapped by the supranational ramifications at level 4.

1.2.2 The Measurement of Scale

Given the diversity of dimensions in which scale can be measured in different industries and activities, it is impossible to divorce measure from function and purpose: What is the measurement for? What decision will be based on it?

If one is seeking to identify the stage of growth at which the need arises for a change of techniques and methods, and to identify this stage in different industrial and social contexts, then one wants measures of scale that are independent of the context, and with which different contexts can be compared. One possibility is to focus on quantities that are meaningful across many contexts and not specific to one industry. Some absolute measures mght be

- Number of people employed
- Physical area or volume occupied by a plant
- Physical mass or volume of daily or annual throughput
- Financial value of the capital employed
- Financial value of daily or annual output

Another possibility is to use relative quantities, such as the following ratios

$$\frac{\text{size of unit being considered}}{\text{size of largest existing unit}}$$

or

$$\frac{\text{size of unit being considered (capacity, annual output)}}{\text{size of relevant context}}$$

Finally we may consider performance ratios, i.e.,

$$\frac{\text{output of useful goods}}{\text{input of scarce resources}}$$

or

$$\frac{\text{Tons/year capacity}}{\text{capital cost}}$$

In manufacturing, all these measures have their roles. Physical measures and ratios will be relevant to engineering problems of scale, numbers of people and their ratio to space will be relevant to organizational and personnel questions, financial scale to questions of financing and insurance. Any of these measures may be problematic or limiting in their own dimensions.

The above measures all describe the system or organization at a particular moment. However, it is often more significant to have a measure that relates to the history or experience of the organization. The simplest such measure is the cumulative production of the product in units, which is used as the basis of the experience or learning curve discussed in more detail in Chapter 6. A related measure is the comparative market shares of different firms in an industry. This has been extensively propounded by the Boston Consulting Group as a basis for strategy formulation (Hedley 1976, 1977). Delombre and Bruzelius (1977) describe a case study from SKF group, a multinational precision engineering company. Their conclusion is that "the correct measure of competitive posture . . . is own market share/market share of biggest competitor." The logic is that the greater experience leads to lower costs; and cost "is relative, not absolute . . . no one knows what a cost ought to be . . . the low cost can at any point in time only be defined by the company which has achieved the lowest cost so far."

1.2.3 The Environment or Relevant Context

The word "environment" is commonly used with a very general meaning. In studying an economic or industrial entity at any level, the systems analyst views it as part of a system, a set of interrelated elements. The entities described by levels 1 to 4 above are not complete systems because the questions raised by scale alternatives have to take account of relationships with the environment. We use the term ***relevant context*** for those parts of the general environment that are relevant to the determination of appropriate scale in a particular case; in other words, the system to be studied is the entity (machine, factory, organization) *and* its relevant context.

In terms of the levels of scale introduced above, we may observe that the organizational environment for anything on level N is the system of level $N + 1$; or, on levels 3 and 4, the environment is other organizations. However, there are many dimensions in which to define *environment*. For example, in physical terms, all production systems are "open" systems, like biological ones, operating by interchange of information, energy, and materials between themselves and their environment; this is a very useful concept when supplies of energy or feedstock are problematic.

It may be argued that the relevant context or environment should be added as the final member of our list of levels, but this could be confusing since the concept of environment is used at all levels. Nonetheless, the scale of the relevant context has increased progressively with advances in the technology and scale of entities on levels 1 to 4.

The use of "market size," whether measured in physical, financial, or other units, has been introduced in the context of measurement of scale. In considering the relationship of an organization to its environment, ratios are the usual measures. Thus even if the organization is unchanged, changes in its environment may alter its relative scale.

What is the relevant context for deciding whether a unit is relatively large? It could be a world total, a national or regional total, or a total within the one organization. Like the word *strategic*, the term *large* is relative. The relevance of different base scales depends upon the degree of interaction between the region, the country, the world, and so on: a low cost per ton of product (e.g., quarried stone) would usually have a more local context than a high cost per ton of product (e.g., semiconductors). Relative or absolute decline in transmission, transport, and/or communication costs may change the relevant boundaries, as can political decisions on the control or decontrol of trade.

Simmonds (1969a, b) has published papers containing carefully researched, empirical studies of scale effects in the Canadian and US chemical industries. He uses as a key measurement the ratio of the "largest single-train plant" (i.e., the largest that depends on one major component) to the total market or production of a country (whichever is the larger). His evidence is that "the size of the largest plant has usually kept pace with the growth of the market." He subsequently (1972, 1975) described this as an "industrial behavior pattern" and suggested the use of such behavior patterns as a basis for classification of industries, particularly for forecasting.

In his second paper (1969b), he uses an examination of relative scale and scale economies to consider the comparative competitive position of Canadian and US firms in the Canadian market, and shows that "across-the-board percentage tariff reductions are ineffectual for industrial nations with relatively small domestic markets such as Canada, in major products such as petrochemical intermediates." Simmonds also points out the various scales of definition of *market*, which indicates some of the problems of measurement and specification arising in the definition of *relevant environment*.

It is also of interest to note the use (without definition) of the term *world-scale plants*: "The cornerstone of our investment planning is to establish world-scale plants" (Hodgson 1978). This quotation is from the chairman's address to the 1978 annual general meeting of the large UK based chemical corporation, Imperial Chemical Industries Ltd. An ICI colleague, in amplifying this point, stressed that the company sought to be "lowest cost producers in every market in which we operate." The pursuit of scale economies in production, and the geographical expansion of relevant contexts by improvements in transport and communication have been mutually reinforcing processes.

1.2.4 Factors of Scale

There are many factors that can affect the choice of scale for an entity, and they can be grouped in different ways, depending on the situation and the research goal. For example, the factors might be political, social, economic, technological, organizational, managerial, financial.

Each group of factors could be further subdivided. As an example of a political factor one might cite security of supply or the desire to create "the largest (smallest, longest, and so on) entity in the world," for prestige to impress customers, or to attract good staff.

Social factors such as the problem of unemployment in a certain town or region could be of crucial practical importance in determining the scale of a business enterprise; one-company towns are vulnerable to technological change.

The political and social factors in many cases require the creation of entities on a scale far from optimal on economic grounds. The role of political and social factors becomes crucial only in the solution of practical problems of scale in a specific location or region, and they therefore tend to be ignored in general or theoretical studies of optimal scale; but their existence must be recognized.

All the other factors are potentially significant in the general determination of optimal scale. Different factors influence the scale of an organization or of its units, in opposite directions: some of them favor an increase of scale, some a decrease. A general feature to be observed is that factors favoring the increase of scale are mainly *internal* to the technology and management of the firm, while the one that favor the decrease of scale are mainly *external* and concern the firm's relationship with its environment.

An analysis of the balance between internal and external factors may be conducted initially in static terms to determine the "optimal" scale. However, such an approach ignores the dynamic behavior of the factors, which will differ from factor to factor. For example, over the period 1950–1970 the internal factors, the technological, managerial, and organizational developments resulting from scientific progress, were relatively dynamic while the external factors, political and social relationships, changed less quickly. Combined with fairly steady economic growth, this meant that scale tended to increase. Subsequently, the external factors have shown more change and variability than the internal factors, resulting in a feeling in some industries that scale should decrease. The dynamic effects of different factors are difficult to determine because their impact is often indirect and hard to observe. There is further discussion of the effect of dynamic learning on scale in Chapter 6.

1.3 APPROACHES TO THE STUDY OF SCALE PROBLEMS

1.3.1 Contexts other than Production Systems

Problems of scale have been studied from the perspectives of many different disciplines and in other contexts besides production systems. Before discussing approaches to scale problems in production systems it is worth reviewing the approaches and concepts developed in two other contexts – biology and human settlements.

1.3.1.1 Scale in Biology

One of the concerns of biologists has been to identify underlying similarities of structure and behavior among widely diverse biological entities. Of fundamental importance is the work of D'Arcy Thompson (1917), now conveniently edited in Bonner's abridged version. Bonner himself in *Morphogenesis* (1952) gave a succinct statement of a general model of the process of development in biological organisms. Although Bonner restricted his general model to biology, we find it remarkably applicable to the growth of scale in technological systems.

Development is separated by Bonner into two broad categories: the "constructive" processes and the "limiting" processes. Constructive processes tend to increase the size or weight (the growth process), change the form (morphogenetic movement), or change the differences of parts in an organism over time (differentiation). Limiting processes check, guide, and channel the constructive processes. They can be external, such as food supply limits, or internal, such as mechanical stresses on the bones. For example, the strength of the leg of a mammalian quadruped varies as the cross-sectional area (L^2) while the weight varies as the volume (L^3). Thus the larger the animal the greater the weight and pressure, so the area of the leg must become disproportionately large to support the animal.

Although Bonner's terminology and case material are exclusively biological, one can trace a close parallelism with technological development. His basic model of development, morphogenesis and differentiation, correspond to two of Gold's points in Chapter 2. He criticizes the confusion between size and scale, pointing out that size increased by mere addition and accumulation (i.e., Bonner's growth) but an increase of scale properly implies a redesign of the form of the plant (i.e., Bonner's morphogenesis). Gold (1974) has previously emphasized that "scale economies are derived from increasing specialization of functions" and hence suggests that

> . . . scale be defined as the level of planned production capacity which has determined the extent to which specialization has been applied in the sub-division of the component tasks and facilities of a unified operation.

There is no doubt that biological models have a pervasive relevance to human activity, ranging from artistic inspiration to engineering technicalities. Arising out of his experience in biology and his perception of the underlying similarities of structure and behavior among widely diverse biological entities, von Bertalanffy (1951, 1968), developed general systems theory and was preeminent in demonstrating similarity of structure between biological and other systems, including social organizations. While this theory does not appear to have won the widespread acceptance or familiarity to which its claims of universality might have entitled it, his work contains a number of examples of biological processes that have relevance to a study of scale in industrial plants and organizations.

For example, he shows that the allometric equation that relates the size Q_1 of an organ to the total size Q_2 of an animal as the animal grows, $Q_1 = bQ_2{}^a$, where a and b are constants, also applies to some aspects of social organizations, such as the relationship between the number of staff and the total number of employees in a manufacturing company as the company grows (von Bertalanffy 1968, pp. 64, 103).

1.3.1.2 Human Settlements

In a general discussion of problems of scale, some mention should be made of patterns of human settlement. Much of the history of civilization is related to the problems of successively larger communities, both in terms of local settlements and at the national and supranational level. Single areas of settlement now range over six orders of magnitude in their population: from isolated houses to cities of several million people.

The scale of towns and cities is principally of significance to other decisions about scale in that the former often define the environment within which the latter are made. This is equally true of national environments, for activities to which this is the relevant measure; and there have been many scale-related arguments for the creation of international activities and supranational entities.

The diversity of circumstances, and the obvious fact that there are specific satisfactions and drawbacks associated with every size, show at once that there is little point in seeking any simple solution to problems of "optimal size" or "optimal mix." The professional urban planners do not themselves appear to have developed clear views on either desirable target patterns, or standard and satisfactory methods, for land use and urban planning. Some of the regional strategic plans in the UK have drawn heavy criticism of their unimaginative and oversimplified techniques. In a critical and pessimistic article, Schneider (1977) remarks that

> . . . planners operate without a conception of an ideal city. Especially in the United States, there is no established norm for size, either with upper or lower limits. There is no economic ideal, no formula for urban productive or consumptive efficiency.

A concept that might be carried over from political science to production systems is that of functionalism, associated particularly with Mitrany (1943, 1965). Functionalists seek the creation of single-purpose institutions specific to each function, whether street lighting, transport, or strategic defense. The appropriate organization size is in each case to be determined by the intrinsic needs of the function and not according to any prior commitment to a particular scale, least of all to such multifunctional composites as nation-states or federations. The logic of Dahl and Tufte's enquiry (1974), *Size and Democracy*, leads them towards a similar conclusion, but they reject the proliferation of changing organizations implied, favoring some simplicity and stability of form. In the same context the passionate work of Kohr (1957, 1977) in defense of smaller nations should be mentioned. But the weight of the argument seems to be with the functionalists rather than the nationalists: as Buchan (1969) quoting Kitzinger (1968) expressed it,

> . . . The national framework is "too big now for some purposes that need smaller units; too small for other functions that need a supranational scale."

The organizational response of many multinational production companies appears to be a compromise between the logic of functional and product–group organization, and the national political realities of the societal environment under whose laws they must operate.

A feature dominating any normative or prescriptive approach to planning the scale or pattern of human settlement is the extent to which it is dominated by the existing pattern. The rate of significant possible change is normally so slow that major change can be achieved only over many decades. To forecast and plan for many decades ahead demands heroic assumptions about the uncertainties, or reflects a scarcely justifiable attempt to create certainties to which future events must adjust themselves.

1.3.2 Disciplinary Approaches to the Study of Production Systems

Problems of scale in actual production systems do not arise under disciplinary labels like the papers in an examination. However, they have been studied from the perspectives of many different disciplines, and the approaches developed can in some cases be conveniently grouped under these headings. The categories below are somewhat arbitrarily divided because subjects overlap and the organization of subjects varies from country to country. Broadly speaking, the authors reviewed relate and refer more to other authors within their category than to those outside it; they tend to use common terms, concepts, and assumptions.

1.3.2.1 "Industry-Specific" Approaches

Industry-specific approaches embody the view that the problem of scale is so technical and industry-specific that it must always be tackled entirely on an *ad hoc* basis in terms of a specific project, and that no useful generalizations can be made from the project, or brought to the project except from earlier, similar projects in the same industry.

The justification for such an approach is clear to the extent that an industry's product or technology is unique in some important respect: electricity cannot be stored, newspapers and ice-cream are perishable. Primary industries, at the interface with natural resources, tend to have unique characteristics, and use of final products by the consumer may also be unique to the product, even where industrially similar: there are differences between a cooker and a refrigerator, or even between a dishwasher and a washing machine. But between primary extraction or harvesting and final use, the industries of intermediate processing and conversion, and transportation and distribution, there tends to be conformity to general patterns, though product storage characteristics lead to some differentiation. Service and information-processing industries may require separate consideration, but their growth of scale and patterns of deployment show similarities to manufacturing industries in terms of organization, if not of technological content.

There are no industries that do not have some characteristics in common with other industries; but insofar as they differ, an enhanced appreciation of the industry's specific problems is likely to be the result of an understanding of the respects in which it is unique. A good example of the benefits of industry-specific research is Gold's study (1974) of economies of scale in Japanese blast furnaces. Gold was able to distinguish between conclusions specific to that industry and conclusions evidently of wider applicability.

1.3.2.2 Engineering Generalization

Engineering literature tends to be industry-specific, but there have been some cross-industry generalizations, such as the concept of a "power-law" (cost = constant × capacity^k; e.g., $k = 0.7$), or the concept of the Reynolds number in hydrodynamics.

In the chemical engineering literature on cost estimation there are extensive tables that give estimates of the parameter k for a wide variety of process equipment and plants. These estimates are based on empirical studies and, though there are physical and chemical laws that provide theoretical justification in some cases, there is usually no information given to assess the quality of the data base or the range of plant sizes on which the estimates of k are based. Thus, if there are significant changes in relative cost factors, such as the ratio of labor to materials, or if projections of cost outside the range of prior experience with size are made, the resulting cost estimates may have substantial errors.

An important aspect of problems of scale, to which the engineering studies pay greater attention, is that of their "multisectional" and "multifunctional" nature. One is not simply comparing black boxes of varying size. This has several implications for the methodology used in analyzing decisions about scale. Technically, the scaling-up rules may be quite different for different parts or functions – a point that follows from the considerations of morphogenesis discussed in section 3.1 above. In planning and construction, one might not necessarily build a perfectly "balanced" plant. Some parts (e.g., site utilities) might have a low marginal cost of extra capacity at the time of constuction, but be impossibly expensive to expand in future years. Sometimes reliability requires redundancy in the number or size of cheap but critical components. The information required to make decisions on plant configuration and size has to be sought at the engineering level; unfortunately, many published economic models and financial decision criteria tend to ignore these refinements.

1.3.2.3 Technological Development

A further level of generalization of engineering approaches can be found in models of technological development (Sahal 1978). These combine the techniques of dimensional analysis and models of growth in order to describe the evolution over time of the relationships between the key features of engineering systems. In particular, Sahal considers the evolution of the maximum size of a unit over time and shows that in a number of cases the data fit the simple models $Y_t = at^b$ or $Y_t = cX_t{}^d$, where t is the time from introduction of the technology, Y_t is the size of the largest unit at time t, X_t is the cumulative production up to t, and $a,b,c,$ and d are constants for the particular technology.

The validity of these generalizations was the subject of much discussion during the workshop (see Chapter 6). If they are valid, they would provide a basis for projecting future developments of the technology.

1.3.2.4 Industrial Economics and Econometrics

Industrial economists have long sought generalized models of input–output relationships in different industries, summarized by "production functions." Such investigations have often been technically deficient for several reasons. For generality a large sample is sought. This may lump together plants of different construction date, design, and other significant factors. The analysis is often at the level of the organization (level 3), rather than at the level of the unit or plant (levels 1 or 2), because there is more published information on economic performance.

A criticism of much of the industrial economics literature (*cf.*, Chapter 2 of this volume and also Gold (1975)) is that it uses static economic models and assumes U-shaped cost curves. Gold (1975) remarks

> Continuing reliance on convenient assumptions in place of exploring the realities of industrial practice has rendered [these] traditional approaches to scale economics widely inapplicable in concept and all but trivial in their posited effects.

While the econometric literature is based on a more comprehensive and consistent theoretical model of the firm and uses more general production functions, cost functions, or profit functions, the choice of the mathematical form is generally dictated by the need to fit the various parameters using standard regression techniques rather than by a consideration of the technical aspects of plant design and operation. Cowing and Smith (1978) review a wide variety of econometric studies of steam-electric generation and conclude,

> . . . the econometric literature on measuring the character of the production technology for steam generation of electricity has increased in methodological sophistication.

> This research has advanced the econometric modelling of production technologies generally. Major advances have been made in a number of fundamental issues. However, these accomplishments merely expand the scope for methodological and empirical contributions. The research to date would suggest that these further developments must be based on a more explicit recognition of the actual physical and behavioral constraints in assembling resources in the generation of electricity.

It is of interest to note that the consensus of the studies is that there are significant scale economies in small and medium units but the scale effects may disappear with larger units.

An econometric study having some relevance to the paper by Betts (Chapter 5) is that by Lau and Tamura (1972) on Japanese ethylene plants. They found that the scale coefficient for capital was approximately 0.6 and the scale coefficient for labor was not significantly different from zero; that is, labor requirements are independent of plant size. Energy and raw material scale coefficients were approximately one, that is, energy and raw materials are used in fixed proportion to output.

The reaction of the engineer to such studies is likely to be "I could have told you so based on my knowledge of plant design and operation." There is an obvious need for the approaches of the engineer and the econometrician to be brought closer together. In particular, the engineer's knowledge of the technology should be combined with the econometrician's concern with the firm's economic goals and constriants.

1.3.2.5 Engineering–Economic Systems

The most comprehensive attempts to combine engineering and economic approaches are the analytical models on the choice of the optimal size of an engineering plant developed by Manne (1961) and subsequently extended by him and a number of other authors (see Manne 1967; Erlenkotter 1973, 1977; Erlenkotter and Rogers 1977). These models have provided considerable insight into the way in which the optimal size and optimal mix of plants is affected by market growth, desired return on investment, scale characteristics of the technology, and the costs of alternative processes or methods of meeting the demand.

Apart from Manne's application of his models to development planning, similar models have been used in practice in order to identify at the preliminary stages of investment planning those plant sizes and configurations that should be examined more closely using detailed engineering studies (Cameron 1974, Ball and Pearson 1976, Smith 1979).

1.3.2.6 Social Science and Organizational Scale

A considerable literature exists on organizations, their sizes, and various structural characteristics. The literature is here only briefly reviewed. Much of it is apparently descriptive, seeking general models and relationships (independent of the particular function or industry of the organization). The literature does not appear to be oriented towards application to specific decisions, although there is no reason why incisive descriptive studies should not be so used if the descriptions include any measures of efficiency or affectiveness.

For example, one of the classic works is Alfred Chandler's epic study (1962) of the growth of America's major corporations. Chandler demonstrated the causal connections between certain types of industry and phases of their development (e.g., the railroads, Dupont Chemicals, General Motors), and the organizational forms adopted. Ansoff (1965) drew extensively on Chandler in his prescriptive work on corporate strategy. One might cite Chandler as providing case studies in "organizational morphogenesis."

Another researcher in the sociology of organization whose work advanced to a prescriptive stage was Joan Woodward (1965). Her work was particularly significant (and widely influential), because it appeared to display a systematic relationship between the technology of an industry and its optimal organizational form, with the implication that firms departing from this optimum would have poorer performance. This strong hypothesis has not been well supported by subsequent research, and Donaldson (1976) claims that its results have been "disconfirmed." This critique was eagerly taken up and emplified by Eilon (1977). Attempts have been made to defend the original Woodward thesis at least at levels near the work flow: e.g., our level 2 rather than level 3 of the total organization. At level 3, size appears to be the main determinant of an organization's structural characteristics; at levels 1 and 2, technology may be determining.

A recent paper in this field is Dewar and Hage (1978): "Size, Technology, Complexity and Structural Differentiation: Toward a Theoretical Synthesis." Each of the four terms in the title is carefully defined in terms that are measurable; structural differentiation is considered both vertically (hierarchical levels) and horizontally (determinants). Technology is defined as "task scope." The terms are then measured for each of 16 social service organizations in 1964, 1967, and 1970, thus giving data not only on the measures but on their rates of change. Correlation and regression analysis are then applied, to try to determine associations and causal connections, and their relative strengths. For example, "Large organizations are and remain complex ones as are organizations with a variety of tasks. But are they both becoming large and adding more inputs at the same time? Which is the stronger causal process?" They found no effect of size on complexity, but suggest that perhaps "the amount of growth was not sufficient to generate the economies of scale necessary before additional administrative specialties could be hired." This type of interpretation is similar in concept to Gold's definition of scale as a function of degree of specialization (see section 3.1 above). Size, rather than technology, is found to be the more important determinant of both vertical and horizontal differentiation.

The literature to which the above is a brief introduction is obviously important to any general study of problems of scale in organizations. It is empirical, quantitative, and seeks generality. But it may be problematical to apply conclusions from public service organizations to situations in manufacturing industries; and the dearth of studies including comparative performance measures is a serious deficiency.

To incorporate questions of scale at the level of the technological unit into these models, one would need a means of translating scale alternatives into their alternative organizational implications; it is not yet evident that any rigorous way has been found of doing this, or even whether any such unique relationships need exist.

1.3.2.7 Control Theory

The literature of control theory contains many contributions from electrical and electronic engineers, and from mathematicians and cyberneticians. It concentrates on technical situations amenable to analytical modeling, computer simulation, and technical experimentation. Its application appears to be very local in origin – the control of automatic machinery or a process plant. But increasingly there have been attempts to extend the scope of the formally structured control systems to larger systems, such as an integrated steelworks complex; and at lest on a theoretical level, the methodology has been applied to larger scale problems such as economic management.

The subject has not been extensively reviewed within the *IIASA* research, but is here noted for the sake of completeness, and with an awareness that it has much to contribute to the methodology of formal control in certain types, and on a certain scale, of organization. The April 1978 special issue of the *IEEE* journal, *Transactions on Automatic Control* is devoted to large-scale systems and decentralized control. In his editorial reviewing the issue, Athans makes a number of significant general observations:

> The inefficient operation of large-scale interconnected physical systems can be attributed to lack of fundamental understanding and modeling of the underlying interactions, the lack of coordinated control strategies in an inherently dynamic and stochastic system.

Athans goes on to point out the difficulty of developing new theoretical tools for both decentralized and centralized control of large-scale systems. He suggests that

> What we need from a theoretical point of view are novel and innovative approaches for comparing alternate decentralized information and decision structures. The current state of the theory does not allow us to do this. The new theories will have to bring in new concepts of solutions, new definitions of what we mean by optimality, with special emphasis on reliable operations, and a more fundamental understanding of the value of information for decision making. In short, we need brand new theories for the future, and this is why the field of large-scale system theory and decentralized control will continue to be an exciting area for both theoretical and applied research in the decades to come.

1.4 DESIGN OF THE BOOK

The book consists of a number of the papers presented at the workshop together with chapters that summarize what emerged in the discussion as the key issues relevant to scale.

We begin with the paper by Bela Gold in which he calls for a revision of the prevailing approaches to evaluating scale economies and the development of a systematic methodology for making scale decisions.

The next three parts of the book group the papers and discussion in accordance with the level of the scale problem on which they focus. In Part 2 "Technology and Scale," we consider scale problems and issues at levels 1 and 2, the unit and the plant. The specific applications in the three papers included are international air transportation, electricity generating units, and ethylene plants. In the discussison of the papers that were concerned with plant and unit size, two issues were of particular concern – the validity of the learning curve for describing the process of size increase and cost reduction (Chaper 6) and the way in which uncertainty about the future modifies scale decisions (Chapter 7).

In Part 3, "Organizations and Scale," we consider problems and issues at level 3, the level of the organization. The presentations included discuss the way in which managerial load depends on the size of the productive activity (Chapter 8), the determination of the appropriate size of organizational units within a large firm (Chapter 9), the importance of formal management systems in contributing to firm efficiency (Chapter 10), the appropriate structure of multiplant, multiproduct organizations (Chapter 11), the application of a formal systems approach to evaluate organizational structure (Chapter 12), and the relationship between the size of management, the size of the organization, and the environment of the organization (Chapter 13).

The major issues arising in the discussion were the relationship between organization size, structure, and the environment (Chapter 14) and the relationship between the size of the organization and the effectiveness of the innovation process (Chapter 15).

In Part 4, "Scale and National Industry Policies," we consider problems at levels 4 and 5, the level of the industry and the nation. We begin with a description of the insights that an understanding of scale effects provides on the debate on the merits of protectionism versus free trade (Chapter 16). Next we include two papers, one from the West and one from the East, which give differing perspectives on national economic policies for small countries. Daly's paper (Chapter 17) considers the options available to Canada, protectionism and self-sufficiency, participation in a common market, or specalization and low tariffs. The paper by Haustein and Wittich (Chapter 18) is concerned with how the GDR should restructure its industry so that it can participate effectively in the CMEA and in world trade in general.

In Part 5, "Concluding Remarks," Chapter 19 describes a number of possible directions for future research on scale based on suggestions of workshop participants and concerns raised in the discussion. Finally, in Chapter 20 we return to the two questions posed at the beginning of this chapter and try to summarize our answers to them based on our review of the proceedings of the workshop.

REFERENCES

Ansoff, H.I. 1965. Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion. New York: McGraw-Hill.

Athans, M. 1978. On Large-Scale Systems and Decentralized Control. IEEE Transactions on Automatic Control. AC-23(2): 105–107.

Ball, D.F., and A.W. Pearson 1976. The Future Size of Process Plant. Long Range Planning. August.

Bolotryi, K.A., and L.I. Itin 1976. Planirovanie optinalnogo razmera predpriyatiya (Planning of optimal size of enterprise). Moscow: Machinostroenie.

Bonner, J.T. 1952. Morphogenesis: An Essay on Development. Princeton University Press. Also 1963 New York: Athenaeum.

Buchan, A., ed. 1969. Europe's Futures, Europe's Choices: Models of Western Europe in the 1970s. London: Chatto and Windus.

Cameron, D. 1974. Three Simple Steps to Determine Optimum Plant Capacity. Long Range Planning. February.

Chandler, A.D. 1962. Strategy and Structure. New York: Doubleday.

Cowing, T.G., and V.K. Smith 1978. The Estimation of a Production Technology: A Survey of Econometric Analysis of Steam-Electric Generation. Land Economics 54(2): 156–186.

Dahl, R.A., and R. Tufte 1974. Size and Democracy. Stanford, California: Stanford University Press.

Delombre, J., and B. Bruzelius 1977. Importance of Relative Market Share in Strategic Planning – A Case Study. Long Range Planning 10. August.

Dewar, R., and J. Hage 1978. Size, Technology, Complexity and Structural Differentiation: Toward a Theoretical Synthesis. Administrative Science Quarterly 23(1). March.

Donaldson, L. 1976. Woodward, Technology, Organizational Structure and Performance – A Critique of the Universal Generalization. The Journal of Management Studies. October.

Eilon, S. 1977. Structural Determination. Omega 5(5).

Erlenkotter, D. 1973. Sequencing Expansion Projects. Operations Research 21(2): 542–553.

Erlenkotter, D. 1977. Capacity Expansion with Imports and Inventories. Management Science 23(7): 694–702.

Erlenkotter, D., and J.S. Rogers 1977. Sequencing Competitive Expansion Projects. Operations Research 25(6): 937–951.

Glagolev, V.N. 1973. Problemy povyshenija effektivnosti mechanizaciji i automaticaiji v promyshlennosti. Tesisy dokladov vsesojuzhio nauchuoi konfezenciji. (Problems of raising mechanization and automation efficiency in industry.) Proceedings of All-Union scientific conference. Moscow: USSR Academy of Sciences publishers.

Glagolev, V.N. 1975. Mechanizacija i automatizacija truda v promyshlennosti Litovskoi SSR (Labor mechanization and automation in the Lithuanian SSR's industry). Vilnius: Mintis.

Gold, B. 1955. Foundations of Productivity Analysis. Pittsburgh, Pa.: University of Pittsburgh.

Gold, B. 1974. Evaluating Scale Economies: The Case of Japanese Blast Furnaces. Journal of Industrial Economics. September.

Gold, B., ed. 1975. Technological Change: Economics, Management and Environment. Oxford: Pergamon.

Hedley, B. 1976. A Fundamental Approach to Strategy Development. Long Range Planning 9(6) December.

Hedley, B. 1977. Strategy and the "Business Portfolio." Long Range Planning 10(1): 9–15 (February).

Hodgson, M. 1978. Chairman's Address to Fifty-First Annual General Meeting of Imperial Chemical Industries Limited. ICI Ltd., London. 19 April. p. 10.

Kitzinger, U. 1963. Britain's Crisis of Identity. Journal of Common Market Studies. June.

Kohr, L. 1957. The Breakdown of Nations. London: Kegan Paul. Also 1975 Swansea U.K.: Christopher Davies.

Kohr, L. 1977. The Overdeveloped Nations: The Diseconomies of Scale. New York: Schocken Books.

Kosygin, A.N. 1976. Guidelines of the Development of the National Economy of the U.S.S.R. for 1976–80: Moscow: Novosti Press Agency Pub. House.

Lau, L.J., and S. Tamura 1972. Economies of Scale, Technical Progress, and the Nonhomothetic Leontief Production Function: An Application to the Japanese Petrochemical Processing Industry. Journal of Political Economy 80(6): 1168–1187.

Manne, A.S. 1961. Capacity Expansion and Probabilistic Growth. Econometrica 19(4): 632–649.

Manne, A.A., ed. 1967. Investments for Capacity Expansion: Size, Location and Time Phasing. Cambridge, Massachusetts: MIT Press.

Mitrany, D. 1943. A Working Peace System. Republished with additions: 1966. New York: Quadrangle Books. (See also: 1965. The Prospect of Integration: Federal or Functional. Journal of Common Market Studies. December.)

Sahal, D. 1978. Law-Like Aspects of Technological Development. dp/78–85. Berlin: International Institute of Management – Wissenschaftszentrum.

Schneider, K.R. 1977. Planning Without Vision – Roots of the Urban Disaster. Long Range Planning 10 (June).

Simmonds, W.H.C. 1969a. Stepwise Expansion and Profitability. Chemistry in Canada. September: pp. 16–18.

Simmonds, W.H.C. 1969b. The Canada–U.S. Scale Problem. Chemistry in Canada. October: pp. 41.

Simmonds, W.H.C. 1972. The Analysis of Industrial Behavior and Its Use in Forecasting. Technological Forecasting and Social Change 3: 205–224.

Simmonds, W.H.C. 1975. Industrial Behavior Patterns: A new dimension for planners. Futures 7(4): 284–292.

Smith, R.L. 1979. Turnpike Results for Single Location Capacity Expansion. Management Science 25(5): 474–484.

Thompson, D.W. 1917. On Growth and Form. Abridged Edition: J.T. Bonner, ed. 1969. London: Cambridge University Press.

Von Bertalanffy, L. 1951. Problems of General System Theory. Human Biology 23: 302–312.
Von Bertalanffy, L. 1968. General System Theory, Foundations, Development, Applications. New York: G. Braziller.
Woodward, J. 1965. Industrial Organization. London: Oxford University Press.

CHAPTER 2 REVISING PREVAILING APPROACHES TO EVALUATING SCALE ECONOMIES IN INDUSTRY

Bela Gold
Case Western Reserve University,
Cleveland, Ohio

Increasing competitive pressures in domestic and international markets have stimulated efforts in many industries to gain what are widely believed to be the advantages of "scale economies" through the building of progressively larger operating units. Such tendencies are apparent in broad sectors of manufacturing – including chemicals, steel, pulp and paper, and cement – as well as in power generation, mining, shipping, and agriculture. This reflects the spread of faith in the benefits of scale increases beyond engineers and industrial managements to governments, which then foster larger operations in the hope of strengthening the competitive position of their industries.

There can be no doubt that scale increases have indeed yielded substantial benefits in many cases. But there is no basis for claiming benefits for all such increases. Indeed, it is important for industrial and public officials to recognize that such generalized expectations rest on cloudy and even dubious foundations. A strong case can accordingly be made for exercising prudence in basing further investments and supporting policies on such beliefs, pending more thorough exploration of the sources and effects of successive increases in scale, and of how these may differ among firms and industries, depending on such factors as their technologies, product mixes, and market conditions.

Accordingly, the first objective of the following analysis will be to review some of the conceptual limitations of the theoretical and empirical literature that has been most widely used as the basis for policy decisions and for further research relating to scale. The second objective will be to present some suggestions for strengthening such efforts on the basis of studies in a variety of industries in several countries during the past decade by our Research Program in Industrial Economics.

Before proceeding, however, attention should be called to three restrictions shaping the following discussion. First, the primary focus is on trying to understand how and why scale is increased by firms rather than on developing speculative models and hypotheses which are too general to permit effective applications to the variety of concrete cases faced by industrial managers and government policy makers. It is hoped, of course, that progressively more general insights will emerge from such explorations of reality. Second, this pragmatic emphasis on seeking to encompass the wide array of potential benefits, burdens, and limitations likely to be considered in decisions to change scale leads to a willingness to develop and utilize highly complex analytical frameworks rather than an insistence on simple models. This is in accordance with the actual rather than the mythical course of scientific progress exemplified by the replacement of the ancient Greeks' beautifully simple conception that all changes in the physical world result from interactions only among earth, air, fire, and water by the extraordinary complexities of modern physics, chemistry, and biology. And thirdly, the following discussion will center around the plant and industry levels of production activities, emphasizing their technology, productivity, and cost, among others, while minimizing concern with related organizational, psychological, and sociological aspects of changes in scale.

2.1 REVISING THE CONCEPT OF SCALE

2.1.1 Shortcomings of Traditional Concepts

One of the odd features of the extensive literature on scale is the fuzziness of the basic concept. Most of the engineering literature regards increases in scale as synonymous with increases in size or capacity. Most of the economics literature concurs in this view, and adds two additional requirements – that the "technology" and the proportioning of input factors remain unchanged – to ensure that increases in scale refer only to enlarged duplicates of smaller units or operations. However, neither of these concepts identifies the potential sources of scale economies or the means by which they might be realized, thus providing no effective focus either for practical decision making or for designing research to develop improved guides for making such decisions.

Such conceptual shortcomings may be illustrated in practical terms by asking what are the significant differences between tripling the production of a given plant by adding two units identical to the original and adding one unit designed to yield twice the capacity and also conforms to the requirements of economic theory that factor proportions and technology remain unchanged. But what would be the source of any hoped-for economies of production? Clearly, all physical input–output ratios would be identical to the original, and would thus yield no gains. Boulding (1948) has suggested that there might be a reduction in salaried staff because, for example, only one president or general manager of production would be needed. This is a highly questionable claim, however, in view of widespread experience in which salaried staff had to be increased more rapidly than associated major increases in capacity in order to cope with the increasingly complex problems of production planning, coordination, and control, as well as related adjustments in materials handling, maintenance, intermediate inventories, and organizational relationships.

A more common suggested source of scale economies in the economics literature involves capital goods, which are obtainable only in a relatively few, widely differing sizes. Hence, it is argued that plants that are too small to utilize fully one such unit, or too large to employ only one and too small to utilize two such units, would be less economical than plants whose needs are met by fully utilizing one such unit. But this "lumpiness of capital goods" explanation seems vulnerable on several counts. First, it offers no explanation for significant economies from progressive increases in scale, inasmuch as plants fully using two or more units of capital goods would not be more economical than those fully utilizing only one. Second, most capital goods are in fact available in a wide range of sizes, and plants can be readily designed to handle a given size; they need not be deliberately designed to bear significant diseconomies. Third, there can be no doubt that larger units of *some* capital goods are more economical than smaller units, but their use would entail a violation of the theoretical requirement of unchanged factor proportions. Moreover, such expected economies cannot serve as the foundation for a general theory unless they can be shown to be widely representative, and until persuasive answers can be provided to the original questions of why and under what conditions larger units or operations are likely to be more economical than smaller ones.

Thus, in our illustrative case, adding two units identical to the original would clearly triple capacity and meet the requirements of unchanged technology and input factor proportions, but it is difficult to identify any resulting sources of scale economies. On the other hand, providing the same increase in capacity by adding a facility designed to yield double the output of the original unit could yield several sources of economies derived from exploiting the benefits of the higher levels of specialization made feasible by the planned larger capacity. This would permit the use of more highly specialized machinery, instrumentation, and automated controls, more extensive materials handling facilities, a greater subdivision of labor tasks, a shift in the composition of skills, and so on.

In short, the preceding helps to illustrate two of the basic weaknesses of the economic theory of scale attributable to its reliance on abstract logic alone. Its requirements of fixed input factor proportions is seldom encountered in cases involving substantial increases to scale precisely because such restrictions tend to minimize or prevent the benefits whose expected realization is a primary motive for considering scale increases. For example, after presenting these standard requirements, Samuelson (1973, p. 28) offers illustrations that do not seem to conform to the requirements of fixed input factor proportions. The requirement that technology be unchanged is likewise contrary to most experience, because extensions of scale beyond past frontiers require an exploration for and an addition to, technological knowledge, as well as modifications of past practices. This may involve altering materials specifications, equipment characteristics, input proportions, operating rates and conditions, labor tasks, maintenance requirements, and so on. In most cases, various kinds of additional technological knowledge are acquired in the course of a determination of how further scale increases would affect the applicability and effectiveness of past operating relationships. Thus, scale increases may utilize 80–95 percent of existing technological knowledge and relationships, but seldom 100 percent.

Our analysis accordingly suggests a definition of scale "as the level of planned production capacity which has determined the extent to which specialization has been applied to the subdivision of the component tasks and facilities of a unified operation." (For a fuller discussion, see Gold (1955, pp. 115–117; 1979, pp. 119–124 *et seq.*).) This definition offers an important basis for differentiating "scale" from "size," for it recognizes that increases in size involving a comparable proliferation of activities, or involving mere duplication of smaller scale relationships, may yield no economies at all. It also offers a practical means for measuring differences in scale and identifies the means whereby management may seek to maximize the potential benefits of increases in scale. Finally, it offers some guidelines for research and analyses of the potential effects of scale changes in different industries and plants – as will be discussed later.

2.1.2 The Irrelevance of "Learning Curves"

Before concluding the discussion of the scale concept, it may be worth noting that it has somehow become confused by some analysts in recent years with the quite unrelated concept of "the progress function" or "the learning curve." The latter purports to explain progressive improvements in the performance of specified operations solely as a result of cumulative experience rather than increases in capacity per unit of time. This approach has commonly been applied by analyzing trends in performance over time through the application of statistical methods on the basis of hindsight.

But the results represent purely descriptive findings devoid of persuasive bases for explaining any observed gains. Such improvements in one or another measure of performance may be due to a host of quite different factors, including changes in product designs, product mix, operating technology, facilities and equipment, management planning and controls, materials quality, labor skills, and labor incentives. But failure to identify and weigh the contributions of these various factors would deny any basis for appraising the past roles and future potentials of each as a guide to continuing managerial efforts to improve results. And ignoring the disparate capabilities of such factors would certainly undermine the projection of aggregate past trends into the future. Although it is conceivable, of course, that such improvements in aggregate performance might also be due in some measure to changes in scale, these would seldom be the primary cause. There would seem to be no difficulty in recognizing their role in such instances, nor any reason to avoid differentiating such distinctive contributions. In the overwhelming proportion of cases, therefore, improvements reflected by the progress function or the learning curve are likely to have little or no relationship to changes in scale.

Nevertheless, in view of the frequency with which such associations have been posited in the economics and management literature, it may not be amiss to call attention to the vulnerability both of the basic concepts involved and of the evidence that has been presented in their support. It seems reasonable, of course, that the repetitive performance of unchanging operations would yield some improvements as a result of increasing familiarity during its early stages. But it is less plausible to expect continuing significant gains from endless repetition alone. Industrial psychologists suggest that, contrary to the old saw "practice makes perfect," repetition merely makes existing behaviors habitual. Hence, if performance improves beyond the relatively early stages, it is important to determine the factors underlying such gains, i.e., who did the learning – R&D staff, production engineers, managers, labor or equipment suppliers – and what did they learn. Such analytical insights are necessary to guide managers both to promote more or faster learning, so they do not simply wait and hope for it to emerge somehow, and to identify the sources failing to provide the contributions of which they are considered capable.

Much of the empirical evidence offered in support of the learning curve also seems based on samples too restricted to warrant the generalizations to which they have given rise. For example, virtually all such findings covering periods longer than a few years are based on statistical trends that have been superimposed on patently fluctuating or intermittent rates of performance improvement, often even ignoring long periods of nonprogress. At the very least, this would suggest that cumulative experience alone does not account for the performance improvement. Hence, there would seem to be no basis for interpreting the findings as measures of any specific influences other than "unexplained change." The results for any firm or industry may be attributable to *external* changes in the availability, quality, and prices of materials, labor, and capital goods inputs; in the level and composition of market demand; and in the price and product quality pressures generated by competitors – none of which is included in the common connotation of "learning." Moreover, most of the *internal* changes generated by design engineers, production specialists, and supervisory staff represent the results not of cumulative repetition of past practices but of active explorations of alternatives to such practices. This may also be termed "learning," but it obviously need not be correlated (except perhaps inversely) with the duration of past practices. Incidentally, while the Boston Consulting Group has demonstrated such improvements in the early developmental stages of technologically vigorous industries – especially those involving recent developments in electronics – without however identifying the relative contributions to such gains of the various quite different factors that may have been involved, it is all too easy to demonstrate the absence of such "progress" for long periods in a wide array of other major industries.

2.2 SOME ANALYTICAL LIMITATIONS OF PAST THEORY AND EMPIRICAL FINDINGS

2.2.1 Economic Theory

Established economic theory offers virtually no significant contributions to an understanding of the sources of past or prospective scale economies. It depicts scale effects – i.e., the effects on minimum average total unit costs of increases in the capacity of plants engaged in identical production activities – in the form of a U-shaped "long-run cost function." This expectation is based on four assumptions, as shown in Figure 2.1:

1. The short-term cost functions, which show the effect on average total unit costs of variations in the capacity utilization of individual plants, are U-shaped
2. The minimum cost points of such short-term cost functions tend to decline for successively larger plants up to some optimal point beyond which minimum costs begin to rise
3. Plants producing identical products by means of identical inputs and technologies commonly cover a wide range of sizes at any given time
4. The overlapping of adjacent short-term cost functions indicates that the diseconomies of underutilization of large plants are greater than those of fuller utilization of smaller plants producing the same output as the large plants

Such elementary economic concepts have been widely diffused among engineers, businessmen, and government officials, and may well have made them very receptive to proposals for continuing increases in scale. Unfortunately, analysis offers little support for any of these assumptions.

The vulnerability of the first assumption on theoretical as well as empirical grounds has been dealt with at length elsewhere. Analysis of its underlying assumptions and an examination of actual cost behavior in a large sample of industries suggested three conclusions: variations in capacity utilization rates may be accompanied by a wide range of

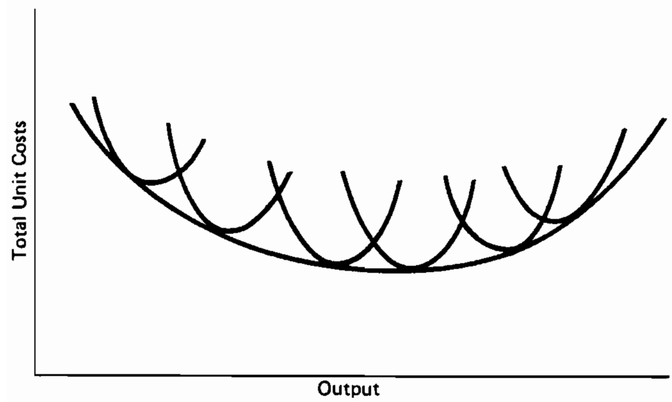


FIGURE 2.1 Theoretical effect of increasing scale on total unit costs.

adjustment patterns in total unit costs, even within the narrow purview of static economic theory; the U-shaped cost function need not be the most common among these; its likelihood tends to decline rapidly as the restrictive assumptions of static analysis are relaxed. For further discussion, see Gold (1966).

Analysis suggests that the second assumption is rooted in three deeper assumptions. One is that production facilities are available in only a few widely differing sizes – which has already been challenged. A second is that larger capital goods yield additional economies, presumably by requiring less investment per unit of capacity – another claim lacking both theoretical and broad empirical support. The third, which is usually only implicit, is that the expected investment benefits of increases in scale would not be offset by accompanying increases in any other costs. But this, too, warrants more serious investigation.

It has already been suggested that the effective operation of larger-scale units tends to require disproportionate increases in salaried personnel and costs in order to cope with the rapidly multiplying complexities of integrating an increasing array of more highly specialized tasks and equipment. Nor is it uncommon for wage rates for comparable skills to be higher in large-scale plants than in small plants, whether because of stronger trade union pressures or other influences. Moreover, consideration must also be given to the relative magnitudes of prospective reductions in unit investment charges and possible increases in wage and salary costs. For example, if total fixed capital charges approximate 10 percent, as is the case for depreciation plus interest in the reputedly "capital intensive" steel industry (for cost proportions in the US steel industry, see Gold (1976)), investment cost per unit of capacity would have to be reduced by *half* by means of an increase in scale in order to reduce total unit costs by only 5 percent, even if all other unit production costs remained unchanged. In short, it would require fairly heroic assumptions about the magnitude of scale economies to support expectations of even reasonably modest reductions in total unit costs on the basis of the theoretical analysis usually presented.

The third assumption on which the scale function, or long-run average cost curve, is based is also highly vulnerable. Logic alone suggests that the greater the economies and subsequent diseconomies of successive increases in scale, the narrower the range of plant sizes that can survive under conditions of effective competition. With respect to the theoretical possibilities, engineering analyses indicate that, within the confines of a given technology and product mix, most industrial processes tend to be characterized by relatively narrow zones within which input–output relationships are most effective – in contrast to the broad "production possibilities frontiers" assumed in the economic theory of production (for further discussion, see Gold (1978)). While these optima are not always identified at the outset, it would seem to follow that every increase in scale yielding substantial improvements in performance would tend to shift the loci of competitive plant sizes upward instead of broadening the range within which scale economies are too small to affect long-term survival. Nor has empirical evidence supported this third assumption on any broad basis. Of course, a variety of studies have reported a wide range of plant sizes within specified industry categories as defined by statistical agencies (for example, see Stigler's application (1958) of his "survivor test" to assess scale economies). But further analysis reveals very substantial heterogeneities within most such categories with respect to the specific products made, technologies employed, factor proportions utilized, and markets served, thereby voiding the claims of relevance to an assessment of the significance of scale economies. Smaller plants serving small isolated markets, for example, need not be under any competitive pressure to increase scale in order to reduce production costs to levels achieved in large markets elsewhere.

Finally, the vulnerability of these three assumptions also undermines both the acceptability and practical relevance of the final assumption. Specifically, it has been suggested that the U-shaped short-term total unit cost function is but one of a variety of theoretically possible and empirically demonstrated patterns, that the shape of such functions may change with increases in scale, and that the range of sizes of plants engaged in identical production activities is likely to be quite narrow in most industries. Accordingly, although underutilization of a given plant's capacity often tends to involve cost penalties, it need not follow that smaller plants can economically provide comparable products (for example, automobiles) for the same markets.

2.2.2 Other Misleading Expectations

Managerial receptivity to proposals for increasing scale has been encouraged not only by economic theory, but also by the long-standing acceptance in the supposedly hard-headed engineering literature of the "six-tenths rule." This holds that each doubling of capacity tends to require increases in investment of only about six-tenths (for illustrative references, proposing exponents of 0.6 and 0.7 for estimating the investment cost of larger equipment and larger plants, respectively, in the chemical industry, see Aries and Newton (1955, pp. 6–7, 15), Bauman (1964, pp. 39, 180), Crowe *et al.* (1971, p. 110), and Peters (1958, p. 93)). Further inquiry reveals, however, that this expectation seems to be rooted solely in the simple-minded view that volume increases more rapidly than the enclosing surface of rectangular, cylindrical, and spherical shapes, and that the output of facilities tends to be correlated with their volume, and investment costs tend to be associated with the size of the enclosing surface. Such a relationship may hold, of course, for some facilities, especially in the construction of hollow shells, such as tanks, furnaces, boilers, pipes, and small buildings. But fundamental shortcomings restrict the range of its applicability.

Specifically, there seem to be few physical or chemical processes for which such fixed relationships hold without boundaries. Even construction activities face growing difficulties under conditions of increasing size, strain, and deteriorative forces. In addition, large facilities often make disproportionate demands on energy inputs, functioning components, instrumentation, and so on. At the very least, these suggest that the exponent indicating the rate of increase of construction costs with gains in capacity for a broad array of processes and industries probably covers a wide range beyond the supposedly universal value of six-tenths. And another implication is that the exponential value for any process is likely to change over successively higher ranges of scale. For an excellent review of the potential sources of scale economies, including experiential as well as theoretical views, see Pratten (1971, pp. 3–19). For earlier reviews of empirical studies of scale that provide penetrating discussions of conceptual problems and the difficulties confronting efforts to undertake effective empirical testing of theoretical expectations as well as illustrations of the wide range of published findings, see Smith (1955), Moore (1959), Walters (1960), and Shuman and Alpert (1960). Empirical estimates of exponents covering a wide array of processes are provided in Aries and Newton (1955, p. 7) and Bauman (1964, p. 146).

In short, reliance on convenient assumptions in place of an exploration of the realities of industrial practice renders the traditional approach of economic theory to scale economies widely inapplicable in concept and trivial in its posited effects. Indeed, it would not be unfair to summarize the latter as suggesting only that there is some optimal size of production unit for any given technology and product mix without specifying what it might be, or what determines it, or the magnitude of its presumed cost advantages compared with progressively larger and smaller sizes.

2.2.3 Some Shortcomings of Empirical Research

Turning from theory to empirical findings, a review of this literature, too, is disappointing. Weiss's review (1971, p. 297) of the published research on scale led him to conclude that it is a "still fairly blank field." The most important reason for the pervasive inadequacies is the tendency to concentrate analyses on the average statistical relationships between plant size categories and some measure of cost or "productivity." In most cases, these involve comparisons for arrays of individual industries, completely ignoring the extensive heterogeneities, cited in section 2.2.1, within such statistical groupings. Hence, resulting differences among plant size groups (usually based on product value or employment), whether cross-sectionally or over time, can seldom be ascribed convincingly to true scale effects alone. Somewhat more useful insights have been provided by some studies of scale effects that concentrated on single industries, though these too vary with the actual heterogeneity of the operations encompassed. But even in the case of coal-fired electric power plants and blast furnaces, with seemingly maximum homogeneity of products and of basic technologies, analyses reveal a substantial list of differentiating factors other than scale that affect performance measures. For an extensive summary of the literature on scale effects in electric power generation followed by a detailed analysis of the factors affecting apparent scale differences, see Huettner (1974, Chapters 2 and 3). Another review of the scale literature followed by a detailed study of the factors associated with scale differences in US blast furnaces is provided by Boylan (1975).

Even more serious is the apparent lack of interest of most researchers in trying to identify the specific forms, major causes, and various effects on performance of changes in scale – they often seem concerned only with the average relationship between two variables that just happen to be ostensible measures of scale and of some associated aspect of performance. It is not surprising, therefore, that the question of the relationship between changes in sheer size and in scale is not even raised, thus ignoring any attendant deviation from the theoretical requirements of fixed factor proportions and unchanged technology. Perhaps this also explains the widespread failure of such studies to explore the limitations of the input, cost, and output data used – including their insensitivity to the changes in the qualities and composition of the inputs used and outputs produced, which are critical aspects of scale changes and attendant benefits.

Most reported results are accordingly descriptive, rather than analytical. They provide no persuasive explanations of the causes of whatever average intergroup differences are reported. Nor do they facilitate efforts to disentangle distinctive scale effects from the interacting effects of such nonscale factors as differences in capacity utilization rates, trade union pressures, changes in factor and product prices, and other innovational developments. Two outstanding exceptions to such largely unanalytical reports were presented to this conference by John Fisher (Chapter 4) and G.G. Betts (Chapter 5).

With respect to another source of empirical perspectives, it is worth noting that when engineering and construction firms are asked for advice on gaining economies through increases in scale, their recommendations most commonly propose only modest increases beyond available experience rather than providing authoritative evaluations of successively greater increases in scale. One reason is that the firms recognize that many unexpected difficulties may lie beyond established scale frontiers, and, since they are not research organizations, they are unwilling to assume such responsibilities unless attendant risks are clearly shifted to the customer. A closely related reason is that few convincing research-based insights are available from any quarter for estimating the effects of major changes in the scale of most industrial processes. Our own research has suggested how extremely limited and vulnerable are the judgments about scale effects even of the central engineering staffs of large companies that have already built a number of plants.

It seems reasonable to conclude, therefore, that in view of the major potentials apparently offered by scale increases in some industrial sectors, and of the costly disappointments in others, there is obviously a great need to undertake serious research on the prospective benefits and limitations of increases in scale in various industries, under differing conditions and at several levels of aggregation. Only through the careful exploration of such varying sectors is needed progress likely to be achieved in understanding the generalizable and nongeneralizable aspects of scale effects.

2.3 SOME EMPIRICAL PERSPECTIVES ON IMPROVING APPRAISALS OF SCALE EFFECTS

In order to develop a fuller understanding of the nature and specific sources of past scale effects, and of the potentials of further changes in scale in particular industries and under varying conditions in product and factor markets, it is necessary to undertake a more systematic and deeper analysis than has been done in most studies of the technological, economic, and managerial factors whose interactions determine such outcomes. In designing such a comprehensive analytical framework, serious consideration should be given to supplementing the past heavy reliance on deductive logic with a number of operational insights suggested by our past empirical research.

2.3.1 On the Influence of Managerial Objectives and Expectations

Sharper perspectives may be provided, for example, by starting with a commitment to learn – instead of continuing to make assumptions about – the primary objectives motivating managerial decisions involving choices among alternative scales and alternative means of implementing them.

In multiplant firms, the desired characteristics of an additional plant are rarely completely independent of other existing facilities. On the contrary, they tend to be defined by the limitations of existing facilities in light of expected competitive pressures and market adjustments. Scale is likely to be only one dimension of the needs to be satisfied. Others might involve changes in the varieties and qualities of products, alterations in the types and qualities of materials and energy inputs, the replacement of old facilities, increasing reliance on labor-saving equipment, and even the shifting of production closer to major markets or sources of purchased supplies. Hence, different new plants may have been intended to serve dissimilar purposes, whether their scales are similar or not; this may lead to distorted evaluations of their relative operating performances if it is assumed that they have identical objectives and that any differences in results are attributable largely or wholly to differences in scale.

Moreover, in evaluating the effects of scale changes, it is necessary to recognize that plant designs can be biased in one direction or another, even within a specified capacity level, so as to accord with managerial preferences. For example, lower capital investment may be traded for higher operating costs, as in the general case of curtailing the sequence of processing operations by purchasing more highly fabricated components, or in the special case of building "peak power" plants rather than more continuously utilized "base load" power plants. Another example involves trading lower capital investment for a shorter plant life, either because an older management seeks to maximize operating performance during its own limited remaining tenure, or because substantial technological advances are considered imminent. Another widely applicable option involves trading higher capital investment for lower labor costs, or greater safety, or better pollution control. And in industries capable of using alternative material or energy inputs, larger capital investment may be used to safeguard such flexibility against the risk of sharp changes in the relative supplies or prices of such substitutes.

Realism also requires recognition of other limitations of the traditional static economic theory. Managerial decisions involving significant changes in scale involve commitments, and hence must be based on evaluations often covering 10 years or more. Within that dynamic setting, managements tend to be less interested in reducing average unit production costs than in increasing profit rates, thus emphasizing the importance of managerial expectations in choosing among scale options. Increases in scale beyond the levels currently estimated to yield minimum unit costs might be encouraged by expected trends in the demand for and prices of the firm's products that promise high profits (for a major example, see Gold (1974)). Greater increases in scale than are currently optimal might also be regarded as justified if the new plants are substantially more economical in using inputs whose prices are expected to increase and whose supply is expected to be limited, or if they permit more effective and less costly pollution controls, or if they permit greater increases in needed capacity within restricted plant sites, or if they offer significant reductions in the time required for construction. Instead, decisions might favor scale increases smaller than those promising minimum unit production costs if prospective competitive pressures and trends in market demand and supply make full utilization of larger scale units unlikely for extended periods. And the practical range of scale options tends to be broader in industries where substantial changes in scale yield only modest adjustments in minimum average unit production costs, and still broader in industries in which production costs account for only moderate proportions of the total cost of sales. Moreover, in industries subject to substantial cyclical fluctuations in output, managements may be less interested in the scale level offering the lowest minimum cost point than in the level yielding the lowest weighted average unit cost over entire business cycles.

2.3.2 On the Role of "Factor Dominance" in Industrial Technologies

In appraising the sources and effects of changes in scale, attention must also be directed to the heavy influence of differences in the technology of industries, especially to the relative roles of various input factors in the productive capacity of plants. What I term "factor dominance" is quite different in concept from the economists' "factor intensity." The latter is concerned with the proportions of cost accounted for by each input category. For instance, the steel industry is commonly referred to as "capital intensive" because economists seem to think that its costs must be dominated by capital charges in view of the enormous magnitude of investments. As a matter of fact, such investments yield such enormous production capabilities that depreciation plus interest payments tend to average less than 10 percent of total costs in major US steel companies, as was noted earlier. On the other hand, the steel industry is undoubtedly "capital-dominated" because the level of mill capacity is almost completely determined by the facilities and equipment that embody the operating technology and fixed investment, and because such production goods also determine the potential contributions of labor and materials.

"Capital-dominated" operations include electric power plants, petroleum refineries, cement mills, pulp and paper plants, steel mills, and a variety of chemical plants. In such establishments, an increase in the speed with which labor performs its tasks merely reduces the man-hours required to service operations, instead of significantly raising productive capacity. Hence, in such industries, increases in scale tend to be achieved primarily by increasing the size and degree of specialization of individual facilities and equipment units, by more effectively integrating successive operations, and by progressively reducing the roles of human production efforts, evaluations, and controls relative to those of machines, instruments, and computers. The resulting benefits generally include increases in the productivity of fixed capital (i.e., in the ratio of capacity to net fixed investment), comparable or even larger gains in output per man-hour, and lesser or no reductions in unit materials requirements. The realization of such benefits depends, of course, on the extent to which the increases in productive capacity are utilized.

"Labor-dominated" production operations, in which output capabilities are determined by labor – such as custom tailoring and shoemaking, manual bricklaying, and various handicrafts – have declined sharply in importance in advanced economies. Even in less developed countries, the potential benefits of increasing the scale of such activities through progressively greater subdivision of tasks and associated specialization of functions tend to diminish rapidly. Efforts to increase the economies of labor-dominated operations tend rather to involve reducing the degree of labor-dominance by increasing the role of capital goods (often by technological advances as well) and by fuller standardization of products. Some examples include the increasing mechanization of shoe and clothing production on the basis of standard sizes and the increasing use of factory-fabricated wall sections in construction. Although output per man-hour may be raised by increases in the scale of labor-dominated processes, at least up to some relatively modest level, such gains are likely to be overshadowed by those from increases in the role of capital goods, from which diminishing returns may not be faced even after manifold increases in scale.

"Materials-dominated" processes are those in which the controlling constraint on output is the richness of the natural resources being utilized or processed. Some examples include primitive agriculture, mining, and fishing, where returns depend very heavily on the fertility of the soil, the richness of the mineral deposit, and the population density of the fishing grounds. Increasing the scale of purely materials-dominated extractive operations is a misleading concept because natural resources are a passive element in the production process. Accordingly, a change in scale of such operations necessarily involves a further subdivision of tasks and the functional specialization of the human and capital inputs. The initial stages of increases in the scale of farming or mining, for example, have typically involved subdivision of labor, and further increases in scale have resulted from the replacement of labor with machinery; later, there was resort to larger and more highly specialized equipment and facilities. Similar means are also employed in increasing the scale of fishing (where there was resort to fish-locating instruments, specially designed boats, and improved refrigeration) lumbering, and ore smelting. Thus, larger-scale extractive operations involve "joint factor dominance" rather than remaining materials-dominated processes. Such combined contributions, however, are unlikely to prevent differentiation of the contributions to scale effects of changes in each input factor. In smelting, for example, there is no difficulty in separating the effects on metal output of increases in the metal content of the ores processed in unchanged smelting facilities, from the effects of processing unchanged ores in larger-scale smelters.

The preceding brief discussion suggests the following hypotheses:

- Scale can be increased, in concept, only by increasing the contributions to capacity of the "dominant factor" in the production process – increases in the contributions of "subsidiary factors" permit only reductions in their respective inputs per unit of output.
- Although some increases in the scale of materials-dominated and labor-dominated processes can be achieved by increasing the subdivision and specialization of labor inputs, especially in the early stages of technological development, significant increases in the

scale of modern production processes tend to be attributable overwhelmingly to increases in the role and contributions of larger and more highly specialized capital goods.

- However, the benefits of scale increases involving the greater dominance of capital goods depend on the extent to which accompanying increases in capacity are utilized, partly because of the inflexibility of capital investments when there are output fluctuations, partly because of the tendency of increasing capital-dominance to decrease the ratio of direct personnel to indirect, less flexible, labor and clerical personnel (including maintenance people, instrument monitors, and bookkeepers), and partly because of the growing needs for salaried technicians and managers to ensure the effective planning, servicing, coordination, evaluation, and improvement of the resulting wider array of more specialized and interdependent operations.
- Because larger-scale modern industrial operations generally involve changes not only in the proportions and specifications of inputs, but also in the range and qualitative characteristics of the product mix, substantially smaller-scale operations can seldom compete with them by producing comparable products for the same markets, even when the larger-scale plants are operating well below capacity.

2.3.3 On the Cost Effects of Increases in Scale

It should be emphasized, however, that none of the foregoing patterns of changes in the "network of productivity relationships" – which I have defined as encompassing the six linkages representing the proportions in which materials, labor, and capital inputs are combined with one another, and also the input requirements of each per unit of output – need have a consistent relationship with changes in "physical efficiency" (for fuller explanations, see Gold (1955, pp. 59–67)). Serious analysis readily demonstrates that physical efficiency is an essentially meaningless concept, reflecting a false analogy from engineering.

Specifically, one can calculate the physical efficiency of an engine *with respect to energy conversion* by comparing the energy content of its fuel consumption with the energy equivalent of the useful power delivered by it. But it is not possible to measure the "physical efficiency" of the engine as a whole, nor of the process of producing it. Such determinations are prevented by the absence of any *important physical* common denominators for combining the input contributions of labor energy and skills, many kinds of materials and supplies, even greater varieties of facilities and equipment, and a wide array of technical and managerial activities; the available measures of numbers of people, volume, and weight are obviously irrelevant (more detailed discussion is provided in Gold (1979, pp. 45 *et seq.*)). Lacking any significant concept or measure of "physical efficiency," and being in any event much less concerned with the physical aspects of performance alone than in their economic effects, managers trying to evaluate the effects of increases in scale, or other sources of past or prospective improvements in operations, are more concerned with resulting changes in costs.

Systematic exploration of the effects of scale changes on total unit costs requires an analysis of three intermediate linkages: their direct effects on the unit input requirements of each factor as well as the proportions of one factor to another, such as were discussed above; resulting interactions with their respective factor prices; and the proportions of total cost accounted for by each of the major input categories. It is important to realize that wide variations may be found within each of these linkages depending on: the means by which scale is increased; attendant changes in the characteristics and mix of inputs and outputs; and the responsiveness of factor markets to the changes in input requirements generated by the increase in scale. An illustration of some of the more common possibilities may point out the analytical complexities that may be encountered in an evaluation of scale effects.

For example, the preceding discussion suggests that, far from providing economies all across the board, many increases in the scale of industrial production processes may yield the following mixed array of changes in input requirements per unit of output at relatively high levels of capacity utilization, compared with smaller scale operations: reduced direct labor, increased indirect labor, unchanged or lower purchased materials (if waste and reject rates are reduced, or if salable by-products are increased), increased energy (to drive the larger contributions by capital goods), increased salaried employment, and reduced investment.

It is traditional in economic theory to assume that factor prices are unaffected by changes in unit input requirements, but this is untenable even under static conditions. One reason is that significant changes in unit input requirements commonly involve changes in the qualitative characteristics of such inputs (for further discussion, see Gold (1978)). Direct labor is commonly saved by scale increases by readjustments in tasks that entail changes in the skill composition of man-hour inputs. The level of inputs of purchased materials may be reduced by shifting from some kinds of materials to others and by altering qualitative specifications to fit more exactly the planned adjustments in product characteristics and output mix. Changes in energy inputs may include shifts among major types and quality grades of fuels as well as alterations in the ratio of purchased to internally-generated power and heat. Significant adjustments may also be expected in the proportions of clerical, technical, and higher-level personnel.

Changes in unit input requirements may also affect factor prices because of reactions in the factor markets to resulting changes in demand. For example, increases in output per man-hour (as labor tends to interpret reductions in man-hour requirements per unit of output) often produce partially or wholly offsetting increases in wages per man-hour, because of trade union demands, or piece rates, or wage incentives, especially if scale increases are introduced adjacent to smaller-scale operations, or within firms having other plants engaged in similar lines of production. When scale increases are introduced by several competitors, resulting combined effects on input factor demands are also likely to affect their prices, encouraging reductions, for example, if total demand for common materials declines significantly, or encouraging increases if the scale changes increase demands for hitherto less available materials. Even salary levels may rise if demand for various kinds of specialists strains available supplies.

The interacting effects of relative changes in unit input requirements and in their respective factor prices obviously determine resulting adjustments in their relative unit costs. Thus, a 10-percent increase in output per man-hour accompanied by only a 5-percent increase in average hourly wage rates would yield a 5-percent reduction in unit wage costs. But how would this affect total unit costs? Clearly this depends on the wage proportion of total costs. If this were 20 percent, as has been the rough average in All Manufacturing Industries in the United States over an extended period, the reduction in unit wage cost would tend to reduce total unit cost by only one percent (for illustrative industrial data, see Gold (1975, pp. 6–12)). Whether even this slight saving were achieved, however, would obviously depend on the direction and magnitude of concomitant changes in the other unit cost categories weighted by their own shares of total unit costs.

Thus, in the illustration presented earlier, the net result of the assumed decreases in direct labor and investment requirements might or might not offset the assumed increases in indirect labor, energy, and salary inputs. In this connection, three points warrant emphasis. First, cost proportions differ very widely among industries – hence, even similar adjustments in unit cost categories might yield quite different changes in total unit costs. Second, changes in the various unit cost categories cannot be evaluated independently of one another in appraising scale effects, because the very means by which scale is increased produces interconnected effects on the component inputs whose contributions must be integrated in the production process. Third, such interacting effects of scale determine not only the estimated level of the minimum cost point of a prospective increase in scale, but also the shape of adjustments in its total unit costs with variations in capacity utilization.

In short, changes in scale may have a wide range of effects not only on the level and shape of the "long-run average cost curve," which defines theoretical expectations of the effects of increases in scale on minimum attainable unit production costs, but also on its successively deeper foundations:

- The level and shape of the theoretical "short-run cost function" for various levels of scale, from which the long-run cost function is derived, as shown in Figure 2.1
- The level and shape of the theoretical physical input–physical output functions on which the short-run cost functions are based (for a detailed critique of these functions, see Gold (1966))

To replace this multilevel structure of oversimplified concepts and "theories" with guides to managements (and governments) that are more solidly rooted in penetrating analyses of the sources and effects of actual changes in scale in various sectors of industry and under a variety of market conditions clearly requires extended research efforts.

2.4 ON THE INTERPRETATION OF EMPIRICAL FINDINGS

In the recent headlong rush to quantification as the basis for broad generalizations in economics, sight has all too often been lost of the relevance of the data and analytical techniques used to the concepts and hypotheses ostensibly being evaluated. Empirical research on the economic effects of scale is illustrative of such unacknowledged but far-reaching deviations from the posited foci of analysis.

To begin with, no effort is even made in most studies to demonstrate that the operations covered by the data presumably measuring differences in scale accord with the defined requirements of using identical technologies and factor proportions. As a result, the findings have no necessary connection whatever with the effects of differences in scale as conceived by the theories supposedly being tested. At best they can reveal only differences that may be associated with differences in relative employment or output levels, and there is no analytical basis for explaining their sources or causes. Unfortunately for the cavalier assumption that such measures are "reasonable" proxies for the scale concept, even the most cursory examination of the characteristics of the establishments encompassed by most statistical series for individual industries demonstrates a wide array of incompatibilities with such claims – including, as noted earlier, differences in product designs, diversity and proportions of product mix, operating technologies, vintages of capital facilities, input factor proportions, input qualities, factor prices, markets served and even capacity utilization rates. Indeed, some more recently constructed plants may even embody significant technological advances without being substantially larger in capacity.

Secondly, in the further interests of analytical simplicity, the effects of scale on performance are usually evaluated by applying inadequate and vulnerable single criteria uniformly to all groups. Even if the criterion used were estimated average total unit production costs, it would not conform to theoretical requirements, unless these could be determined for the minimum cost point of each scale level. But few even pretend to do so, relying instead on actual average costs per unit of output for annual data, which often encompass substantial variations in capacity utilization rates for individual establishments, as well as considerable differences in output fluctuations among establishments in an industry category. However, data limitations force most studies to settle for even less persuasive measures, such as annual averages of: value-added per unit of direct labor (wage earners or man-hours) or per unit of estimated physical output; and estimated physical output per unit of direct labor. One need hardly belabor how tenuous the relationship is between these "proxies" and the total unit production costs at optimum levels of capacity utilization, which the former are supposed to represent.

Third, it is important to remember that performance must be measured against intended objectives rather than against targets arbitrarily assumed by outsiders. After all, one cannot even evaluate the relative superiority of two runners going around a practice track without knowing the distances over which each is trying to maximize his performance. As was discussed at length earlier, managerial decisions to increase scale may reflect a variety of objectives, among which achieving the lowest level of unit production costs at optimal levels of capacity utilization may be common but far from universal. And even when this is a primary target, it is based on each management's particular expectations concerning prospective future changes in input factor supplies and prices, in the level and product composition of demand along with associated product price levels, in the pressures exerted by competitors, and in any relevant governmental regulations and policies. Accordingly, evaluating the effects of increases in scale through uniform application of any single criterion, which ignores differences in the objectives and expectations underlying the relevant managerial decisions, poses obvious problems to meaningful interpretation.

The preceding consideration also raises questions about when scale effects should be evaluated. As tends to be true of all other major innovations as well, the effects of scale adjustments within a plant tend to change over time (for a more detailed discussion, see Gold *et al.* (1979, Chapter 13)). During the early period of operations, performance is often limited by the need to overcome technological difficulties, to develop needed operating skills and supervisory controls, and to achieve increasing capacity utilization. Thereafter, results are frequently affected by external changes: in factor prices; in the level, product composition, and geographical dispersion of demand; and in marketing pressures from competitors – and by the progressive internal readjustment of various stages of operations to facilitate maximum realization of scale potentials. Over still longer periods, the effects of early increases in scale interact with the effects of comparable or even larger scale increases by competitors – as well as with possible changes in technologies and other developments influencing costs and prices – thereby tending to reduce the profit margins of the pioneers. At any given time, however, evaluations are likely to reflect a variety of vintages in each size (or scale) group – ranging from relatively recent to progressively older plants – thus reflecting different stages in their evolving patterns of cost and other performance effects.

Two basic analytical weaknesses are at the root of the foregoing problems of evaluating the actual effects of changes in the scale of production operations:

1. Reliance on the telescopic perspectives of far-off observers scanning too broad a field of phenomena
2. The substitution of hindsight evaluations of *ex post facto* findings for efforts to understand the actual processes of decision making leading to specific scale adjustment choices.

The first of these rests on the dubious assumption that the most important effects of scale increases are essentially similar across the entire industrial horizon. One can, of course, survey the topography of the earth from distances great enough to minimize the significance of mountain chains. But one would not do so if the purpose of such observations were to build roads. Studies of scale that are intended to evaluate the effects of specific prospective changes, or the desirability of altering current policies affecting such changes, must come closer to the examination of actual cases and, thereby, we can learn gradually the extent to which it may prove possible to develop subgroupings of scale changes that are similar enough to support some eventual generalizations about their effects.

Hindsight perspectives are often unsatisfactory for a variety of reasons, and may be increasingly misleading as the period between the original decision and the time of evaluation is lengthened. One reason is the difficulty of uncovering the objectives and expectations on the basis of which managements had to make scale commitments in the face of uncertainties about the future. Thus, actual outcomes may be evaluated on the basis of uninformed assumptions about the managerial objectives that determined the forms taken by the scale changes. Another reason is that unexpected developments after completion of a new facility tend to foster a succession of managerial responses that may affect important aspects of inputs, operations, and outputs, thus altering the effects attributable solely to the change in scale. Nor can the explanations and assessments offered afterwards always be accepted as accurate, especially when offered by newcomers, or by those associated with disappointing outcomes, or by those seeking credit for any demonstrable gains.

In view of the rudimentary state of our knowledge about the nature and effects of scale changes in various industries and differing market conditions, sound progress may well require a substantial array of case studies tracing the entire process of decision making, implementation, and successive evaluations in "real time" – beginning with the estimated effects of alternative scales presented to management, along with management's definition of objectives and related expectations; exploring the bases for the final choices made; and then appraising the nature of, and reasons for, all major changes in managerial policies, in addition to evaluating various aspects of performance periodically.

2.5 CONCLUDING OBSERVATIONS

1. Changes in the scale of production undoubtedly have important potentials in many industries, but few of them seem to have been explored systematically – including their economic and organizational as well as their technological effects. Hence, mythologies are still widely accepted – both those of many American economists claiming no significant scale benefits, and those of European economists as well as engineers in various parts of the world claiming widely available benefits of substantial proportions. Although I have examined a number of such evaluations prepared by company personnel to guide decisions by their managements, these, too, seemed to me to be dangerously vulnerable – especially in view of the enormous financial commitment that would have been involved in implementing them. Moreover, there are an increasing number of examples of large-scale industrial plants whose results appear to be disappointing to their managements. Hence, there are important practical needs to be met.

2. There are no general theories or empirical findings as yet that offer authoritative bases for estimating the scale potentials for industries at large, nor even for assessing the maximum scale potentials for any industry, firm, or plant. But intensive research should provide increasingly useful guides to the probable directions of scale increases in different factor-dominated technologies, along with the changes in factor and cost proportions likely to result under various market conditions. And it should also be possible to develop more systematic bases for estimating the prospective physical input–output and cost effects of specified forms of scale increases involving capacity expansions of 25–50 percent, under specific technological and market assumptions (though any particular set of such assumptions might prove wrong).

3. Finally, our field studies suggest that effective analysis of the effects of past and prospective changes in scale require a thorough grasp of the relevant technology, operating problems, and specific managerial pressures as well as of associated economic linkages, organizational repercussions, and impacts on human inputs. Expertise in economic theory or in statistical methodology alone is unlikely to yield the required range of knowledge.

REFERENCES

Aries, R.S., and R.D. Newton 1955. Chemical Engineering Cost Estimates. New York: McGraw-Hill.

Baumann, H.C. 1964. Fundamentals of Cost Engineering in the Chemical Industry. New York: Rheinhold.

Boulding, K.E. 1948. Economic Analysis. New York: Harper.

Boylan, M.G. 1975. Economics of Scale in the Steel Industry: The Case of U.S. Blast Furnaces. New York: Praeger.

Crowe, C.M., A.E. Hamielec, T.W. Hoffman, A.I. Johnson, P.T. Shannon, and D.R. Woods 1971. Chemical Plant Simulation. Englewood Cliffs, N.J.: Prentice-Hall.

Gold, B. 1955. Foundations of Productivity Analysis. Pittsburgh, Pa.: University of Pittsburgh.

Gold, B. 1966. New Perspectives on Cost Theory and Empirical Findings. Journal of Industrial Economics. April. (Reprinted in Gold 1971.)

Gold, B. 1971. Explorations in Managerial Economics: Productivity, Costs, Technology and Growth. London: Macmillan. Also (1971) New York: Basic Books. Japanese translation (1977) Tokyo: Chikura Shobo.

Gold, B. 1974. Evaluating Scale Economies: The Case of Japanese Blast Furnaces. Journal of Industrial Economics 23: 1–18. September. (Reprinted in Gold 1979.)

Gold, B., ed. 1975. Technological Change: Economics, Management and Environment. Oxford: Pergamon.

Gold, B. 1976. Tracing Gaps Between Expectations and Results of Technological Innovations: The Case of Iron and Steel. Journal of Industrial Economics. September. (Reprinted in Gold 1979.)

Gold, B. 1978. Interactions Between Technological Changes and Factor Prices. Revue d'Economie Industrielle. Spring. (Reprinted in Gold 1979.)

Gold, B. 1979. Productivity, Technology and Capital: Economic Analysis, Managerial Strategies and Government Policies. Lexington, Massachusetts: D.C. Heath – Lexington Books.

Gold, B., G. Rosegger, and M.G. Boylan 1979. Evaluating Technological Innovations: Methods, Expectations and Findings. Lexington, Massachusetts: D.C. Heath – Lexington Books.

Huettner, D.A. 1974. Plant Size, Technological Change and Investment Requirements. New York: Praeger.

Intriligator, M.D., ed. 1971. Frontiers of Quantitative Economics. Amsterdam: North-Holland.

Moore, F.T. 1959. Economies of Scale: Some Statistical Evidence. Quarterly Journal of Economics. May.

Peters, M.S. 1958. Plant Design and Economics for Chemical Engineers. New York: McGraw-Hill.

Pratten, C.F. 1971. Economies of Scale in Manufacturing Industry. London: Cambridge University Press.

Samuelson, P.A. 1973. Economics. 9th Edition, New York: McGraw-Hill.

Shuman, S., and S. Alpert 1960. Economies of Scale: Some Statistical Evidence – Comment. Quarterly Journal of Economics. August.

Smith, C. 1955. Survey of the Empirical Evidence of Economies of Scale. Business Concentration and Price Policy. Princeton, N.J.: Princeton University Press.

Stigler, G. 1958. The Economies of Scale. Journal of Law and Economics. October.

Walters, A 1960. Economies of Scale: Some Statistical Evidence – Comment. Quarterly Journal of Economics. February.

Weiss, L. 1971. Quantitative Studies of Industrial Organization. In (1971) M.D. Intriligator, ed., Frontiers of Industrial Economics. Amsterdam: North-Holland.

Part 2

TECHNOLOGY AND SCALE

OVERVIEW

The papers in this part are concerned with scale decisions at levels 1 and 2, the level of the equipment and the plant, in which the characteristics of the technology tend to be dominant.

Dathe's paper describes the historical development of aircraft size and the technical and other factors that appear to be limiting further size increases. This paper is representative of scale effects in transportation, not only in aircraft but also in ships, transport vehicles, and pipelines. While Dathe's paper contains no formal model of factors determining the size of the aircraft, there are some papers that attempt to develop formal models of size determination. For example, Kendall (1972) developed a model of the way in which ship size is determined by length of voyage, quantity and value of cargo, and the size-dependent capital and operating costs. This model was used for planning port facilities.

Fisher's paper is representative of a continuing interest in scale effects in electricity generation. He develops a model of the way in which the optimal size of generating units is determined by the relationship between the economies of scale in capital costs and the diseconomies of scale in unit reliability. By careful analysis of available data on unit cost and performance he was able to assess the extent to which utilities have installed units that were too large. By contrast, most of the literature that considers scale effects in electricity generation is concerned with the optimal mix of different types of generating units (*cf.* Anderson 1972). Another paper that was presented at the workshop, but is not included in this volume, and which was concerned with scale effects in electricity generation is by Spinrad (1979).

Bett's paper is a description of the factors determining ethylene plant scale. He gives a comprehensive and unique description of the factors that led to a rapid increase in scale over the period 1950 to 1975 and how, when some of these factors changed, the process of scaling up came to an end. His description is a useful basis for understanding the behavior of the industry and the way in which it responds to external influences (see Cantley 1979).

Finally, this part concludes with reviews of two of the main issues that arose in the discussion of these and other presentations relating to technology and scale:

1. Scale, technology, and the learning curve – the extent to which generalizations about the process of progressive size increase and cost reduction can be made

2. Coping with the uncertain future – the way in which uncertainty about future markets and the performance of technology affects scale decisions

Other papers presented at the workshop but not included in this volume and which were related to technology and scale were Derkx *et al.* (1978), Sahal (1978), and Gustavsson (1979).

REFERENCES

Anderson, D. 1972. Models for Determining Least-Cost Investments in Electricity Supply. Bell Journal of Economics and Management Science 3(1): 267–299.

Cantley, M.F. 1979. The Scale of Ethylene Plants: Background and Issues. WP-79-43. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Derkx, H.H.J.M., A. Kamerman, and A. van der Rijst 1978. How Experience and Attitude Affect Steel Plant Productivity. Iron and Steel International. October: 319–329.

Gustavsson, S.-O. 1979. Motive Forces for and Consequences of Different Plant Sizes. Goeteborg, Sweden: Chalmers University of Technology.

Kendall, P.M.H. 1972. A Theory of Optimum Ship Size. Journal of Transport Economics and Policy 6: 128–146.

Sahal, D. 1978. Law-like Aspects of Technological Development. Discussion paper dp/78-85. Berlin: International Institute of Management.

Spinrad, B. 1980. Scaling and Learning in Nuclear Energy. CP-80-17. Laxenburg, Austria: International Institute for Applied Systems Analysis.

CHAPTER 3 PROBLEMS OF SCALE IN INTERNATIONAL AIR TRANSPORTATION

Johannes M. Dathe
Industrieanlagen-Betriebsgesellschaft mbH,
Ottobrunn, FRG

Dass uns werde klein das Kleine
und das Grosse gross erscheine. . . .

(what is really small should become small for us
and the really big should appear as big)

Gerhard Tersteegen

3.1 INTRODUCTION

Today, problems of scale are an important issue for people who study the economic and societal causes and consequences of our industrial development. A supertanker that is sized to fulfill the requirements of economical transportation may become a "super offense" to coast regions in case of an accident. The centralization of the electrical power supply, which has definite advantages in balancing partial malfunctions in routine service, may become the cause of a major breakdown of large network areas in the case of a sudden cumulative overload. And the number of fatalities at the Chicago air accident last May would have been smaller than 276 if the DC-10 had not been a high-capacity aircraft with more than 300 seats.

After these experiences, nobody should be surprised that – in spite of pronounced gains in efficiency that are achievable only with large technical units – many people are reluctant to believe in the advantages offered by sheer technical magnitude. In short, skepticism of this kind is expressed by the slogan, "Small is beautiful."

Answers to questions about scale can only be given after careful analysis. The problems related to technical or organizational scale show big differences for various branches of industry. The case that is studied here is international air transportation carried out by airlines and by unscheduled carriers. Civil aviation is a field with a lot of prestige but it has been science oriented in most of its technical and economic decisions for several decades.

Some kind of a historical cornerstone (like the year 1973 for the energy market) can be dated at about 1970 for the field of air transportation. It is therefore useful to divide the analysis into two parts: (a) the rationale of the development until 1970, and (b) the status of the seventies and the prospects for the future as they are viewed today. Some of the dimensions of scale we will consider here are aircraft size (measured in gross weight, wing span, body width, payload, or number of passenger seats), aircraft cruising speed, and yearly transportation volume (measured in passenger-miles sold or in ton-miles of freight sold).

3.2 LESSONS FROM THE HISTORY UNTIL 1970

3.2.1 Aircraft Size and Economy

The history of scheduled air transportation as we know it today had its starting point in the twenties. The development of the size of civil aircraft during these 5 decades can best be characterized by a plot of gross weight of big aircraft as a function of the year of introduction of that type into service (Figure 3.1). It is fascinating to see that both the state of the art and considerations of economical air transportation induced a steady increase in the gross weight of aircraft which were and are representative of the technological state achieved in service. The first "giant aircraft" with 12 engines was the Dornier "Whale" DO X in 1928 with its 52 metric tons. It had the same outsider role as the Bristol "Brabazon" in 1949 or the project of an aircraft powered by a nuclear reactor in the sixties which never reached the development stage. These outsiders demonstrate that in the mainstream of aircraft history, weight increases of considerable magnitude took place but that the maximum size possible at that time was not introduced into widespread operation.

The last step in this weight history is the current Boeing 747 "Jumbo Jet" with a gross weight of about 350 metric tons. All the airline aircraft introduced before were of the so-called "narrow-body" type with a maximum fuselage diameter of about 3.4 m. The B 747 was the first member of the "wide-body" family with a diameter between 4.8 and 7.2 m and is now joined by the Airbus A 300 and the McDonnel Douglas DC-10. Wide-body aircraft are less comfortable for the passengers, but have important economic advantages for the airlines. Many analysts believe wide-body transport aircraft will have a great future.

Virtually all airlines that are operating over long distances and are members of the International Civil Aviation Organization (ICAO) of the United Nations decided to use aircraft of the 350-ton class. There were two major reasons for this decision. First, air traffic in congested areas has reached such a volume during the rush hours that the steadily increasing number of "flight motions" could no longer be handled without considerable delays. The step from 180 to 380 seats per aircraft helps to decrease the number of aircraft that arrive at and depart from each airport. Secondly, the cost per ton-kilometer or per seat-kilometer is decreasing with growing transport units. This is shown in Figure 3.2. where the direct operating cost (personnel, maintenance, fuel, landing, fees, insurance, capital interest and depreciation) for short- and long-range aircraft is given as a function of aircraft design payload; see Simon (1970). In certain intervals of these curves cost reductions of 20 percent or more are achieved only by an increase of the design payload. On the other hand, it can be seen from these curves that at the design points of

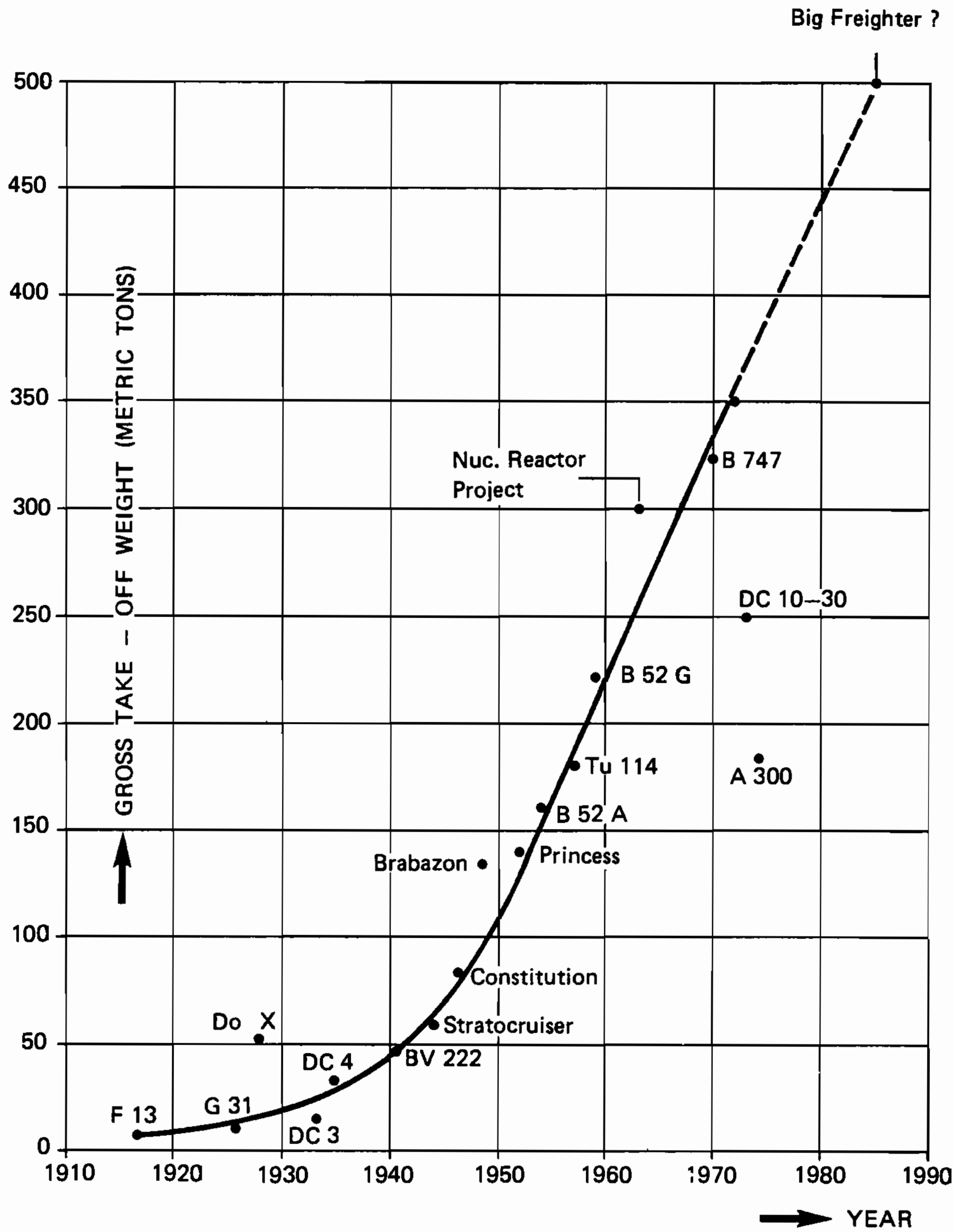


FIGURE 3.1. Weight history of big aircraft.

the Airbus A 300 (range 5,000 km, payload 35 tons) or of the Boeing B 747 (range 11,400 km, payload 70 tons) no marked improvement could be expected from a further increase in weight.

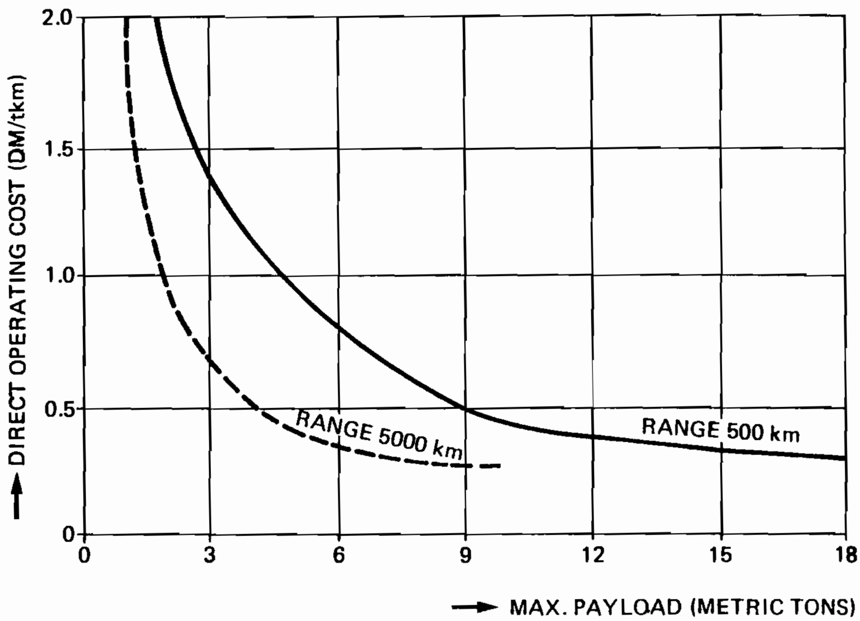


FIGURE 3.2 Direct operating cost as a function of the design payload of transport aircraft.

The increase of gross weight was not the only reason for the increase in the payload: there was considerable technological progress through the decades that resulted in a steady reduction of the structural weight of the aircraft and in a decrease in fuel consumption. In spite of the higher aircraft speed in 1970, the fuel needed to fly 100 km in a long-range aircraft was reduced from 4.5 kg per ton in 1950 to 4.0 kg in 1970; see Simon (1970). Only the jet engine introduced to civil aviation in 1955 allowed both growth in unit size and qualitative improvements. The associated increase in speed from 600 to between 900 and 1,000 km/h made the transatlantic flight shorter and more attractive, and the cruising altitude of 10,000 m and over made passengers more comfortable than ever before.

As already mentioned, the growth of aircraft unit size was a necessity because of the increases of air traffic volume. On the other hand, the higher efficiency of larger aircraft resulted in reductions in ticket prices which induced further growth of air transportation demand.

During the past 5 decades the yearly growth rate of transportation volume (expressed in passenger-miles sold) had many oscillations for various reasons; see Figure 3.3. These data are from the international statistics of the ICAO countries, but do not include important regions such as the Soviet Union and the People's Republic of China. The growth rates had a mean value of 29 percent during the thirties, 14.5 % in the fifties and sixties, followed by 9 % in the early seventies. Manufactures in the United States are still optimistic about a further growth rate around 7 % during the following decade.

Equally informative is Figure 3.4, in which yearly passenger-miles sold are plotted against time. There was an unbroken rise until 1968. Smaller increases around 1970 were

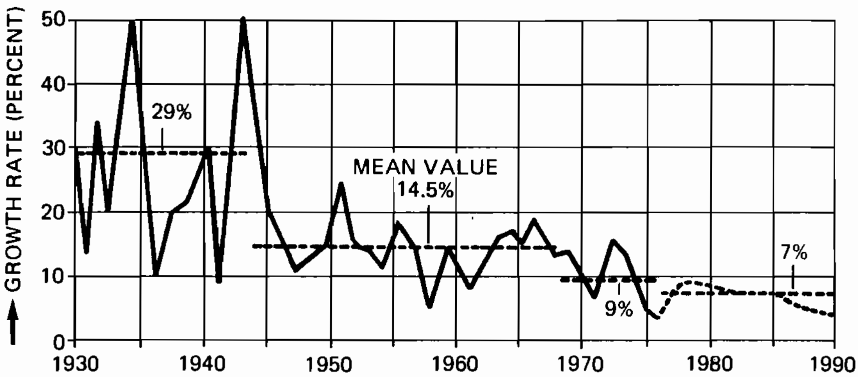


FIGURE 3.3 Growth rates of passenger miles sold in scheduled ICAO airline service.

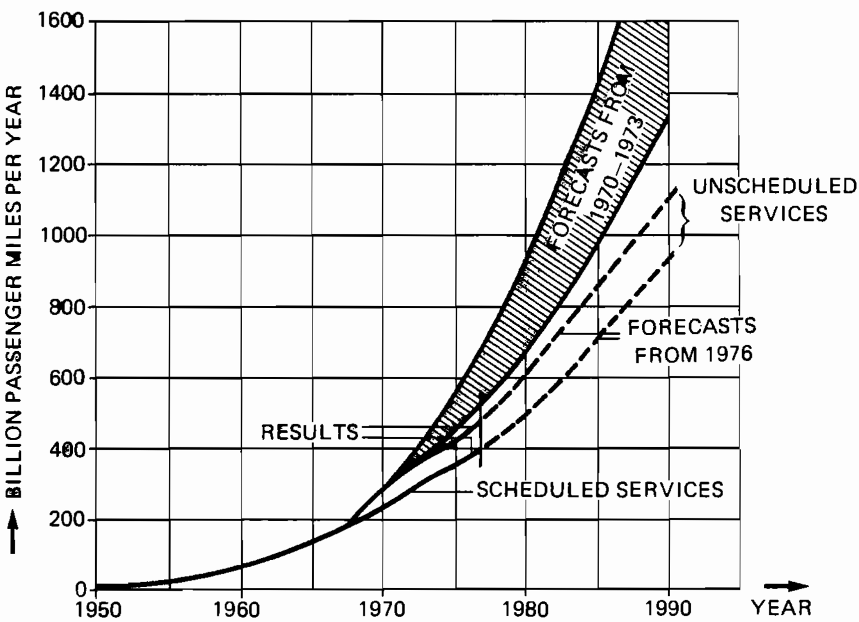


FIGURE 3.4 Growth of ICAO passenger services.

interpreted as temporary difficulties. The very optimistic trend extrapolations of 1970 were modified in 1973, the year of the oil crisis. But the development between 1973 and 1977 remained below these estimates. The latest forecast presented in this chart was worked out by the US manufacturer McDonnell Douglas; see Geddes (1977). These values are slightly reduced in comparison with the 1975 forecast by the same source. However, it should be understood that this forecast tries to make a kind of "self-fulfilling prophecy": This is one possible scenario of future development; another possibility would be that the total figure for 1990 remains below 1,000 billion passenger-miles.

Since the late sixties, the unscheduled passenger carriers are of growing importance as competitors of the airlines with their scheduled operations. There is no doubt that these carriers were inducing price reductions in tickets on the whole international market and are an important factor in paving the way to mass transportation by air.

3.2.2 Air safety

It is a question whether the development to larger air transportation density and to bigger unit sizes had disadvantages for air safety. According to *Flight* (1978), during the 2 decades in which the big changes took place, there was a reduction in the number of passenger fatalities per 100 million passenger-km; see Figure 3.5. Of course, there can be doubts whether the passenger-kilometer is the correct parameter for this comparison. But if one compares aircraft-kilometers or numbers of landings, there is a decrease in the magnitude in the ratio of 2:1 during the years under consideration. A careful analyst is not convinced when reference is made to the favorable casualty statistics of the last years. Maybe it was just good luck. In principle, at least, it is possible that there are some risks that might change the picture if the number of big aircraft grows further or when some modifications occur in the environmental or economic conditions that are of importance for air safety.

To study the relative safety level of big air transportation units, we should consider the following two model situations:

(a) air traffic in the form of *n* aircraft of one type that seats 360 passengers

(b) air trafic in the form of 2*n* aircraft of another type that seats 180 passengers.

The following general statement can be made: Many casualties, if not a majority, are caused by a human error of a single person. Today there exist a large number of technical and procedural provisions that lead to a "fail-safe" situation (i.e., that prevent a fatality) in case of human error. Continuous progress was made during the last 2 decades in developing such provisions, but there is no reason why they should not be realized with model (b) as well as with model (a). So, in the rare case when all the provisions are unable to prevent the fatal consequences of a human error, system (a) is in a marked disadvantage compared with system (b), because the number of passengers affected can be twice as high with (a) as with (b). Fortunately, the probability of this situation is low today. The accident involving a DC-10 in Chicago in May 1979 was a shock to the public. In its consequences, it showed that the authorities, with the means at their disposal, are effective in preventing similar future dangers for air passengers.

There is also the possibility of midair collisions. This is an attribute of the existing air-traffic control system in a time of growing aircraft density and is only one link of the chain of events caused by a human error. Progress was made by reducing the number of "near, but avoided midair collisions" during the last few years. In general, the probability that *m* passengers are killed in a midair collision is the same in both cases (a) and (b).

The third point of comparison is aircraft motion on or near an airport, including take-off and landing procedures. According to casuality statistics, the zone with a diame-

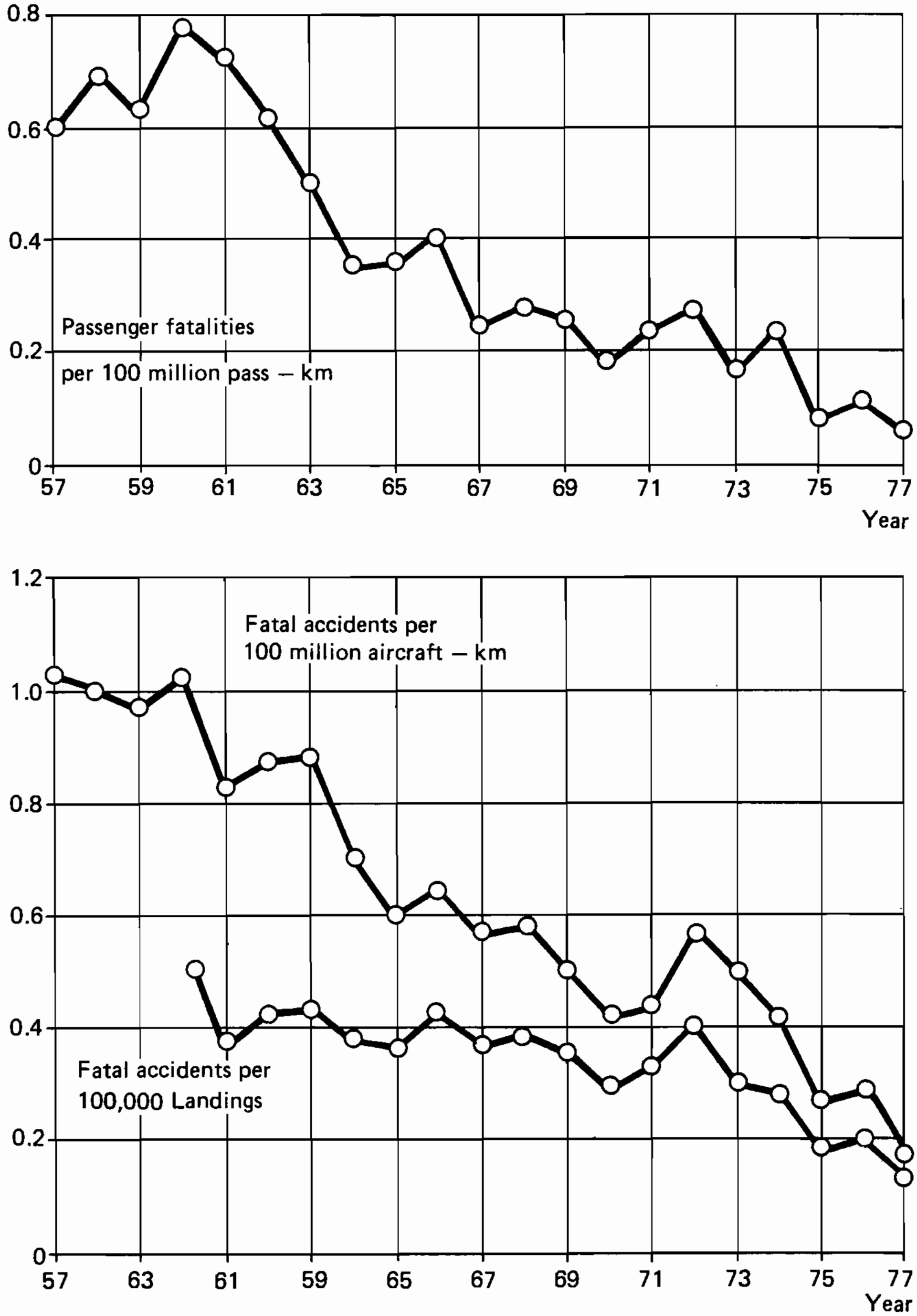


FIGURE 3.5 Fatality frequencies in scheduled ICAO airline services.

ter of about 50 km around an airport is the most dangerous area of air traffic. The situation in this area is watched by means of an analogue representation on a number of radar screens. Today, the observation of the situation which changes continuously is facilitated for the air control personnel by semiautomatic techniques. The observation task for the single air controller becomes more and more difficult when the number of objects on the screen increases. G. Fechner's law in psychophysics tells us that, if the number of stimuli (here: number of objects to be observed) increases, the human potential for reaction grows only with increments that become smaller and smaller. With the high air-traffic density during rush hours, there is the severe danger of stimuli overload of the human air controller. It is now important to understand that in our model situation (a) with big aircraft, this danger is reduced considerably because the same traffic volume is realized with half the number of aircraft in case (b).

In summing up these air safety comparisons, we should have in mind that our model situations were idealized and simplified with respect to reality. But we get the impression that "in toto" there is no marked safety disadvantage if a larger fraction of the air traffic volume is carried by high-capacity (big) aircraft.

This statement is only true for the high level of air safety enforced by the regulations for air control personnel in, for example, the FRG today:

- special training for 3 yr (after college degree)
- detailed program for breaks from working; interruption after every 2 hr of service
- 33 working hr per week
- 4 weeks of treatment at a health resort after every 5 yr of service
- retirement at an age of 53.

There are plans for a higher degree of automation in air traffic control, especially in the United States, for the eighties. But even when they are realized, the human air controller will remain the decisive factor for ensuring a high degree of air safety.

3.2.3 Environmental Aspects

After safety, the dimensions of airport facilities and aircraft noise are two important factors in air transportation. The scale of the aircraft has a pronounced influence on the space that the airlines require of the airport authorities. The history of the area requirements for aircraft in landing position on the airfield is shown in Figure 3.6 for long-range aircraft and medium-range aircraft separately. This is the consequence of the development of aircraft weight shown in Figure 3.1. The diameter of the area needed for an aircraft in landing position (if safety distances are included) remained nearly constant until the mid-fiftes. It then rose suddenly and reached its preliminary maximum for the Boeing 747 and the Airbus A 300 around 1970. These dimensions determine the size and locations of bridges for passenger embarkment and disembarkment, the scale of airport buildings, and so on. To a certain extent there is a competition between airport authorities to fulfill such requirements to enable the airlines the use of modern, high-capacity aircraft. On the other hand, there is the insight that the problems with passenger transportation to and from the airport, with services for the passenger at the airport, with baggage handling, and so forth, may prevent a further increase of aircraft size and seat capacity over the maximum level reached today. The "super airport" of Dallas, Texas with its further growth potential is only a singular example. It demonstrates what would be needed if aviation had to further increase its dimensions, but it could not be afforded in most places in the world.

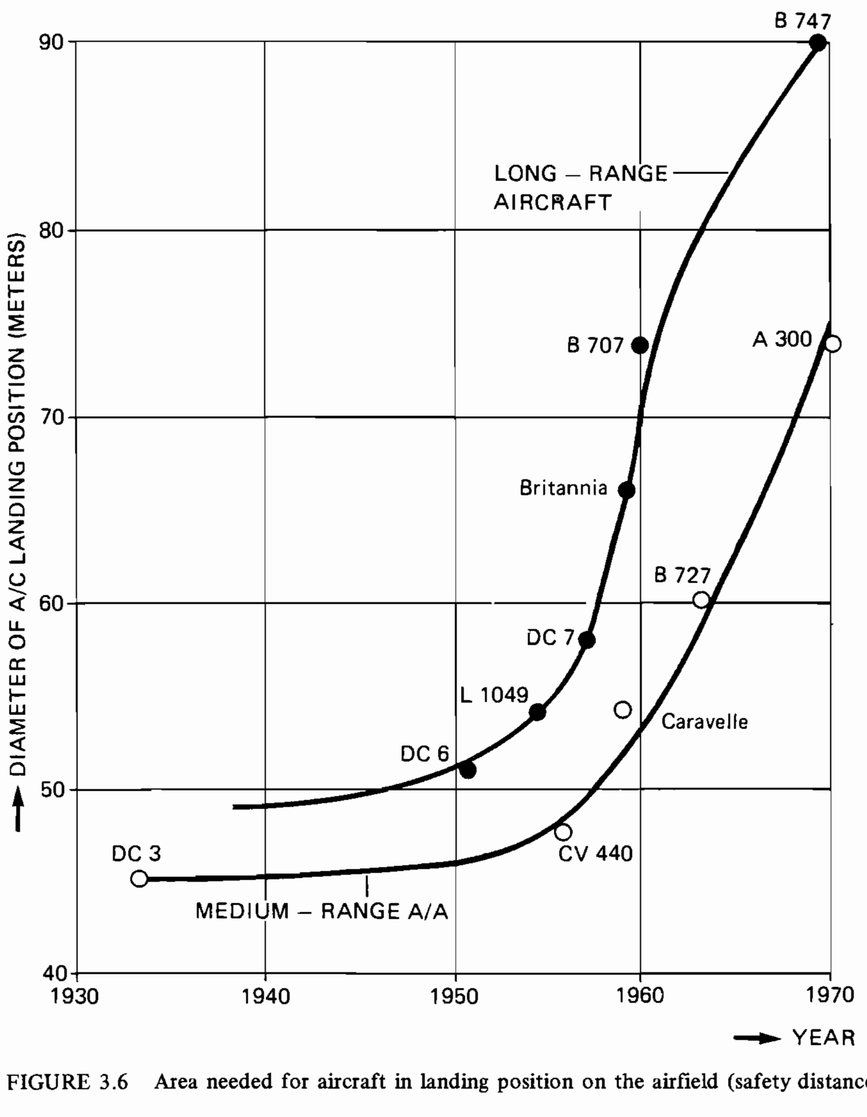


FIGURE 3.6 Area needed for aircraft in landing position on the airfield (safety distance included).

The next point, aircraft noise, is an important factor in selecting aircraft and engine types and in the striving of enraged groups of the population to prevent new airports. During the last 10 yr, considerable progress was made by the engine manufacturers in reducing noise and smoke emitted by airline aircraft. The major nuisance on and around airports still perceived today is created by the engines developed during the late fifties and early sixties. With a new generation of engines primarily installed in wide-body aircraft, a sound pressure reduction of at least 10 decibels was achieved in comparison with the Boeing 707/DC8 four-engine transport fleet. These reductions are enforced by noise standards specified by the US Federal Aviation Administration (FAA) for the airworthiness certification of new aircraft.

A noise reduction is achieved by a new design principle of the nacelle in which a novel sound-absorbing material is used. And the steeper gradients of the take-off and landing flight paths are responsible for a reduction of the perceived noise near the airport.

Design modifications are economically justified with relatively big engine units because they are needed for heavy aircraft. Because of the high reliability level that was reached with turbofan engines, it was possible to equip the 150-ton Airbus A 300 with two engines only.

The progress can best be shown if we compare the area with 90 decibels or more near an airport from which an older and a modern aircraft are taking off. The length of this "noise carpet" is more than 60 km for the Boeing 707, but only 11 km for the Airbus A 300.

The reduction of pollutants in the exhausted air is another problem with which the airlines are confronted. For a few years, newly designed smokefree combustion chambers have been installed whenever an engine is overhauled. During this decade, engine manufacturers are steadily but slowly reducing the amounts of unburned hydrocarbons, carbon monoxide, and nitrogen oxide in engine exhaust.

If we review the environmental effects discussed, it is obvious that big aircraft have their most severe handicap with respect to their dimensions, their numbers of passengers, and the resulting handling problems at the airport. However, the problems of noise and exhaust are more easily tackled with bigger unit sizes.

3.3 PRESENT STATUS AND FUTURE EXPECTATIONS

3.3.1 Air Passenger Market

The high growth rates in international air transportation during the sixties induced the airlines to order the new big Jumbo Jets. But an economic *baisse*, especially in the United States, damped the air travel demand. Increases and differentiations in the tariff system and more intense competition led to reduction in the return on investment for the airlines. The mean seat occupation factor for all ICAO countries was 59 percent in 1960, but only 52 percent in 1971; see Khare (1976). There was an increase in 1974, the year after the oil crisis, but this was caused by an out-of-service action for whole parts of the fleet, including Jumbo Jets.

In contrast to the airlines, the unscheduled air transportation carriers had, in 1972 and 1973, growth rates of about 15 percent. In the meantime, the airlines have participated in the unscheduled part of air transportation, which is a vitalizing force for tourism. In the international market segment, about 25 percent of the traffic volume is bought unscheduled.

The demand situation in the passenger market is now stable. There was a record year in 1978, e.g., the airlines located in the United States had a passenger increase of 17 percent. The yearly growth rate of 7 percent forecasted for the next decade, (see Figure 3.3), seems not impossible. But of course, this is only one scenario for the development, and a certain slowdown during the eighties cannot be excluded from consideration.

By no means have the present returns on investment the same growth rates as the volume transported. Many airlines are in the red and some of them may be forced to merge with stronger competitors. The obstacles to higher efficiency are obvious to us when we are told that, for example, Lufthansa allocates 32 percent of the total cashflow to its personnel and that between 20 and 24 percent of direct operating costs are fuel costs. These are two factors that have their own growth rates that can be influenced only to a very small extent.

The forecasts for the next decade reflect these facts: because of their higher efficiency, the airlines will use wide-body aircraft for long and medium ranges whenever they can be used economically. But for certain market segments, narrow-body aircraft will stay in service even in the foreseeable future.

From the 4,335 jet aircraft that are now in service among ICAO airlines, already 770, or 16 percent are wide-body aircraft (Boeing 747, DC-10, Airbus A 300, and so forth). According to Geddes (1977), McDonnell Douglas, in its forecast for 1991, predicts a total number of jet aircraft of 5,950, of which 2,760, or 46 percent, will be wide-body aircraft. Only because of this increase in the mean seat capacity per aircraft will it be possible to realize the predicted duplication of the air traffic volume during the next 12 yr, (see Figure 3.4), with a growth of the number of jet aircraft of only 26 percent. This is an important result.

If we look at this forecast in terms of transport capacity, the seat-miles offered per year are a better parameter than the sheer numbers of aircraft per category; see Figure 3.7. In the short- to medium-range market, the fraction contributed by wide-body aircarft will increase, but the narrow-body types which have the lower seat capacity will maintain their importance for routes with smaller transport demand. On the other hand, for long ranges the relation between wide-body and narrow-body aircraft will undergo a steady change in favor of the big units.

Following this optimistic view which is not unrealistic but has certain ingredients of a "self-fulfilling prophecy" the manufacturers are preparing for a new generation of passenger-transport aircraft in addition to extensions to the current programs. To name just a few new projects, we have for the short- to medium-range market the narrow-body Boeing 757 and Airbus A 300 B 10, and the wide-body Boeing 767 and McDonnell Douglas DCX-200. A project for long-range flights is the Boeing 777, which has a wide body but a smaller seat capacity than the Boeing 747 "Jumbo Jet."

All intentions suggest that during the next 10 to 12 yr the number of big aircraft of the 360-seat class will increase as will the number of aircraft with smaller seating capacity. Recently, the Boeing Corporation stated that it would be possible to develop a "Super Jumbo Jet" with a seating capacity for 620 passengers in the mid-eighties. They added in this announcement that the realization of these plans would depend on demand and that airport capacity rather than aircraft technology would be the limitation. It can be expected that many airlines today are not enthusiastic about prospects of this kind. In the long run, the energy problem will become a decisive factor and this might lead either to a limitation of aircraft size or, perhaps, to the operation of a small number of "super size" aircraft; see section 3.4.

To withstand the economically dangerous position between strong price competition on the one hand and a tendency toward increasing cost on the other hand, the airlines must look for other improvements than a further increase in the aircraft unit size. The most promising technological option may be a steady decrease of the structural weight of transport aircraft. It seems possible to replace certain parts of the airframe made of aluminium-based alloys by composite materials such as plastics reinforced with carbon or boron fiber; see Flemming (1976). It is expected that a development time of at least 10 or 12 yr will be necessary until such a replacement is realized for about 70 percent of the aircraft structure. The target set for the year 2000 is a passenger aircraft with a given gross weight and the same cruising speed as today's aircraft, but which carries a payload 50 percent heavier.

All these forecasts are for a so-called "surprise-free" development, i.e., a number of more pessimistic scenarios would have to be studied before the picture obtained would be complete.

3.3.2 Supersonic Transportation

Until now, all criteria and reasons for choices in passenger aviation have been surprisingly just as rational as those in other sectors of industry, the economy, and society. In discussing supersonic transportation, we are entering a "domain of irrationality" within air transportation. It is worth notice that one of the leading nations in aerospace technology, the United States, at present does not participate in this field of activity.

Seen historically, it was a dramatic advance when in the late fifties with the introduction of long-range jet aircraft the flight duration across the North Atlantic was reduced from 18 to 7 hr. This fact, together with a permanent increase in flight safety, flight-plan reliability, and comfort, induced a shift of many sea-going passengers to the air. Compared with this impetus, the further reduction of the flight time from Paris to Washington to about 4 hr on the Anglo-French Concorde in the seventies is attractive from a rational point of view only for a small fraction of extremely busy managers. These critical remarks must be made in spite of the fact that the Concorde is a success technologically.

Many important arguments against the introduction of supersonic transport were already made by Bo Lundberg (1964) in his famous study.

Important aspects of his statements are as true today as the were in 1964, especially those related to the lack of efficiency and the noise ("sonic boom") problem.

The few scheduled supersonic flights offered by Air France and British Airways were very attractive and had a relatively high usage rate when they were introduced; but in spite of the higher fares, these flights do not bring back the money that they cost. There is no doubt that fuel shortages and price increases will first hit this sector of aviation.

The surprise-free forecast shown in Figure 3.7 yields a demand for 118 supersonic transport aircraft in 1990 in the ICAO countries. This is an impressive figure, but for an aircraft type after the Concorde, it would not be enough to reach the "break-even" point.

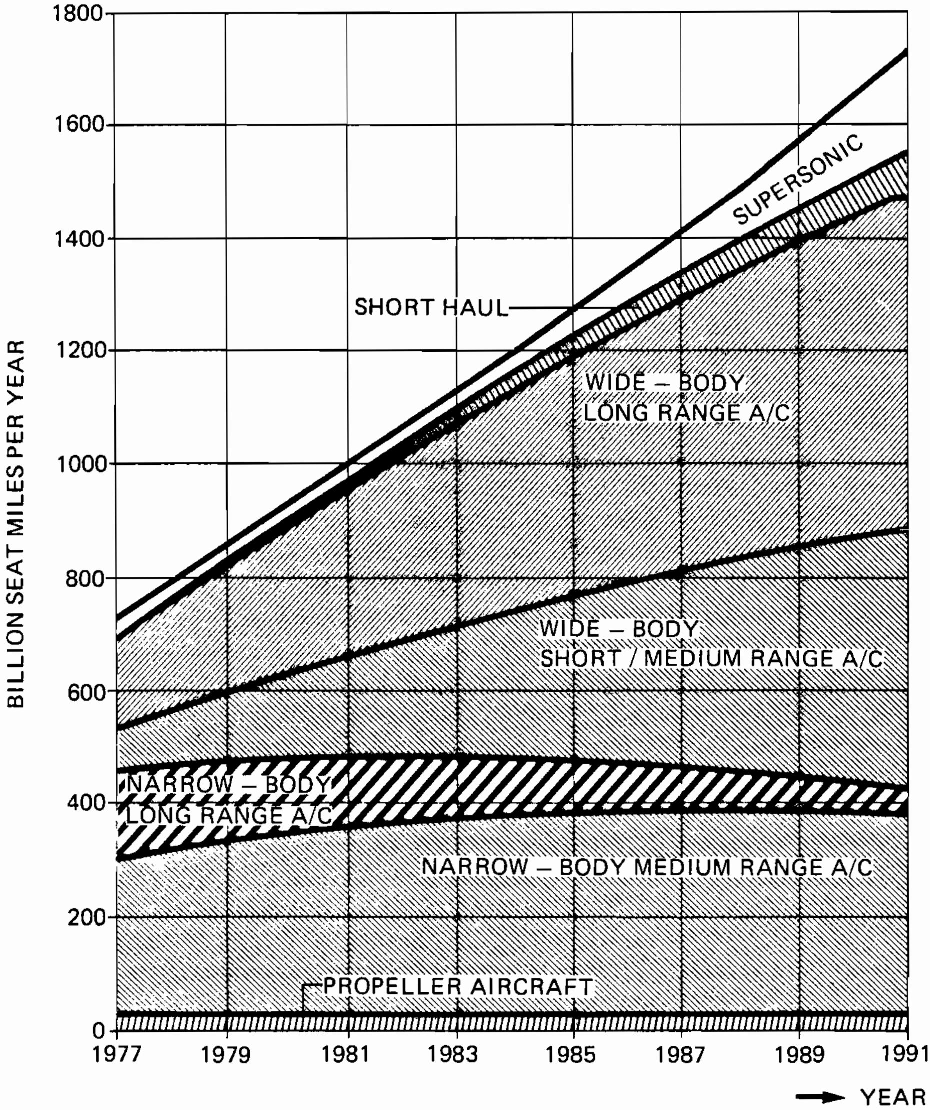


FIGURE 3.7 Distribution of offered ICAO seat capacity to aircraft categories.

3.3.3 The Air Freight Market

So far, mainly passenger transport by air has been considered. But the air freight market is of growing importance. In the fifties, the first scheduled flight services were introduced that were pure freight transports. In the decades before, a growing number of million ton-kilometers were moved every year, but they were carried under the floor of the passenger cabins of airlines. Today there exist a number of passenger aircraft types that can be converted to transport freight, e.g., the Boeing 707-320 C for a maximum load of 40 tons and the Airbus A-300 C 4 for 43 tons. A pure cargo aircraft in which the main fuselage section can be used by loading through a front door is the Boeing 747-200 F, the freighter version of the Jumbo Jet, with a maximum load capacity of nearly 91 tons.

The economic situation of the freight market during the seventies was more or less parallel to the passenger market: there was a shrinking return on investment between 1968 and 1974 and a nearly steady growth of demand and of freight volume through all the years; see Koehler (1976). In most forecasts, it is expected that this increase will bring a considerably higher freight/passenger ratio in 1990 than we have today.

With respect to the problems of scale, there is the interesting question of whether this development will lead to a project for a "big freighter" in the gross-weight class of about 400 to 500 tons. This could be an economic solution, especially under the conditions of the "post-oil age."

3.3.4 The Long-Term Energy Problem

There is no doubt among the kinds of primary energy the oil fraction will decrease, whether in 20 or in 40 yr. Some people believe that the shrinking oil supply will first be reserved for automobiles and aircraft, and later on only for aircraft. Even if this process were carried out with a high degree of order and discipline – there can be no long-term planning for air transportation without considering the "post-oil age."

Shortly after the oil crisis in 1973/1974, the first project drawings of transport aircraft with large hydrogen pods on their wings were shown. Studies on hydrogen-propelled aircraft have been carried out to explore the many problems that had to be solved, for example, to store liquid hydrogen in the lower part of the fuselage, to use the expected reductions in structural weight for maintaining a sufficient payload capacity, and to establish safe procedures for handling hydrogen and operating with it.

A much more attractive alternative both the aircraft designers and for the airlines would be the use of more conventional synthesized fuel from, for example, coal liquifaction. Economic factors and possibly pressures that are to a large extent unknown today will be of importance in determining the direction of further development. With these prospects, the possibility of economically attractive aircraft weighing more than those in "Jumbo Jet" class cannot be excluded.

3.4 FINAL ASSESSMENT

The history of air transportation proves to be a field well suited to the study of technological, economic, and environmental problems of unit scale. The development until 1970 was characterized by a steady increase of aircraft size which was unquestionably caused by the fact that the economically and technologically justified maximum unit scale was not reached before. Fortunately for aviation, the Boeing 747 "Jumbo Jet," the biggest aircraft, became a success after initial difficulties. As a consequence, so-called wide-body aircraft are being introduced in growing numbers both in the long-range and in the short-to medium-range market. For the next decade, a well-balanced mix of smaller and larger aircraft is expected in both markets. For nearly 10 yr a passenger transport aircraft with about 380 seats (i.e., the Boeing 747) has been thought to be the reasonable maximum – an opinion which is expected to be maintained. The main obstacles to further growth are handling difficulties for passengers and baggage at the airports in addition to load factor problems.

By using larger numbers of big aircraft (of current size) it would be possible to duplicate the air traffic volume between 1978 and 1991 while increasing the number of aircraft by only 26 percent. This is one of the keys for maintaining a high level of air safety; the other keys are engine technology, electronics, and operation procedures for air control.

Not aircraft size but transport with supersonic speed is the "domain of irrationality" in aviation both from an economic and a societal point of view. If we follow a line of rational arguments we should expect a shrinking of the currently very small supersonic market.

Around the year 2000, or earlier, aviation will be confronted with the "big oil-shortage problem." It depends on many parameters external to aviation whether this is only a new challenge for aircraft and engine designers which can be tackled successfully, or whether the volume of air transportation will decrease and finally stabilize at a considerably lower level.

A lesson from this case study, which may be important in the consideration of other industries, is that it is possible to keep a large-scale technology "under control" and to minimize the adverse effects for the environment and for society. Two prerequisites for achieving this seem to be a steady development for decades in an expanding market with many innovative inputs from related areas and a long tradition of applying scientific methods to all aspects of the field of activity.

REFERENCES

Flemming, M. 1976. Research for Modern Fiber Technology. Aerospace International 3: 22 - 27.

"Flight." 1978. Safety in Air Transportation. Flight International 115: 184 – 187.

Geddes, J.P. 1977. The International Air Transportation Market Until 1991 - The Recent McDonnell Douglas Report. Interavia 32: 987 – 990.

Khare, G. P. 1976. Efficiency Problems of Scheduled Air Transportation. Interavia 31: 927 – 928.

Koehler, S. 1976. A Penetrating Change of the Air Freight Market. Interavia 31: 929 – 930.

Lundberg, Bo K.O. 1964. Pros and Cons of Supersonic Aviation in Relation to Gains or Losses in the Combined Time/Comfort Consideration. Journal of the Royal Aeronautical Society 68: 611 - 630.

Simon, E. 1970. Die Wirtschaftlichkeit der Großflugzeuge (The Efficiency of Big Aircraft.) Der Mensch und die Technik, Beilage der Süddeutschen Zeitung, 16. April 1970: 9.

CHAPTER 4 THE OPTIMAL SIZE OF SUBCRITICAL FOSSIL-FUELED ELECTRIC GENERATING UNITS

J.C. Fisher
General Electric Company,
Schenectady, New York

4.1 INTRODUCTION

The term economy of scale denotes the often-observed decline in unit cost of a product or service that accompanies an increase in the scale of the equipment or activity by which it is produced. When applied to the cost of electricity, economies of scale can exist for an entire utility system or for any of its component parts.

In this paper, I examine a portion of the U.S. utility system, the generating unit. A generating unit is a boiler-condenser-turbine-generator heat engine and electric generator, together with accessory equipment and structures. A utility plant is a named geographic site that may contain one or more units. The units are constructed from time to time over a period of many years, with the result that the plant capacity tends to increase over time.

U.S. utilities report annual statistics to the U.S. government on a plant-by-plant basis, and statistics for about 96 percent of the capacity of utility-owned fossil-fueled steam-electric plants are published by the Energy Information Administration of the U.S. Department of Energy (1977). The statistics include capital cost and many factors that affect it, such as unit number, type of construction, fuel capability, heating value of fuel, steam pressure and temperature, turbine design, geographic region, and year of commercial operation. When a new unit enters commercial operation, the capital cost of the plant jumps upward by the capital cost of the unit, so that capital costs of units can be determined from the sequence of annual costs as correlated with the addition of new units.

I have collected and organized statistics for all fossil units reported to enter commercial operation in the United States from 1958 through 1977, about 750 units in total, and together they form the basic data base for this analysis and others to follow.

Many factors affect the cost and performance of generating units. In order to reveal the economy of scale, it is necessary to strip away the influence of as many extraneous factors as possible. Different technologies can lead to different scale effects. We cannot assume that scale effects are the same for a 2400psi/1000F/1000F/3600rpm subcritical steam unit, a 3500psi/1000F/1000F/3600rpm supercritical steam unit and a 2400psi/1000F/variable-F/3600 & 1800rpm cross-compound unit. (In these shorthand expressions for technology, the first term is turbine inlet steam pressure in pound per square inch, the second is inlet steam temperature in degress Fahrenheit, the third is reheat steam temperature, and the fourth is generator speed or speeds in revolutions per minute.)

In this analysis, I focus on a single well-established technology, 2400psi/1000F/1000F/3600rpm. I omit the earliest 14 such units, which entered commercial operation from 1960 to 1964, because they may represent immature technology; this leaves a final data base of 136 units entering commercial operation from 1968 to 1977.

4.2 CONSTRUCTION COST OF 2400/1000/1000/3600 FOSSIL UNITS

Several parameters, aside from scale, can be shown to affect construction cost. In a preliminary mutiple-regression analysis, based on Formula 1 in the Appendix, I explored the influence of a number of parameters. They are listed here together with the factors by which regression analysis suggests that they influence construction cost. Each factor is followed by its 95-percent confidence limits in parentheses. (Later in the paper the regression formula is modified and the magnitudes of these factors are revised.)

1. Order of construction of a unit at a plant site, compared with units constructed after units 1 and 2.

    a. Unit 1 cost factor: 1.18 (1.09 – 1.28)
    b. Unit 2 cost factor: 0.86 (0.80 – 0.93)

2. Fuel capability, compared with coal-burning capability.

    a. Oil or gas only cost factor: 0.73 (0.66 – 0.81)

3. Type of construction, compared with conventional construction and once-through cooling.

    a. Outdoor or semi-outdoor cost factor: 1.05 (0.95 – 1.16)
    b. Cooling tower cost factor: 0.97 (0.84 – 1.11)

4. Type of ownership, compared with investor ownership.

    a. Public or cooperative cost factor: 0.85 (0.76 – 0.96)

5. Region of United States (Federal Power Commission (FPC) classification), compared with region 3.

    a. Region 1 cost factor: 1.59 (1.38 – 1.83)
    b. Region 2 cost factor: 1.16 (1.03 – 1.30)
    c. Region 4 cost factor: 1.17 (1.03 – 1.33)
    d. Region 5 cost factor: 0.85 (0.76 – 0.95)

e. Region 6 cost factor: 1.29 (1.08 – 1.53)
f. Region 7 cost factor: 1.41 (1.17 - 1.70)

6. Year of commercial operation, compared with 1975.

a. 1965 – 67 cost factor: 0.42 (0.37 – 0.49)
b. 1968 cost factor: 0.45 (0.39 – 0.53)
c. 1969 cost factor: 0.54 (0.47 – 0.62)
d. 1970 cost factor: 0.57 (0.49 – 0.66)
e. 1971 cost factor: 0.58 (0.47 – 0.70)
f. 1972 cost factor: 0.79 (0.69 – 0.89)
g. 1973 cost factor: 0.76 (0.67 – 0.88)
h. 1974 cost factor: 0.93 (0.81 – 1.07)
i. 1976 cost factor: 1.36 (1.16 – 1.60)
j. 1977 cost factor: 1.32 (1.15 – 1.51)

When corrections are made for all of these extraneous factors, the underlying scale economy is revealed as follows, compared with units rated 400 – 499 MW.

a. Rating < 300 MW cost factor 1.12 (0.98 – 1.29)
b. Rating 300 399 MW cost factor (0.88 – 1.09)
c. Rating 500 – 599 MW cost factor 0.95 (0.86 – 1.05)
d. Rating 600 – 699 MW cost factor 0.98 (0.87 – 1.11)
e. Rating ⩾ 700 MW cost factor 0.92 (0.79 – 1.07)

All the above factors were determined by multiple regression using Formula 1 in the Appendix.

An examination of the results of this preliminary regression shows some very plausible regularities. For example, Figure 4.1 shows the rate of inflation of construction cost as reflected in unit cost relative to that for 1975. Since 1968, the rate of inflation has been roughly constant. (Prior to 1968, during the 1950s and 1960s, construction costs declined and the costs for 1965 – 67 may reflect the end of the turnaround.) The fluctuations from year to year in Figure 4.1, where a smooth progression of inflation would be expected, suggest the level of uncertainty in the data due to causes not identified (or not correctly parameterized) in the regression formula.

Figure 4.2 plots the trend of cost with unit rating, and suggests that a scale economy exists over the full range of ratings in the data base.

The other cost factors are generally plausible, with the exeption of those for type of construction, where unconventional construction would be expected to reduce cost and cooling towers would be expected to add to cost. However, these factors are highly correlated with region, and their uncertainties are large. I judge them to be instances of statistical fluctuation.

In order to obtain the best possible estimate of the scale effect, I dropped from the data base the 16 units that went into commercial operation from 1965 to 1967 and retained only the 120 units that went into operation from 1968 to 1977 when inflation was essentially constant. I dropped the variables representing type of construction as

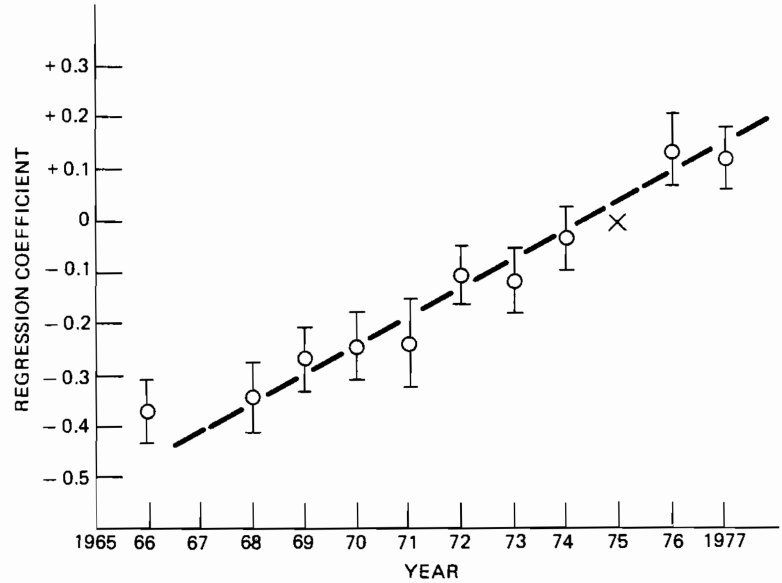


FIGURE 4.1 Regression coefficients for log ($/kW) versus year of commercial operation, as determined from Formula 1 in the Appendix. Inflation was roughly constant over the time period 1968–1977. The error bars indicate the 95 percent confidence limits and suggest the magnitude of the residual uncertainty in the data.

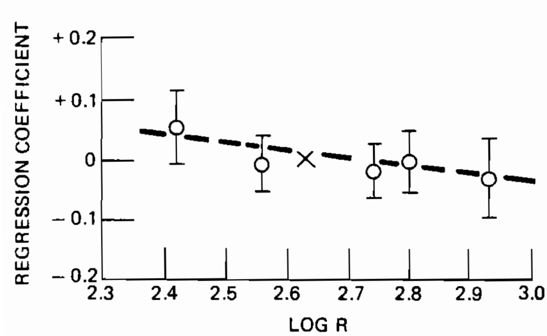


FIGURE 4.2 Regression coefficients for log ($/kW) versus log R where R is rating in MW, as determined from Formula 1 in the Appendix. The negative slope indicates an economy of scale.

lacking statistical significance. Then I modified the regression formula to assume a constant percentage annual inflation and a scale economy according to the relationship

$$(\$/\text{kW}) \sim R^{-a} \tag{1}$$

where R is unit rating and a is a constant. The inflation rate and the constant a are determined from the revised regression Formula 2 in the Appendix.

The final regression parameters, together with their 95-percent confidence limits, are summarized in Table 4.1.

TABLE 4.1 Regression parameters for the construction cost of 2400/1000/1000/3600 units, based on Formula 2 in the Appendix. (Figures in parentheses are the 95 % confidence limits.)

1.	Unit 1 cost factor	1.21	(1.11		1.31)
2.	Unit 2 cost factor	0.87	(0.80	–	0.94)
3.	Oil or gas only cost factor	0.74	(0.67	–	0.82)
4.	Public or cooperative cost factor	0.85	(0.75	–	0.96)
5.	Region 1 cost factor	1.51	(1.33	–	1.71)
6.	Region 2 cost factor	1.08	(0.96	–	1.21)
7.	Region 4 cost factor	1.15	(1.03	–	1.29)
8.	Region 5 cost factor	0.80	(0.72	–	0.90)
9.	Region 6 cost factor	1.21	(1.04	–	1.41)
10.	Region 7 cost factor	1.33	(1.11	–	1.59)
11.	Annual inflation (%)	12.7	(11.3	–	14.2)
12.	Scale economy exponent (Eq. 1)	0.16	(0.02	–	0.29)

The effect of scale, as determined from the revised regression, is given by the relationship

$$(\$/\text{kW}) \cong (\text{Rating})^{-0.16} \tag{1a}$$

so that a doubling of rating reduces the construction cost per kilowatt by about 10 percent. If this were the full story on capital costs, the larger a unit the lower would be the cost of its electricity. However, we must consider the manner in which reliability affects cost, and the manner in which scale affects reliability.

4.3 RELIABILITY COST OF 2400/1000/1000/3600 Fossil Units

Reliability, expressed by the capacity factor, is an important contributor to the overall cost of capital per unit of electrical output. (The capacity factor CF is the ratio of output actually achieved over a period of time to what would have been achieved at full design power.) For a base-load unit, the capital cost per kilowatt of average capacity actually achieved is

$$(\$/\text{kW})_{\text{actual}} = \frac{(\$/\text{kW})_{\text{construction}}}{(\text{CF})} \tag{2}$$

If a unit costs \$200/kW to construct, but has a capacity factor of only 0.5 averaged over the year, the capital cost per kilowatt of average annual capacity is 200/0.5 = \$400/kW.

If capacity factor were independent of rating, it would not affect the scale economy. However, analysis of fossil unit performance (Fisher 1978) shows that the base-load capacity factor declines with unit rating according to a relationship of the form

$$(\mathrm{CF}) \cong (b)^{R/1000} \tag{3}$$

with b = 0.70 for mature subcritical fossil units of mature design. Although it was derived from analysis of a population of fossil units somewhat different from those for which the construction cost scale economy in Eq. (1a) was determined, I believe it is close to the mark. (Although this brief analysis considers only the influence of reliability on the capital cost of base-load units, its influence on cycling units is essentially the same (Fisher 1978).)

When construction cost is modified by the base-load capacity factor to obtain the actual capital cost of a generating unit, the result in terms of the independent scale effects for construction cost and capacity factor is

$$(\$/\mathrm{kW}) \cong (R/1000)^{-a} \quad (b)^{-R/1000} \tag{4}$$

realized capital cost	construction cost	CF reliability adjustment

The capital cost per kilowatt of actual average capacity is a minimum when

$$R/1000 = a/\ln b \tag{5}$$

If a = 0.16, b = 0.70 as determined in this paper and by Fisher (1978), the value of $R/1000$ that minimizes capital cost is

$$R/1000 = 0.16/\ln 0.7 = 0.45$$
$$R \approx 450 \text{ MW}$$

There is substantial uncertainty in this value, since the 95 percent confidence limits are very wide, approximately 100 – 800 MW.

Overall, the analysis suggests that if construction cost and capacity factor were the only significant factors affecting the cost of electricity at the power plant, the optimal rating of a generating unit would be somewhere near 350 – 500 MW. (This optimum is somewhat larger than values I have previously estimated and have reported in informal talks. In earlier analyses, I included units with cross-compound turbines driving 3 600 rpm and 1 800 rpm generators along with the simpler 3 600 rpm units driving a single 3 600 rpm generator that are analyzed here. Since cross-compound units tend to cost more (although they compensate by tending to be more efficient and require less fuel) and since the proportion of cross-compound units tends to increase with unit size, the presence of cross-compound units in earlier data bases tended to mask the decline in construction cost of 3 600 rpm units. Since the present analysis applies to a single technology, I believe it gives a more reliable estimate of overall scale economy.)

4.4 EFFICIENCY COST OF 2400/1000/1000/3600 FOSSIL UNITS

Other scale-related factors can influence the cost of electricity, and it is possible that a change of efficiency with rating could have such an influence.

Unfortunately, the data from the Energy Information Administration (1977) report heat rate (BTU/kWh, inversely related to efficiency) only by plant, not by unit. Hence it is possible to relate heat rate to unit rating only for plants that consist of a single unit or of two or more identical units. The complete data base contains 36 such plants for which good heat rate values were available, providing a truncated data base for estimation of the dependence of efficiency on rating.

An analysis suggests that heat rate depends on the type of fuel, on the energy content of coal if used as a fuel, on the use of cooling towers, and on unit rating. A regression based on Formula 3 in the Appendix suggests only marginal significance for each of these variables, although in aggregate they are significant. The influence of rating on heat rate is to add to the basic heat rate of about 10,000 BTU/kWh an increment $(0.28 \pm 0.70) R$ where R is the rating in MW. There is thus a hint of an increase in heat rate (a decrease in efficiency) as rating increases, but the amount is small and uncertain. Perhaps the best that can be said is that there is no evidence that efficiency improves as rating increases.

4.5 OTHER SCALE-RELATED FACTORS THAT MAY AFFECT THE BUS-BAR COST OF ELECTRICITY

At this point, the data on operating units suggest an optimal rating of about 350 – 500 MW. This rating is close to the 200 – 300 MW estimated optimum for British coal-fired units (Abdulkarim and Lucas 1977), and I do not believe that the difference is statistically significant.

Several additional factors tend to favor smaller-size units. Smaller units afford greater siting flexibility because they can be dispersed and sited closer to the load. Smaller units allow a reduction in reserve margin for a given system reliability, since the loss of a small unit is less demanding on reserve capacity. There may be a reduction in maintenance personnel for smaller units because of their higher reliability. Although these factors may be of significance, I have not quantified them.

A final factor may be of considerable significance. Now that the limits to scale economy have been established for units designed and constructed one or two at a time, and there is no longer any motivation to seek economies by designing and building ever-larger units, it becomes possible to standardize a design and replicate a large number of identical units. This opens up the possibility of a new dimension in scale economy – the achievement of cost reduction through experience with repetitive construction of identical units.

Experience in many industries suggests that repetitive production of a series of identical items results in cost reductions of 10 – 30 percent for every doubling of cumulative production, or equivalently

$$\$/\text{unit} \sim n^{-c} \qquad (6)$$

where n is the cumulative number of units produced and c is a constant. Note the similarity of this relationship to the scale economy relationship (\$/kW) $\sim R^{-a}$. Indeed it is the same relationship with the dimension of scale as the number of identical units n in place of the rating R.

I assume that Eq. (6) with an appropriate choice of c represents the decline in cost with unit number to be expected for a series of identical generating units, and consider a hypothetical example to illustrate its potentical significance for the capital cost of electricity. I assume that the construction cost of a 1000 MW unit is \$500/kW; that a doubling of rating will decrease construction cost per kilowatt by 10 percent, which is close to the figure determined in this paper; and that a doubling of the number of identical units will decrease construction cost per unit by 15 percent, an amount representative of experience in other complex technologies. Then I consider the costs of several alternatives for constructing 1 000 MW of new capacity as summarized in Table 4.2.

TABLE 4.2 Hypothetical example of the influences of unit scale and experience scale on the construction cost of 1000 MW (\$/kW).

	One 1000 MW unit	Two 500 MW units	Four 250 MW units	Eight 125 MW units	Comment
Unit 1	500	556	617	686	\$/kW for first unit increases because its size decreases.
Unit 2		472	525	583	
Unit 3			484	537	
Unit 4			446	496	\$/kW for subsequent units decreases because of accumulating experience
Unit 5				476	
Unit 6				457	
Unit 7				439	
Unit 8				421	
Average \$/kW	500	514	518	512	Average construction cost stays roughly constant
\$/kW adjusted for capacity factor	714	614	566	536	Average capital cost per kW of available capacity declines as unit size decreases

With the parameters chosen for this example, the smaller the unit rating the lower the capital cost in \$/kW of actual average capacity. The optimal unit size shrinks to the smallest size that can efficiently utilize 2 400 psi/1 000 F/1 000 F steam for 3600rpm generation. This technological lower limit is probably in the range 100 – 200 MW. Because I expect the true scale parameters for generating unit construction and operation to be close to those used in the hypothetical example, I expect the optimal rating for a series of generating units to be at this technological limit.

The replication of a series of identical generating units opens up an entirely new and profoundly different avenue for reducing the capital cost of generating capacity. The economy of scale assumes a new form, and manifests itself as the reduction of cost that can be achieved through the scale of operations in replicating large numbers of identical units. I believe that the potential for cost reduction along this new avenue is substantial.

APPENDIX

In the regression formulas, the symbol δ (unit 2) stands for 1 when the unit number is equal to 2 and for 0 when the unit number is not equal to 2; and in general δ (V_j) for any discrete variable V_i stands for 1 when $i = j$ and for 0 when $i \neq j$.

The preliminary regression for construction cost does not assume that the inflation rate is constant over the full time period or that the scale economy is monotonic over the full range of ratings. Rather it seeks to discover such ranges if they exist, and to determine the significance of other variables that could mask the underlying scale economy. The formula relates construction cost to the cost of a standard unit (unit number 1 or 2, coal fuel capability, conventional construction, no cooling towers, investor ownership, region 3, entering commercial operation in 1975, rating 400 – 499 MW). Each variable is assumed to multiply the cost of the standard unit by a factor that can be determined from multiple regression. The regression is linear when the dependent variable is log ($/kW).

Formula 1

$$
\begin{aligned}
\log(\$/\text{kW}) = \; & A_0 \\
& +A_1 \;\; \delta\,(\text{unit 1}) \\
& +A_2 \;\; \delta\,(\text{unit 2}) \\
& +A_3 \;\; \delta\,(\text{oil or gas fuel only}) \\
& +A_4 \;\; \delta\,(\text{unconventional construction}) \\
& +A_5 \;\; \delta\,(\text{cooling towers}) \\
& +A_6 \;\; \delta\,(\text{public or cooperative ownership}) \\
& +A_7 \;\; \delta\,(\text{region 1}) \\
& +A_8 \;\; \delta\,(\text{region 2}) \\
& +A_9 \;\; \delta\,(\text{region 4}) \\
& +A_{10} \;\; \delta\,(\text{region 5}) \\
& +A_{11} \;\; \delta\,(\text{region 6}) \\
& +A_{12} \;\; \delta\,(\text{region 7}) \\
& +A_{13} \;\; \delta\,(1965-67) \\
& +A_{14} \;\; \delta\,(1968) \\
& +A_{15} \;\; \delta\,(1969) \\
& +A_{16} \;\; \delta\,(1970) \\
& +A_{17} \;\; \delta\,(1971) \\
& +A_{18} \;\; \delta\,(1972) \\
& +A_{19} \;\; \delta\,(1973) \\
& +A_{20} \;\; \delta\,(1974) \\
& +A_{21} \;\; \delta\,(1976) \\
& +A_{22} \;\; \delta\,(1977) \\
& +A_{23} \;\; \delta\,(<300\,\text{MW})
\end{aligned}
$$

$+A_{24}$ δ (300 – 399 MW)
$+A_{25}$ δ (500 – 599 MW)
$+A_{26}$ δ (600 – 699 MW)
$+A_{27}$ δ (> 700 MW)

The revised regression for construction cost drops the years 1965 – 67, and assumes a constant rate of inflation over the time period 1968 – 77. It assumes a monotonic scale economy. It drops the variables (unconventional construction) and (cooling towers) as not statistically significant.

Formula 2

log (\$/kW) = A_0
$+A_1$ δ (unit 1)
$+A_2$ δ (unit 2)
$+A_3$ δ (oil or gas fuel only)
$+A_4$ δ (public or cooperative ownership)
$+A_5$ δ (region 1)
$+A_6$ δ (region 2)
$+A_7$ δ (region 4)
$+A_8$ δ (region 5)
$+A_9$ δ (region 6)
$+A_{10}$ δ (region 7)
$+A_{11}$ (year of commercial operation)
$+A_{12}$ log (rating)

Each of the 36 units in the heat rate data base entered commercial operation in 1974 or earlier, providing a minimum of 3 yr of experience for each in the statistical record that extends through 1977. Units entering commercial operation from 1975 to 1977 were not included, because their operating experience was judged to be too limited. Since heat rate is reported on an annual basis and varies from year to year, I took the lowest annual heat rate in the record as a measure of the inherent efficiency of the corresponding unit. (If commercial operation began after July of the year, I did not count that year because it could have had more than its share of cold weather.)

The regression for heat rate H assumes that heat rate can depend on rating (the variable of interest for possible scale effects), and on other variables that could mask the influence of rating. The other variables considered are cooling towers, type of fuel, heating value of fuel (if coal), and year of commercial operation. The year of commercial operation is included as a variable because goverment regulations may require progressively more thorough stack-gas cleanup, with resulting increases in heat rate.

Formula 3

$H =$ A_0
$+A_1$ (rating)
$+A_2$ δ (cooling towers)

$+A_3$ δ (oil or gas fuel)
$+A_4$ (*BTU*/lb if coal fuel)
$+A_5$ (year of commercial operation)

REFERENCES

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

Energy Information Administration. 1977. Steam-Electric Plant Construction Cost and Annual Production Expenses 1977. Washington, D.C.: United States Department of Energy. Earlier volumes cover a period of 30 yr.

Fisher, J.C. 1978. Size-Dependent Performance of Subcritical Fossil Generating Units. Palo Alto, California: Electric Power Research Institute.

CHAPTER 5 IMPLICATIONS OF PLANT SCALE IN THE CHEMICAL INDUSTRY WITH PARTICULAR REFERENCE TO ETHYLENE PLANTS

G.G. Betts
BP Chemicals Limited,
London, UK

5.1 WHY ARE ETHYLENE PLANTS LARGE?

The tendency for continuous increase in plant scale is a feature of the high-technology industries producing electrical energy, man-made materials, and transportation facilities, as they approach maturity. Large scale is only one prerequisite for their continuing economic viability and survival. The chemical industry, and ethylene plants in particular, must also look for a continuity of operation, a secure of feedstock (raw materials), a secure, balanced, and minimum level offtake for its products, an economic supply and distribution framework for its raw materials and products, commercial or manufacturing back-up to cover loss of product due to plant breakdowns, access to competitive process technology, and in-depth experience in the specification, design, operation, and maintenance of large-scale units.

Ethylene and its associated C_3/C_4 olefins are the major building blocks for the chemical industry, about 80 percent of whose production is synthetic materials in the form of plastics, fibers, and rubbers. Since the early 1950s, the ethylene produced in Western Europe has been predominantly based on the steam cracking of naphtha. Ethylene and its associated olefins have progressively replaced other chemical building blocks such as acetylene based on coal or coke, or ethanol produced by the fermentation of biomass. The consumption of ethylene in Western Europe over the 25-yr period 1955 – 1980 has increased by a factor of 12 and the total installed capacity of ethylene plants has increased by a factor of about 15 (Figure 5.1). BP Chemical and its associated companies in common with other early entrants to the petrochemical industry, has built a succession of ethylene plants of increasing capacity (Table 5.1) since the 1950s.

The high growth rate of the production of organic chemicals both world wide and in the UK during the period up to 1973 (Figure 5.2 and 5.3) was encouraged by declining ethylene prices made possible by a combination of lower feedstock costs, technical improvements in steam cracking permitting higher yields of the primary product ethylene, and hence lower feedstock-to-ethylene ratios, and the economies of plant scale. This state of affairs came to an abrupt end following the OPEC oil crisis at the end of 1973. The effect on the selling price of ethylene has been such that there was in real terms a decrease up until 1973 which was then sharply reversed (Figure 5.4); with the recent Iranian oil crisis the price is likely to continue to rise.

TABLE 5.1 Ethylene plants of BP chemicals.

	Date Commissioned	Capacity (tpa) (approximate) T/A
UNITED KINGDOM		
Grangemouth (Scotland)		
No. 1	1952	30,000
2	1957	30,000
3	1961	75,000
4	1967	250,000
Baglan (Wales)		
No. 1	1964	50,000
2	1973	340,000
Wilton (50/50 ICI/EP)		
No. 6	due 1979	500,000
FRG (Erdölchemie)		
Dormagen (50/50 BP/Bayer)		
No. 1	1957	20,000
2	1960	40,000
3	1963	85,000
4	1970	360,000
5	1978	360,000
FRANCE (Naphtachemie)		
Lavera (43/57 BP/Rhone-Poulenc)		
No. 1	1953	20,000
2	1960	35,000
3	1966	135,000
4	1972	480,000

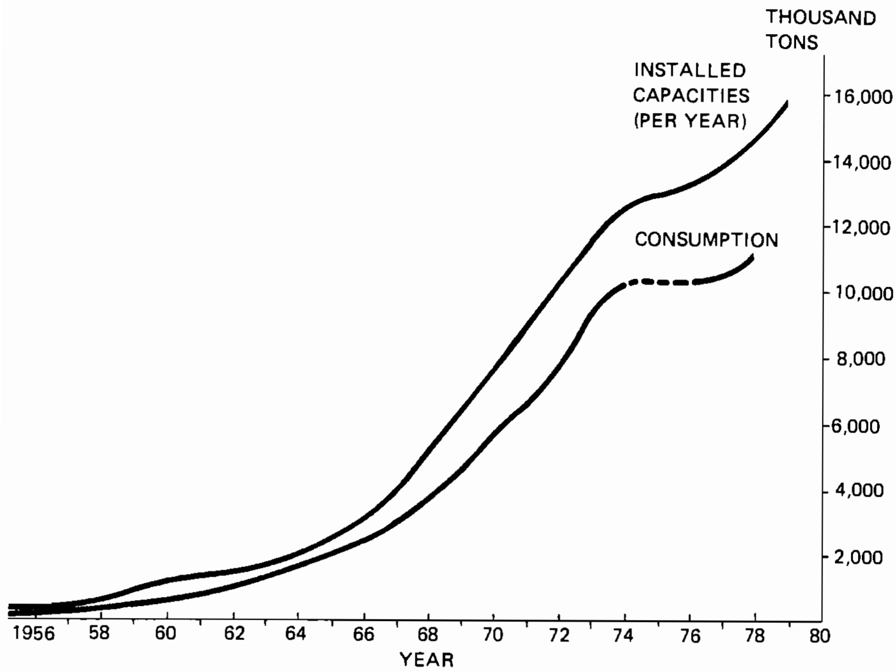


FIGURE 5.1 Ethylene plant capacities and ethylene consumption in Western Europe.

5.2 WHAT MAKES AN ETHYLENE PLANT A LARGE PLANT?

In the early 1950s, ethylene plants typically had a capacity of around 30,000 tpa (metric tons per year). Now plants of 500,000 tpa are not unusual, and the largest single-train plant being built will have a capacity of 680,000 tpa. The throughput of process materials may approach 2 million tpa. This serves to emphasize that what constitutes a large plant is a dynamic and not a static concept, and what is regarded as world scale capacity is continualy increasing. What then are the characteristics of an ethylene plant that make it a large plant?

To build a 500,000-tpa ethylene plant including its gasoline treater unit might require a capital investment in excess of £ 150 million, excluding its associated offsite and supporting infrastructure.

The process complexity of a typical world scale 500,000-tpa ethylene plant is indicated by its facilities: it might contain about 20 distillation towers, 150 drums or pressure vessels, 200 heat exchangers, between 3 and 8 major compressors and turbines and, say, 5 to 8 pairs of cracking furnaces, together with pumps, smaller compressors, start-up boilers, flares, cooling towers, and other ancillar devices.

An indication of the control complexity of a typical world scale ethylene plant is its, perhaps, 600 control loops, possibly 50 or more on-stream analyzers, 1,800 trip initiating devices, in excess of 1,000 temperature measurements, a control panel 70 m in length, and an on-line computer with possibly 1,500 inputs controlling key process equipment and providing management data.

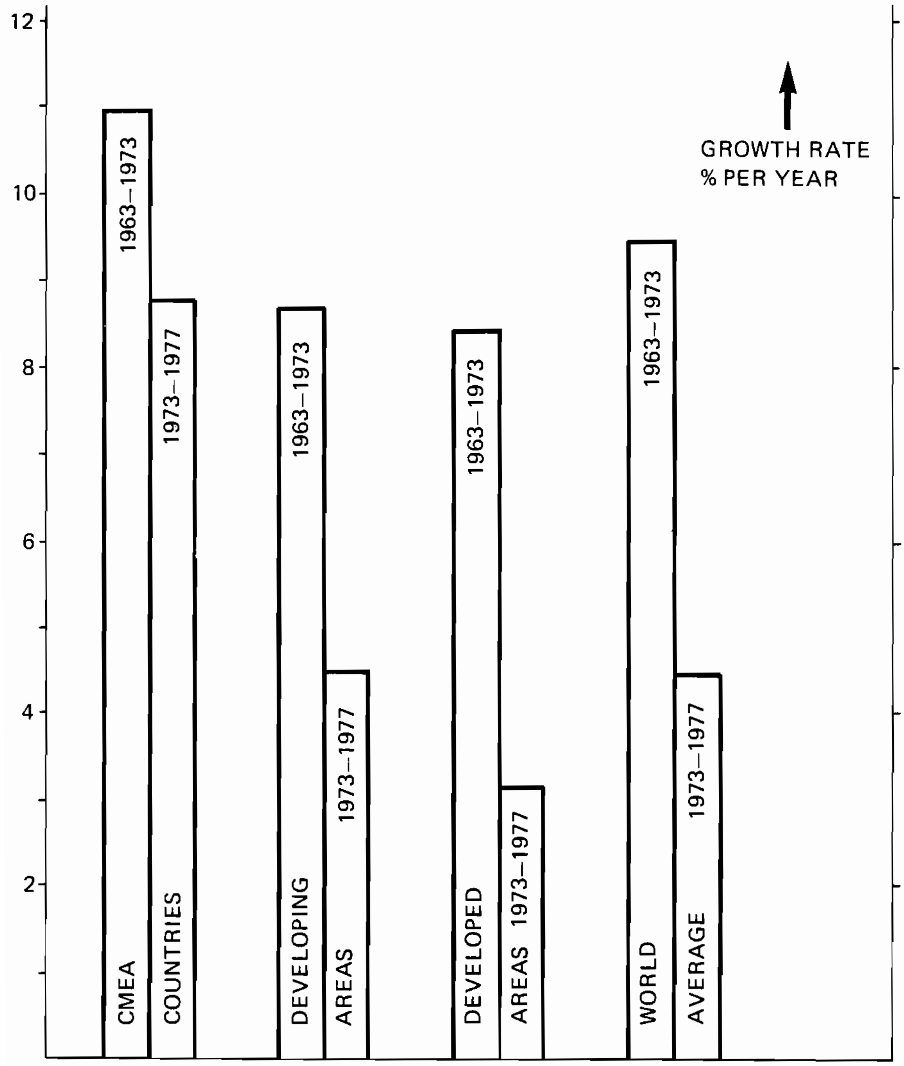


FIGURE 5.2 Growth in world chemicals production: 1963–1973 and 1973–1977.

An indication of the physical size of large ethylene plants is the plot area: a typical 500,000 tpa plant might be about 80,000 m^2 (20 acres). This is six to seven times larger than the area required by an ethylene plant built about 30 yr ago with a capacity of 30,000 tpa. The largest column might be up to 10 m in diameter and have an overall length of 100 m; each cracking furnace might contain a total length of about 600 m of alloy steel tubing, and the total plant compressor horsepower might be in excess of 100,000.

An indication of the construction effort required to build a world scale plant is that the total manpower required would be in excess of 5 million man hr, and the peak labor force in excess of 1,500 men.

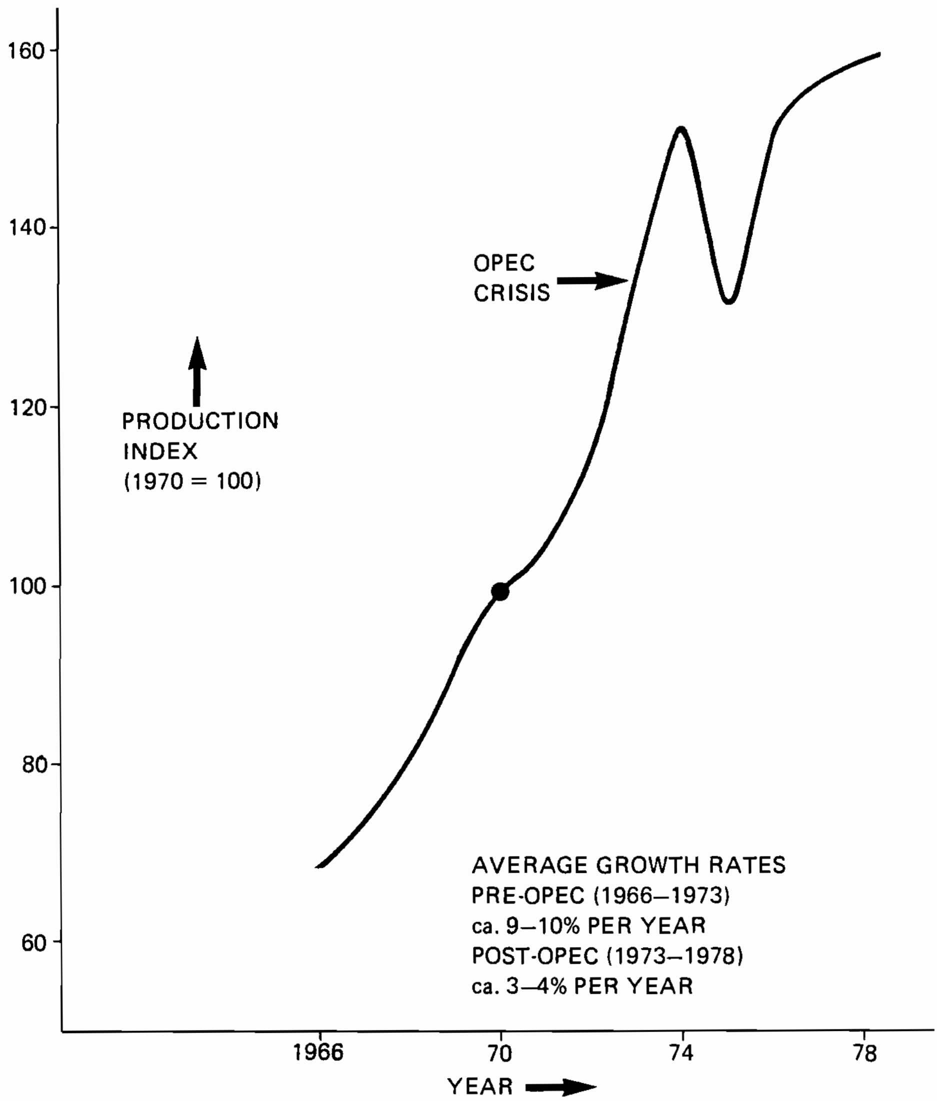


FIGURE 5.3 UK organic chemicals production: 1966 – 1978.

5.3 TO WHAT EXTENT DO ETHYLENE PLANTS EXEMPLIFY THE PROBLEMS OF PLANT SCALE?

By most criteria, such as throughput, capital investment, process and control complexity, physical size, and construction effort, ethylene plants clearly exemplify large plants. In the chemical industry however, they are by no means unique. The construction of plants of ever-increasing size for the manufacture of ethylene and other commodity products has been a general feature of the chemical industry for the past 30 yr.

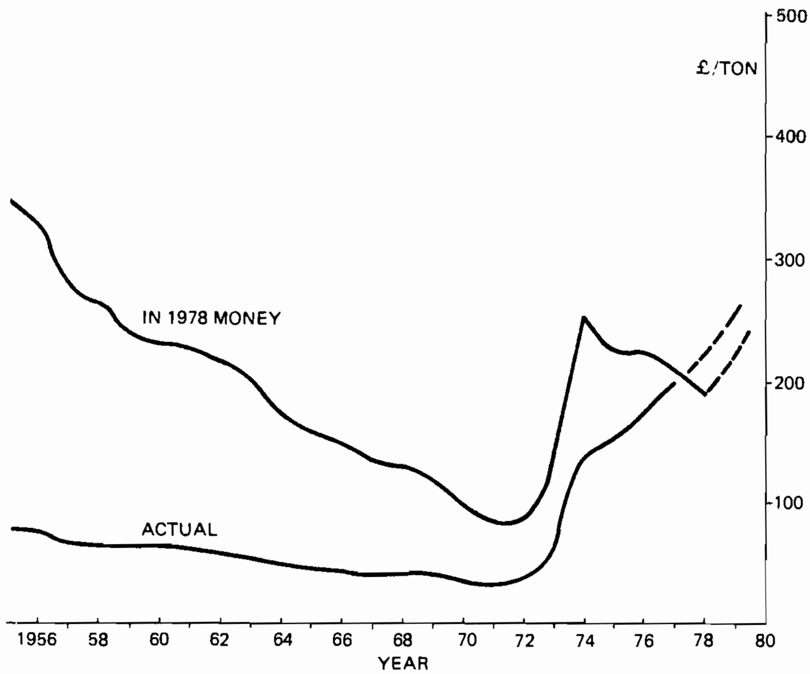


FIGURE 5.4 Ethylene selling prices: actual and in 1978 money terms.

Ethylene plants are not, however, single-train plants in the strict sense of the term. The steam pyrolysis of hydrocarbons gives rise to the formation of carbon and this requires the cracking furnaces to be taken off-line periodically for decoking. To maintain continuity of output, multiple furnances have to be installed.

Other examples of large-capacity plants are ammonia and methanol plants; the largest ones currently being constructed are about 510,000 and 825,000-tpa capacity, respectively. These plants are usually single-train throughout, thereby attaining the maximum economies of scale. At the other extreme, large-scale plants for the manufacture of, for example, chlorine from brine, largely consist of multiple electrolytic cells and decomposers and hence the economies of scale are not as great.

In contrast to power stations and chemical plants producing and methanol, ethylene plants do not produce one product.

The principal products from a naphtha-based plant are, by weight,

Ethylene	29 % (with ethane recycle)
Propylene	14 %
Butadiene	4 %
Fuel oil	5 %
Methane	16 %
Gasoline	24 %
Hydrogen/C_4s	8 %

Equipment and control systems are required to separate and purify these products. Ethylene and propylene have low boiling points, – 104°C and – 48°C, respectively, at atmospheric pressure, and compression and cryogenic systems are required to separate the various olefins and to handle them in the liquid form.

Because ethylene plant feedstocks such as naphtha or gas oil are refinery products, about a third of which is returned as gasoline and fuel oil, these plants tend to be located adjacent to refineries. To avoid transportation costs, ethylene plants often have built around them downstream plants to convert the ethylene, propylene, and C_4 fractions into chemical commodity products. Also, to protect and support the upstream investment it becomes increasingly desirable for the downstream user plants to be of large capacities. Ethylene plants differ in this respect from, for example, power stations, which do not have satellite units or extensive tankage.

Ethylene technology is readily available from a number of major engineering contractors, rather than being dependent on the grant of a licence from a manufacturer. However, because the initial investment in large ethylene plants and their minimum economic base-load is high, it is not an attractive business area for new entrants because entry fees such as these are high. Nevertheless, new entrants particularly in developing countries may justify the construction of world scale plants on the premise of obtaining a sizable proportion of export business.

Ethylene plants for a number of reasons are rarely built as exact replicates of one another, and each plant has its own specific features with respect to feedstock range, product quality, local site conditions, and so on, and only limited scope for offsite prefabrication under factory conditions. In their design and constructional aspects therefore, ethylene plants differ from the aircraft industry, which has a production line product manufactured under controlled workshop conditions by skilled workers, who learn from experience and who have the motivation of continuity of employment.

5.4 WHAT ARE THE PRESSURES PROMOTING INCREASE IN PLANT SCALE?

The perceived incentive for building large manufacturing units such as ethylene plants is economic, namely, to take advantage of the reduced investment and fixed operating costs per ton of product. The capital cost of an ethylene plant is believed to be influenced by scale, (Figure 5.5) and this together with the other changes in input ratios improves ethylene production costs and potential profitability given a fixed selling price.

Nevertheless, the capital cost of ethylene plants has in real terms been increasing since about 1968, and has more than canceled out the benefits of scale. The other factors that have increased capital costs are design changes necessary to increase the on-stream reliability of the plant, to improve its safety and environmental performance, and to provide the flexibility to process a wider range of naphthas; in the UK there has also been the effect of decreasing construction labor productivity. These have all had the effect of increasing the cost of ethylene plants in real terms from 1968 onwards (Figure 5.7).

During periods of manufacturing over-capacity, the prices of commodity chemicals may be barely adequate to give a normal return even on historic investment. Particularly during an era of escalating capital costs, a manufacturer who is planning new ethylene capacity to meet future demand will often have difficulty in forecasting an adequate

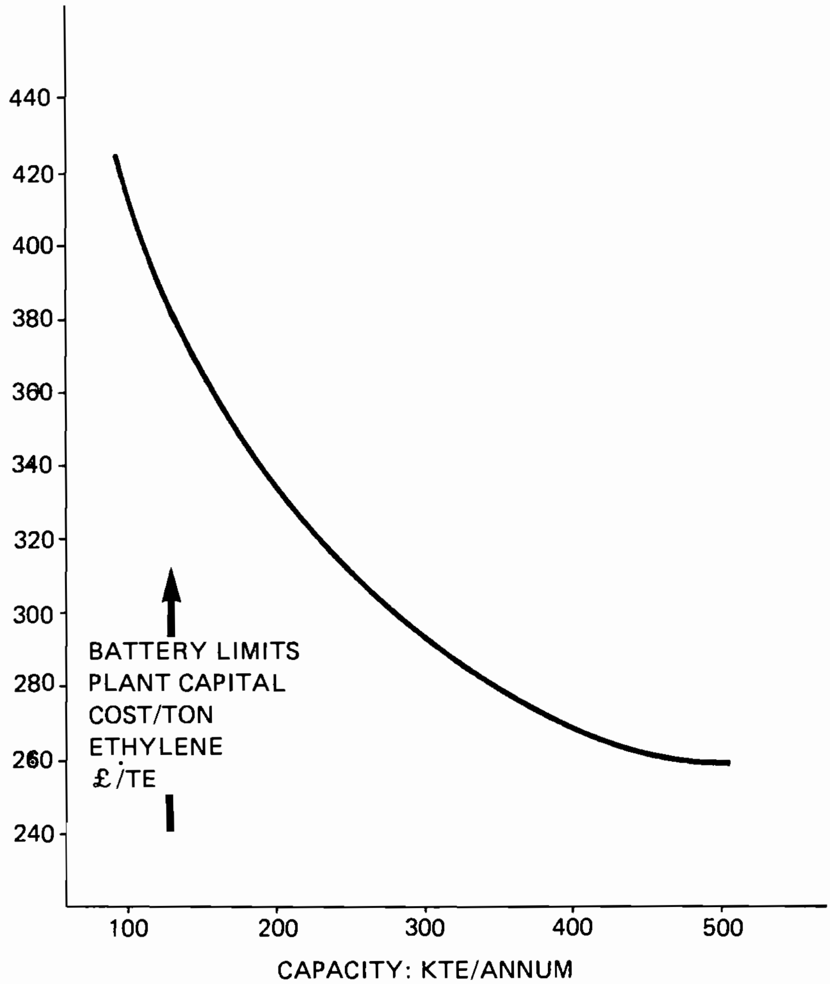


FIGURE 5.5 Ethylene: 1979 capital costs (mid-1979 cost basis; high-severity cracker (25 – 30 percent ethylene yield) plus associated gasoline treatment and butadiene extraction units; feed stock: naphtha). KTE stands for kilotonnes ethylene.

return on the new investment. Unless the manufacturer is in some special or protected position, other than not to buid at all, there is often little choice other than to build the new plant of at least current world-scale capacity.

If the cost and price of ethylene and other olefins can be maintained at a lower level in real terms, for example, by using the economies of scale, then this may well help to generate additional uses for ethylene and other olefins and/or lead to a further displacement of natural materials. In recent years, however, the price of naphtha has had the predominant effect on ethylene price (Figure 5.8) and not plant scale, and this trend is likely to continue.

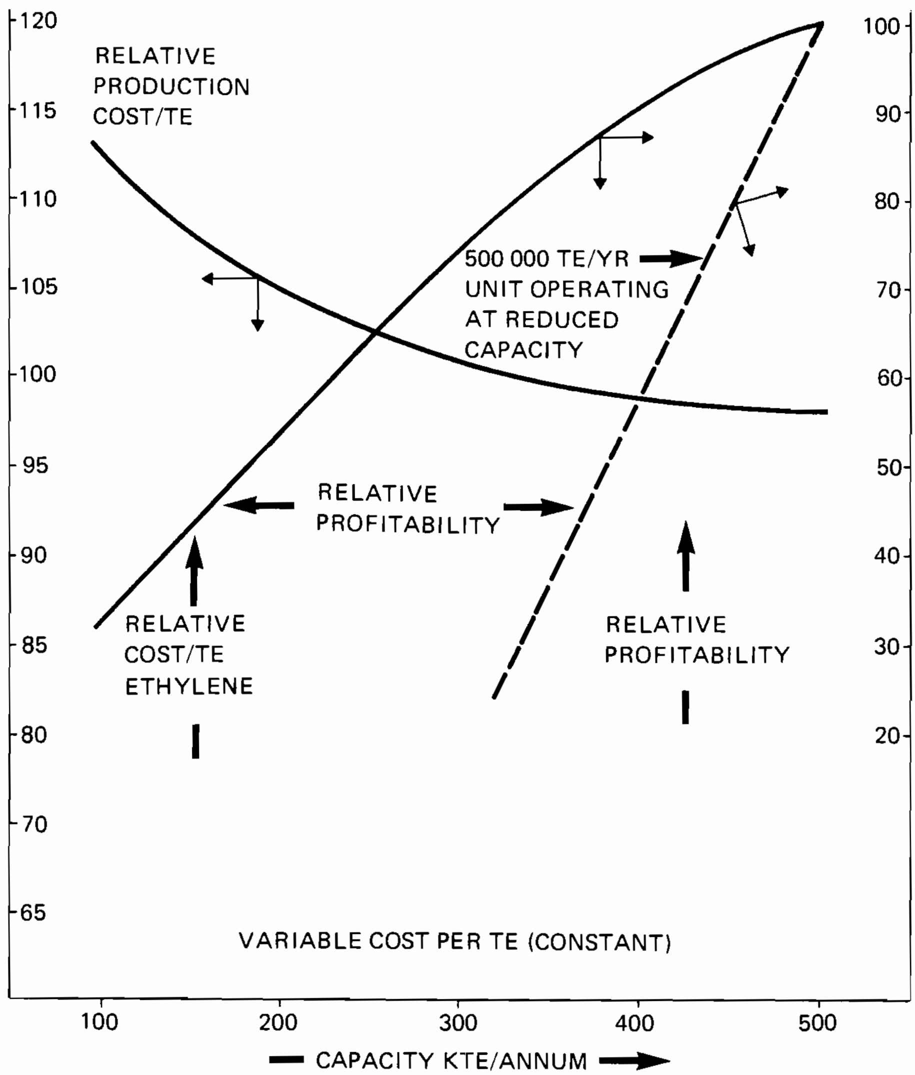


FIGURE 5.6 Ethylene: 1979 production economics. KTE stands for kilotonnes ethylene.

Technological development, both process and mechanical, encourages the construction and operation of new larger-capacity plants. On the process side, the development of higher cracking severities in the 1960s encouraged the shutdown of older and smaller ethylene plants designed to crack at low severities. The incorporation of the capacity shutdown into new plants could be justified economically by the saving in variable and fixed operating costs. Mechanical developments such as high-capacity centrifugal compressors made it possible to build plants of larger throughputs, and improved metallurgy made possible the use of higher temperatures in furnace tubes to attain higher cracking severities; the recovery of heat by steam generation at a high enough level permits turbine drive of the compressors at high thermal efficiencies.

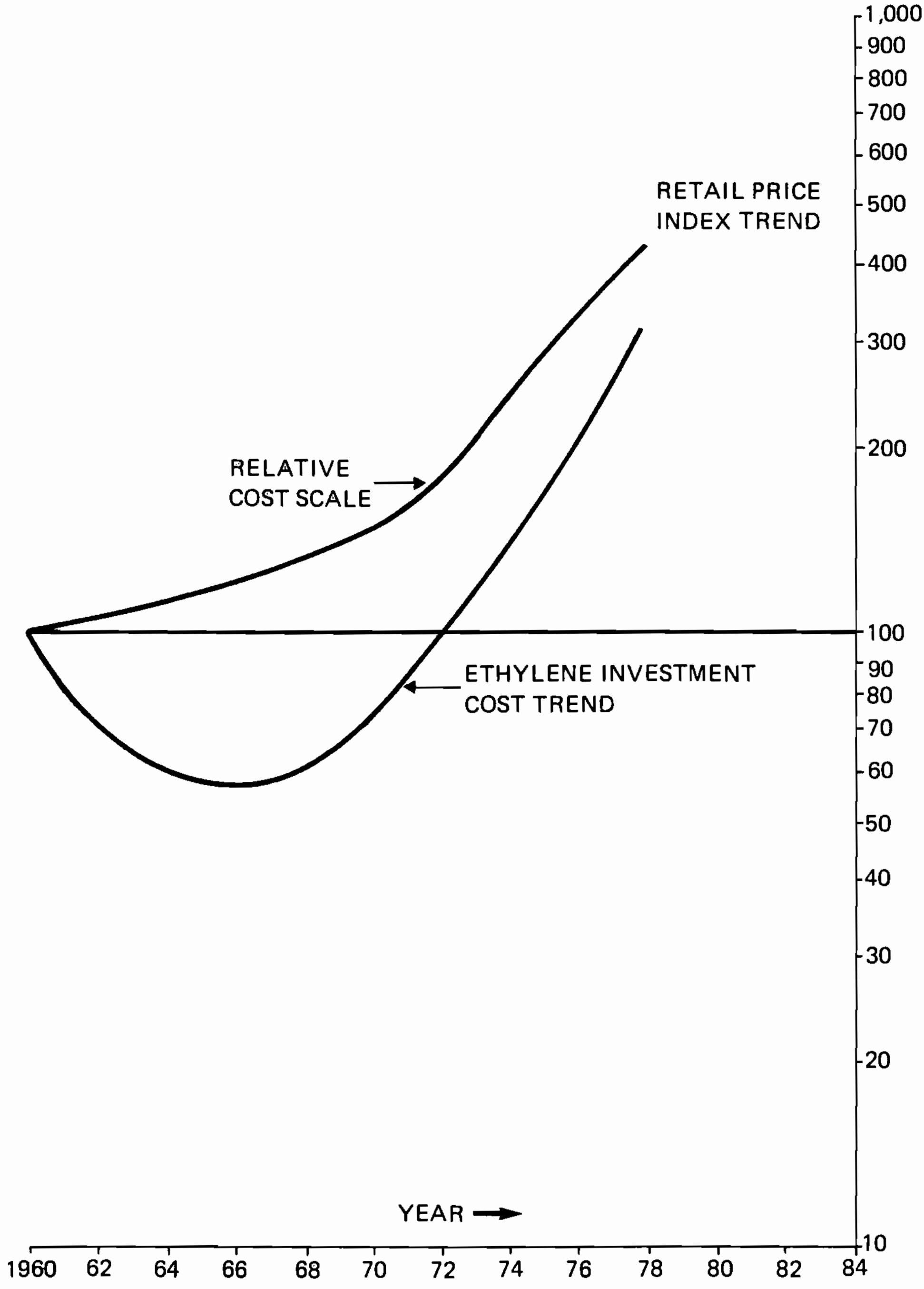


FIGURE 5.7 Trend of retail price index and ethylene investment cost over the period 1960 – 1978.

Even though fewer than the expected benefits, for example, in terms of reduced capital costs per unit of product, may have been experienced in other or earlier large-capacity plants, there is a tendency to believe that the lessons have been learned and are understood and will not be repeated on the next project. This is only to be expected for example in the chemical industry which has a successful history based on change and willingness to adopt new technology. Experience undoubtedly does enable both engineering contractors and manufacturers to move along the learning curve and to build up confidence.

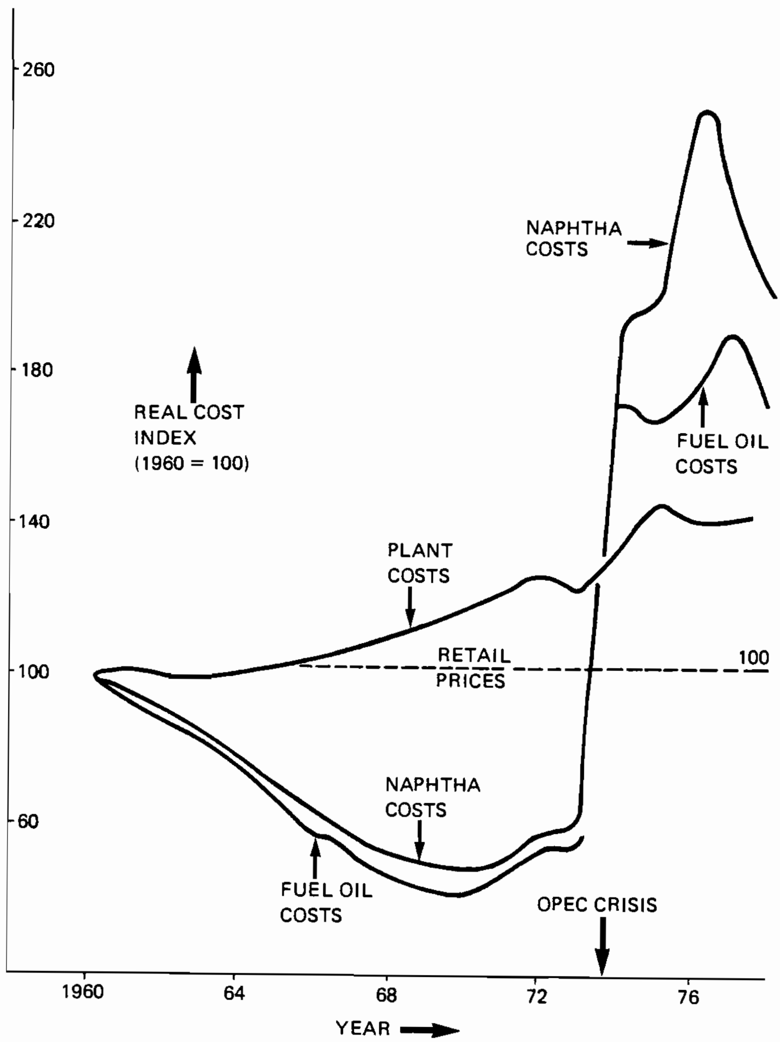


FIGURE 5.8 UK cost trends in real terms: 1960–1978.

An economic climate of rapid capital cost escalation has the effect of making the capital investment per ton of product in large-capacity plants appear in historic accounting terms to be cheaper, and this encourages the tendency to build in anticipation of demand. Governmental tax allowances and investment grants are often related to new capital investment and this promotes the installation of capital intensive projects.

The ownership of plants by the financial mechanism of asset sharing or jointly-owned companies enables manufacturers to aggregate their ethylene requirements and to build a large plant with the economies of scale without the necessity of phasing their projects, provided that the benefits are not negated by additional logistic costs.

Institutional reactions that manifest themselves in the difficulties, delays, and uncertainties associated with obtaining planning consents for green field sites in new areas, particularly for chemical manufacture, have the effect of concentrating manufacture in existing sites and encourage the building of large-scale plants.

5.5 WHAT ARE THE PRESSURES ACTING AGAINST INCREASE IN PLANT SCALE?

When the capacity of ethylene plants exceeds around 300,000 tpa, it is no longer possible to transport some of the larger distillation columns from the equipment fabricator to the site because of their diameter or height. Such columns either have to be built by site fabrication as a single unit, or shop-fabricated twin units have to be installed. The size of the compressors may be such that they have to be split and built in series. Even if the site is close to a harbor that permits sea transport of the equipment, its physical size or weight may require special widening of roads and strengthening of jetties and bridges, and so on.

The perceived economies of scale in building a large ethylene plant are to some extent offset by the cost of transporting the olefins or downstream products from an integrated multiplant site to the market. When downstream user plants are built adjacent to the ethylene plant, the movement of large tonages of products can place a serious strain and require capital expenditure on dock/harbor/rail/road facilities, that may eventually reach a physical limit. Insofar as the movement of ethylene itself is concerned, this has been simplified by the development of gas pipeline grid systems such as in Texas and the Rotterdam/Antwerp/Ruhr areas.

The massive investment required to build large-scale plants can give rise to financing requirements necessitating high debt-to-equity ratios and overseas borrowings. Adequate insurance both of the investment and consequential loss of production are also important considerations.

A large-scale plant requires a high base-load to be economical. For example, the relative profitability of a 250,000-tpa ethylene plant operating at full load is the same as a 500,000-tpa plant operating at about 80 percent of its capacity (Figure 5.6). A firm requires secure long-term commercial contracts that must usually be negotiated in a highly competitive situation. It should be borne in mind that although the economies of scale suggest a reduced investment per unit of additional capacity, the commercial risk associated with top-end sales is often higher.

The large single-train plant gives rise to concern as to its reliability in terms of on-stream days per year. This concern is less if the larger plant is supplying ethylene into a multiproducer grid system or if older possibly smaller ethylene plants exist at the same site and can be kept on standby. If the large-scale plant stands in isolation, great attention will have to be paid and additional capital spent on the installed sparing of appropriate items of equipment, liberal warehouse sparing, back-up facilities, larger product storage, avoidance of innovation, more extensive plant monitoring, and so on.

Large-scale plants require a large site construction force. It is now recognized by the industry that this leads to lower labor productivity, poorer control, lower quality of labor, and an increased tendency for industrial disputes.

Large-scale plants inevitably give rise to higher quantities of gaseous and liquid effluents discharged both continuously and under upset conditions. Planning consents are often more difficult to obtain and may require significant additional capital expenditure to reduce these discharges. Similar considerations apply to the hazard analysis, for example, the containment of higher quantities of flammable materials kept as in-plant inventories or as stocks of feedstock and products.

Even if society in the industrially developed countries is able to accept the overall social benefit of large-scale manufacture at least at a national level, this does not necessarily influence local opinion in the area where the large-scale plant is to be constructed and operated, where it may be opposed. Similarly, although the argument that manufacture should be concentrated into a few large-scale plants may have economic merit nationally, this is less likely to be acceptable if it means shutting down smaller regionally dispersed manufacturing units.

5.6 WHAT ARE THE PROBLEMS AND PENALTIES POTENTIALLY ASSOCIATED WITH INCREASE IN PLANT SCALE?

It should be recognized that the design and construction of chemical plants of a size significantly larger than the current norm in itself represents an innovative step for both the engineering contractor and the operating company, always from the engineering and constructional aspects, and sometimes from the process and operational aspects. Many of the problems inherent in plant scale and the potential discontinuities of the economies of scale are difficult to foresee. Once the commitment has been made, corrective action may be too late or not completely effective.

It is in practice virtually impossible to differentiate to what extent penalties particularly in capital cost arise solely owing to increase in plant scale, and not to other changes and other interactive factors introduced or occuring at the same time.

Although the number of ethylene plants world-wide is large, very few plants have been built within a sufficiently closely defined geographical area, within a sufficiently short time span, and under identical conditions to make meaningful cost analysis comparisons. Also, the chemical industry, which operates in an internationally competitive market, is understandably unwilling to disclose its confidential financial data in the detail necessary for a meaningful comparative analysis.

One way to avoid the risk of scale is to duplicate or even exactly replicate existing plants, or to introduce only small step changes. This assumes that the problems experienced and solved will be avoided. However, if as is now the case, the rate of increase in demand for ethylene is declining, new plants will be built with less frequency. This, together with the changing feedstock situation, makes changes to the design increasingly probable (particularly to handle feedstock flexibility, increasing energy costs, or that the same equipment is no longer available), even though plant capacities may not continue to increase significantly.

Some of the problems in increasing plant scale are self-evident. The increased physical size of equipment and piping affects their transportation, lifting, wall thicknesses, and length of welding run. Pipe diameters can reach the size where wall thicknesse, thermal expansion, and stressing consideration require a much more sophisticated design of pipe supports. All of these factors have an obvious influence on capital cost. Clearly, the larger the physical size of equipment and alloy pipework the more restricted is the choice of fabricators because of the increasingly nonstandard nature of the equipment, and this can result in unexpected delay in delivery to the site. Although large construction sites can be divided into smaller areas, so that equipment and piping may be located in several such areas, the equipment and piping to be built and tested as complete systems and sectionalization will not be wholly effective.

Experience seems to indicate that estimates, made at the time of financial sanction of the project, of the cost of the process equipment items such as towers, vessels, exchangers, and so on are in the event reasonably reliable. Also, until the physical size of items reaches the dimensions at which they require twinning or site fabrication, the economies of scale seem in practice to be realized roughly in accordance with the 0.7 factor rule. There is however a tendency to underestimate the materials accounts that cover piping, insulation and painting, structures, buildings, foundations, sewers, fireproofing, electrics, and instrumentation. The piping and instrumentation accounts in particular can be seriously underestimated. The cost of the materials does not have at all the same favorable relationship with plant scale as does the cost of process equipment. The reason is that until the design of the plant has been developed to the stage of piping layouts and materials takeoffs, the estimates of the quantities required are based on experience factors related to the process equipment accounts. This is certainly a convenient but possibly questionable assumption. Even if these factors are believed by the engineering contractors, who usually develop the initial estimates, to be based on a similar plant, the particular local conditions, design requirements, and practices of clients can give rise to quite different results. It should always be recognized that capital cost estimating is not an exact science but is largely based on past experience plus a great deal of engineering and cost interpretation, and on the state of the supply, fabrication, and labor market.

The essential question is, however, to what extent and why does increase in plant scale reduce the expected economies in the actual cost of these materials accounts. The answer no doubt lies in the fact that large equipment and piping requires ample access for construction, maintenance, and operational purposes. This results in a more liberal layout of the plant plot area. The cost of most of the materials accounts is directly affected by equipment layout and plot size. One indirect but important effect of low values in the initial capital cost estimate of the plant for the materials accounts is that this can result in too low a level of bulk ordering of materials or reserving of shop fabrication space. Ultimately, the site construction activity will be adversely affected if there are insufficient or late deliveries of materials.

The high proportionate cost of instrumentation is no doubt due to the increased need to protect both equipment and personnel, to help to approach more closely optimal operation and higher on-stream time of the plant, and to monitor carefully its environmental performance. These factors require more and more data from the operating unit and more sophisticated control, which increases the cost. The increased cost can be more easily justified economically as the capital investment and cash flow through the plant become larger.

The element of capital cost most likely to overrun the early estimates is the construction manpower. The smaller ethylene plants built in the UK say 20 yr ago had a site construction cost element of the order of 10 – 12 percent of their total investment. Larger plants built around 10 yr ago had a site construction cost at around 25 – 30 percent of the total. Now world-scale plants may have a site construction cost of say 40 percent of the total and take 5 yr to build. These percentages will to some extent, however, be artificially inflated because the cost of construction labor in money-of-the-day terms is disproportionately affected by escalation since it is incurred towards the end of the project schedule.

A number of explanations have been advanced for the high proportion of construction man hours and low productivity. Again, the real question is to what extent can this be solely attributed to plant scale and is not also due to a changing social environment, different climatic conditions, local industrial relations and the influence of taxation and other financial incentives, and changes and modifications to the original design and specification of the plant to incorporate more recent operational experience.

There is a body of opinion that if a site work force peaks to more than about 500 men, productivity begins to decline. Another theory is that although understandably the plant owner wishes to see site work commence as soon as possible, better planning of the construction activity and labor productivity would result if work commenced in the field later, when the bulk of materials had been delivered to the site and a clear run of work was available. Any factor that delays or interrupts the smooth flow of site construction work causes a significant increase in the overall project schedule, which adversely affects this important element of capital cost. This is in addition to the extended period of financing, cost of site supervision, and site overheads such as the hiring of major construction aids. Some of these upset factors have already been mentioned, for example, late design changes, late deliveries of equipment and materials, inadequate storage facilities, corrective engineering, reinspection, and retesting.

Clearly, as the size of the construction work force grows so does the problem of attracting the right quality of tradesmen. Poor quality work requires more inspection and corrective engineering, which means more cost and delay. With ethylene plants, the problem is often exacerbated by the simultaneous construction of user plants, probably by multiple contractors at the same large construction site. Studies have been published of other industries that seek to relate strike proneness to size of work force. The probability increases when more than about 1,000 people are involved, which may be some sociologial threshhold related to the number of communication links. With regard to management, large construction sites call for the most experience and skills, but it would be a mistake to assume that such experience and skills are completely transferable or relevant to differing locations in the world.

The building of large plants creates peak loads on the equipment/materials/fabrication industry; sometimes there is limited choice of suppliers; there are peak demands on skilled construction labor and on the engineering contractors; and there are peak site engineering demands on the manufacturer to achieve quick plant maintenance turnarounds. In the engineering contracting industry a large plant requires the employment of temporary agency manpower at a higher cost; and with a quality of work that makes it more difficult to supervise. It also makes it essential for the contractor to seek and promote business on a world-wide basis, which is highly competitive, in an attempt to even out the extremes of work load on his home office.

The effect of plant scale on operating staff is an important factor affecting productive efficiency. Any physically large equipment erected on a relatively small area initially has an overpowering effect on the human mind, particularly if the equipment is rotating and creating noise or contains flammable or toxic materials. However, provided that there is a period of familiarization and training, this quickly disappears and confidence is gained. The high risk inherent both in terms of the investment employed and loss of sales revenue due to a prolonged plant shutdown requires a very high level of plant supervision and operation. Fewer decisions can be left to chance or to spontaneous individual reactions. Each foreseeable operational occurrence has to be covered by standardized routines and procedures. This can narrow the breadth of a job, which requires careful attention so that individual job satisfaction and motivation are not adversely affected. Recently, one of the ethylene plant contractors surveyed the time taken to start up a a number of ethylene plants by the oversimplified but convenient criterion of "oil in" to "production of specification product." The data indicated that the more experienced the operating supervision the more the likelihood of a shorter start-up. Although there was no direct relationship with plant scale, the large-scale ethylene plants were owned by very experienced manufacturers with, for example, carefully prepared training and familiarization programs.

The presentation of the vast amount of control information has to be carefully considered, particularly with respect to the correct grouping and emphasis of emergency displays, and industrial psychologists have been employed to assist in such problems. Incidentally, industrial psychologists have also been used to build teams and to promote good communication between the staff of engineering contractors and the client's personnel, some of whom work in the contractor's office during the design phase of large-scale plants.

It is perhaps unfortunate and misleading that large-scale plants are usually characterized as capital intensive, whereas in reality they are manpower intensive. A large manufacturing plant is the product of an immense input of technical, engineering, craft, and managerial skills. If the original expectations of a more productive efficiency of capital investment and operating economics are to be achieved in practive, then this requires the coordination and motivation of many diverse skills and interests at different stages of the project towards the overall objective, which manifests itself in the operating plant.

5.7 CONCLUSION

There may be a tendency to persist with a too simplistic and possibly outdated approach to the economies and productive efficiency of plant scale, and hence there is a need for some of the currently perceived beliefs at least to be challenged. The incentive to do this lies principally with the manufacturers, because they bear the ultimate penalties (or benefits) of the robustness of their investment assumptions and decisions. However this presentation has suggested that any analysis of the implications of plant scale is far from simple. No doubt a great deal will always have to be left to informed judgment and technical experience applied to the particular project and particular circumstance under consideration.

CHAPTER 6 SCALE, TECHNOLOGY, AND THE LEARNING CURVE

K. Tsuji
International Institute for Applied Systems Analysis,
Laxenburg, Austria

The learning curve, which represents the relation between accumulated production and the cost per unit of production (or a similar index), was used in several papers at the workshop and stimulated lively discussions. On one side, the usefulness of the concept was questioned, and on the other side, the relation was treated as a natural law. Also, the implications and the nature of "learning" and their relations to the changes in scale and technology were discussed.

This chapter will summarize the important points of the discussions and attempt to clarify the usefulness and limits of the learning curve concept based on a literature survey and on a conceptual, hierarchical model. For a basic understanding of learning curves, the reader is referred to Yealle (1979 b).

6.1 DISCUSSION SUMMARY

The fuzziness of the learning curve concept and the learning itself is partly due to the word "learning," which can arbitrarily refer to individual, managerial, organizational, and societal learning; without a specification of the subjects and the objects, the meaning of the world becomes rather obscure. This was most clearly expressed by Gold, who asked, "who learns what."

In particular, Gold expressed his doubts about the usefulness of the learning curve for the blast furnaces (Figure 6.1) presented at the workshop by Derkx and Kamerman (Derkx *et al.* 1978).

Gold: I am very much troubled by the use of the term learning curve. Do we use the term learning curve to cover everything: improvements in technology, the learning of people how to do the same job better, changes in the nature of the product, changes in the nature of inputs, feedstocks and stuff of these sort? Is that what the learning curve is? In which case we don't have to use "technology" any more. We can simply say that everything represents the learning curve. A learning curve drawn for a blast furnance from 1896 up to now necessarily reflects the changes in furnance technology as well as in the qualitative characteristics of its inputs and outputs.

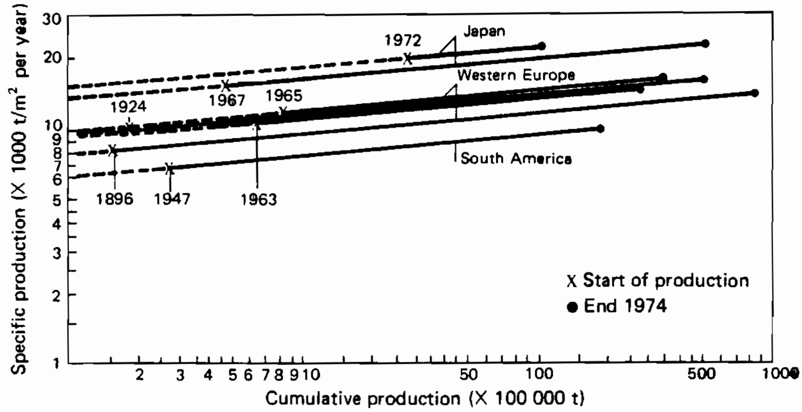


FIGURE 6.1 Graph of the effect of dynamic scale economies in steel plants, showing equal relative increases in productivity for all plants. Source: Derkx *et al.* (1978).

Derkx: Learning is the basic experience of labor inside a plant. They learn to do better. But this experience also results in overcoming bottlenecks in the facilities. Experience not only has the effect of doing the same job better, but results of experience can also be incorporated in investments to improve existing facilities or when constructing new facilities. A good example for me is the increasing hearth diameter of blast furnaces. In the 1960s the Japanese tried to increase the hearth area and then the next bottleneck was the cooling capacity. Thus technological development is going step by step from bottleneck to bottleneck based on the experience with the facility. In this respect our curve could be considered as the envelope of step-by-step technological development.

Gold: The increase of hearth diameter in Japanese blast furnances was due to a basic change in the theory. The Japanese developed a new theory. The original theory about the combustion being confined to the core of the blast furnace led to the decisions of the American industry that there would to be no gain from increasing the hearth area. With the change in basic theory, the Japanese were able to jump up from 1,000-ton to 10,000-ton-per-day blast furnaces (Boylan 1975, Gold 1974). Now this represents a change in technology and theory rather than merely eliminating bottlenecks; this involves developing some new ways of doing things. I do not think that the learning curve concept is quite discriminating enough to differentiate among the variety of factors associated with scale effects.

Derkx: We are talking about specific production. We have compared the big Japanese furnance of 10,000 tons per day with the specific production of smaller blast furnances. This is a way to compare technological content of the facility, the specific production, and the way it goes, better or not. This is the reason why we compare specific production.

A reaction to this argument was:

Schenk: In summary learning means different kinds of changing technology, different kinds of changes in society, etc. Even with large jumps in technology, there is a stable pattern of learning in different industries.

Sahal: In fact, learning involves a multitude of activities (e.g., material handling procedures, plant layout, scheduling of labor, etc.) and, therefore, it is conditional upon the participation of a great many agents in the production process. The theoretical basis of the learning curve is straightforward: it is isomorphic to the Pareto distribution (see Sahal 1979b). The learning curve is far from a trivial regularity . . . it is one of the most important law-like relationships to have been observed in this area.

In his presentation later on, Sahal claimed that the accumulated experience will determine the maximum size of a plant and also the technological change (Sahal 1979a, c, d). In other words, it is not only that improvement in the efficiency of the existing production processes is conditional upon learning or accumulation of relevant experience. Learning also determines the scope for development of new techniques. He demonstrated this by the results from the analysis of historical data on advances in the efficiency of computational devices, farm tractors, turbogenerators, and transportation equipment. The strong emphasis on the role of learning brought up some counter-arguments.

Van Dalen: You've suggested that the general trend toward improved efficiency in certain sectors, or the "volume scale" (the more you produce, the more efficient you get), is largely due to a learning process within the industry. Couldn't this also demonstrate the drift of technological improvement in the world as a whole? The improved efficiency of tractors, for example, could depend on the development of the ball-bearing industry, or the development of lightweight metals or a new kind of synthetic rubber. These products were not designed specifically for tractors but were developments in general. The tractor manufacturer does not learn from his own business, but benefits from technological drift in the world.

Learning has too much of a passive connotation and doesn't take human ingenuity or inventiveness into account. I have observed that there are a host of ideas available for implementation, but they are considered by managers to be impractical or too risky for business. Things are not hampered by technological ingenuity, but by an apprehension of economic considerations. The idea of learning from operations should have less emphasis than has been brought out so far.

Rosegger: Mr Cantley started with a taxonomy that was more than ornamental, which I hoped would sort out size effects from experience or time effects. I am uncomfortable with measures of learning that suggest there are prolonged periods of "forgetting" with respect to a particular technology. I am not sure what one does statistically, but conceptually we can do better than to say those are periods when we learned from errors only, which is the implication of the measures when the rate of learning is negative. We must look for the underlying causes, not subsume them.

Rochlin: The fuel economy of automobiles shows no learning curve. It shows a correlation with the price of fuel relative to other goods, but not a correlation to the number of units produced.

External factors may in fact be the most important ones. One of the driving forces in tractor innovation was that people were killed because tractors turned over. Perhaps a better example is the three-point tractor hitch; this was largely a consequence of the shift toward urbanization, which made the size of the average farm increase greatly. Now, to say that the three-point hitch caused something is to turn the dependent and independent variable around.

As for the question if the learning curve is a law:

Price: There has been an assumption that learning curves automatically give you a lower production cost. This is not so. If you have cumulative experience, you have the potential for having lower production costs through learning: but there is nothing automatic or inevitable about the achievements of lower costs. There are some companies who don't learn; you can also have countries with industries such as the Japanese, that start with disadvantages, but adjust their industrial policies to get a high rate of diffusion in their country and hence to take a leap forward along and hence down the learning curve in cost terms.

Betts: I think there are two aspects of the learning process, strictly related to the problem of plant scale. One aspect, one area of the learning process is, let me call it, the rate and final level of utilization of fixed assets. Let me characterize it like this. Just supposing that we have built over there a half-a-million-ton-per-year ethylene plant. It is standing there. It is a fixed asset, it is inert, it won't bite us at all if we leave it alone. We have put into it an immense amount of investment and this has to be properly used, because investment is in fact the result of past successful efforts and therefore has to be carefully utilized. There is undoubtedly in real life an actual process which takes place in which human beings learn to face up to, control, and manage such fixed assets. It is the interface between inert things and the human mind. And there is a barrier somewhere in that interface. This demonstrates itself, in the rate at which this plant reaches its design or full capacity. This may be long and slow, and sometimes it may never reach it. The rate at which the human beings involved manage to achieve a greater output for that fixed asset in terms of tons an hour, preferably tons a day, much more preferably tons a year, much more preferably still tons a decade, that is what I describe as one aspect of the learning process.

The other aspect is the process of making better, more reliable or more credible investment decisions. This is a management problem. For example, we may find that in increasing plant scale, we have introduced an innovation. I am not talking of change in the manufacturing process or about some fundamental technical innovation, I am simply talking about building something much bigger. We may find that in practice and in contrast to our original expectations, and the basis on which the original investment decision has been made, that for example, the plant cost half as much more to build and twice as long to build. That is learning. We may find that it takes say 3 years to reach its design output or it never reaches its design output or in fact we may find ways in which to get a more rapid build-up of output. That is learning.

Now, learning by whom? This experience will be learnt only by the people actually involved. In other words, if you look at that, their process of learning, and you say, suppose that we build a duplicate plant 5 years later, undoubtedly, in my view, the benefit of their process of learning – provided that they have analyzed the data to find out why it is that the plant has cost more and the output has been less, is repeatable at least within their particular organization and circumstances. To look at such experience in a macroscopic way, as things on graphs, plotting everything against everything as though the whole of human race naturally performed according to some God-given law, is not valid.

Also Gold talked more about the learning curve concept in relation to the scale problem. This is summarized in his paper (Chapter 2).

The problem here seems to be whether the term learning should be restricted only to the process of people getting to do the same job better and better, or be allowed much broader implications. As indicated earlier in this section, the arguments come from the fuzziness of the term *learning*. Thus, the nature of learning and the sources of learning must be looked into in detail in order to clarify the point raised in the discussion.

6.2 SOURCES OF THE LEARNING CURVE

The learning curve was first introduced as an observed relation between accumulated production and the cost of production per unit in terms of direct labor hours in the airframe production industry (Wright 1936). In this industry, the learning by the assembly workers from the repetition of a complex task was considered to be the cause of the relation, and thus the term "learning curve" was adopted. Since then, similar phenomena have been observed in a wide range of industries manufacturing everything from electronic instruments to chemical products; the possible causes of these phenomena are controversial and have been discussed in the literature rather extensively.

Andress (1954) made the distinction between "learning in the literal sense on the part of workers" and "a whole series of other factors among which management innovations appear most significant," and concluded that the former learning is the predominant factor. But his observation was restricted to the aircraft industry.

Conway and Schultz (1959), who demonstrated the phenomenon for electronic assembly products and electromachine products, argue that the learning from repetition of the same task is not very important but such factors as changes in tooling, production methods, design and volume, improvements in quality, and improvements in management have a more significant effect. Levy (1965) has divided various possible factors into three classes that he refers to as planned learning, autonomous learning, and random or exogenous learning. The first two correspond to the causes pointed out by Conway and Schultz above, and the last one implies the improvement based on information acquired unexpectedly from the environment.

Baloff (1966) suggested the possibility of the existence of similar phenomena in a wider range of manufacturing industries:

> . . . the phenomenon usually results from an integrated adaption effort on the part of a variety of direct-labor, indirect-labor, and technical personnel, and it relies primarily on "cognitive" rather than manual "learning."

In terms of who learns what, engineers, supervisors, machine operators, maintenance men, quality-control personnel, and other indirect-labor employees can all make contributions through

> . . . such cognitive activities as redesigning the product or process, altering raw-material and end-product quality specifications, evolving more effective maintenance procedures and finding the proper operating "balance" of a manufacturing process.

More recently Bodde (1976) discusses in detail the forces that make the learning curve operational. In addition to the efficiency increase in labor in which he included not only direct labor but also maintenance personnel, supervisors and staff members, he points out other factors such as the introduction of new technological processes and improved methods, effects of substitution in the products, technical conservatism that results in product redesign, standardization of the product, and economies of scale and development of a common experience base; Badde points out that the last factor is important especially in multiproduct companies. Hedley (1976), from his association with the Boston Consulting Group, summarizes the variety of factors as:

- productivity improvement due to technological change and/or "learning" effects leading to adoption of new production methods
- economies of scale and of specialization
- displacement of less efficient factors of production, especially investment for cost reduction and capital-for-labor substitution
- modifications and redesign of products for lower costs

Certainly the range of potential factors that cause the learning curve phenomenon has increased as the areas of observation have increased from the restricted area of the airframe industry to more integrated areas such as the electricity, steel, and chemical industries. Beyond this point, it is more appropriate that one recognize the learning curve phenomenon as a description of the evolution of an industry rather than just enumerating every possible cause. In this macroscopic viewpoint, the causes of the learning curve phenomenon can be related of the major events that have taken place during the evolution of the industry. The Department of Prices and Consumer Protection (1978) describes this clearly:

In the course of their evolution, most industries expand their scale of plant to achieve real-terms reduction in unit costs. This process is reinforced by two other factors; technological advances, which are usually embodied in newer, larger plants; and a process whereby managers and operators learn from experience how to operate particular technologies and facilities more effectively.

Thus, in addition to the various factors mentioned before that can be attributed to "learning", change in scale and technological advance must be considered as the major sources of the learning curve phenomenon.

6.3 EVOLUTION OF MANUFACTURING SYSTEMS

Before proceeding to the discussion of the implications, usefulness, and limits of the learning curve concept, it is perhaps wise to find a structural basis on which the relations between the general learning effect, scale increase, and technological advance can be discussed.

We will structure this on the basis of hierarchical decomposition and time dichotomy of the various causes raised in the previous section. First, let us consider the hierarchical level of an industrial organization. Here we follow the level classification by Cantley and Glagolev (1978); the unit level, plant level, organization (firm) level, and industry level. Each level in this classification is not very strictly defined; however, the classical observations in the airframe industry or the electronic or electromechanical assembly products in Conway and Shultz (1959) are examples, on the unit level. Cameras, time recorders, basic paper products, and glass containers are examples on the plant level or the firm level, and electricity and petroleum production are examples on the industry level.

Casual observation on the different levels suggests that in the higher levels, more factors would enter as the source of learning phenomenon. To look at the unit or the plant level implies that there is a given fixed plant (asset) and the learning implies the general effort toward increasing the efficiency of the given asset. Let us use the term "learning by doing" in order to express this subclass of learning. Stated differently, this learning by doing would correspond to the description by Joskow and Rozonski (1979): reduced labor requirements as tasks become routinized through repetition, more efficient production and labor scheduling and improved production control by management, improved routing and handling of material thanks to the engineering department of a firm which redesigns the capital equipment utilized by workers and makes changes in the operation of the plant.

This class of learning is the one associated with the learning curve in the traditional sense, i.e., for airframes, electronic products, electromechanical products, or more recently power plant construction (Kennedy and Allen 1979). The essential feature of this class is that the production process being looked at (or which will be looked at) is fixed and all sources of learning come from the minor improvements in an effort to utilize the given production facilities. In this case, it is assumed that the production process under consideration can be isolated from the surrounding environment and no major technical changes or scale changes occur over the time period of ibservation. In this situation workers learn in the sense that they do better and better as they work on the same (or a similar) job. Also workers and managers will get better in utilizing the given production facilities. Minor improvements may be achieved by design changes or changes in methods. That is, this class of learning is the result of a continuous effort towards better utilization by everyone who is related to the particular (isolated) production facilities.

In the higher levels, the learning becomes more sophisticated in nature and perhaps it is rather difficult to describe exactly who learns what. The people in research and development in a company learn about new theory or new technological know-how. The people in top management acquire information concerning the operation of the company. Here we refer to this class of learning as the "accumulation of knowledge." These possible sources of learning at different levels are also discussed in Cantley and Sahal (1979) from whom Figure 6.2 is taken. In this figure, the arrows represent information transfer, or "learning." They are of three kinds:

1. The circular arrows represent learning occurring cumulatively over time within a particular entity on its own level.
2. The vertical arrows represent transfer of information or know-how between levels.
3. The horizontal arrows represent transfer between an entity and other entities on the same level – whether or not within the same higher level

Thus, in addition to the possible learning inside each level, they stress the importance of information exchange across the boundaries of these levels.

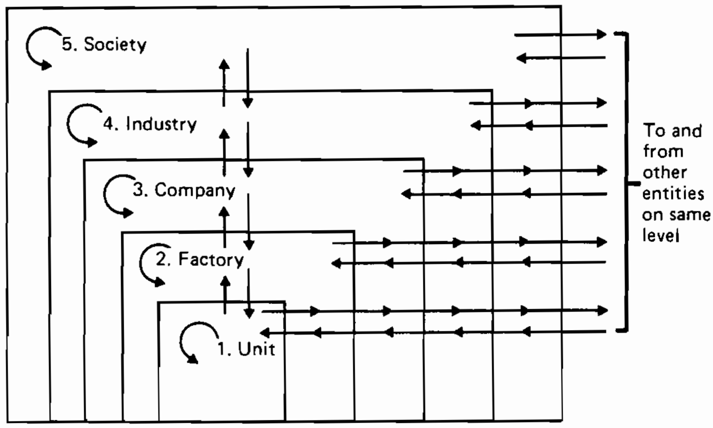


FIGURE 6.2 Levels and directions of learning or information transfer. Source: Cantley and Sahal (1979).

The second basis of structuring is related to the time horizon. Roughly speaking, the learning from repetition occurs in everyday work in a somewhat continuous manner, while design changes and other supervisory changes occur less frequently. Further, scale changes and major technological advance or changes occur much less frequently. These horizons were indicated in Lofthouse (1974), in which he used the terms "short run," "long run," and "very long run." He noted that such a treatment is always under criticism; how long is short or long or very long. But the horizons often serve adequately for structuring purposes. Learning by doing, defined above, tends to occur during the short run where no scale changes and technological advance would occur. In the long run, scale changes are allowed and in the very long run, technological changes are allowed. This implies also that the learning by doing is conditioned by scale and technology, and scale is conditioned by technology.

The discussion so far can be summarized into the hierarchical structure shown in Figure 6.3. Note that not all of the information flow is indicated; for example, there may be direct information acquisition in every box, which is not shown. Figure 6.2 would effectively supplement Figure 6.3. It is interesting to note that the structure is similar to one of the well-known hierarchical structures, i.e., the multilayer structure in which any action in the upper layer specifies the condition for the lower layer and activates actions in the lower layer; thus the actions in the upper layer take place less frequently than the lower-layer actions (Lefkowitz 1966, Mesarovic et al. 1970).

The broader class of learning, i.e., accumulated knowledge, resides perhaps in every box and it is rather difficult to show in the diagram. The role of this type of learning is conditioning; a certain level of accumulated knowledge enables a technological change or a scale change to happen.

6.4 USEFULNESS AND LIMITS

In the previous section, we sorted out various causes of the learning curve phenomenon

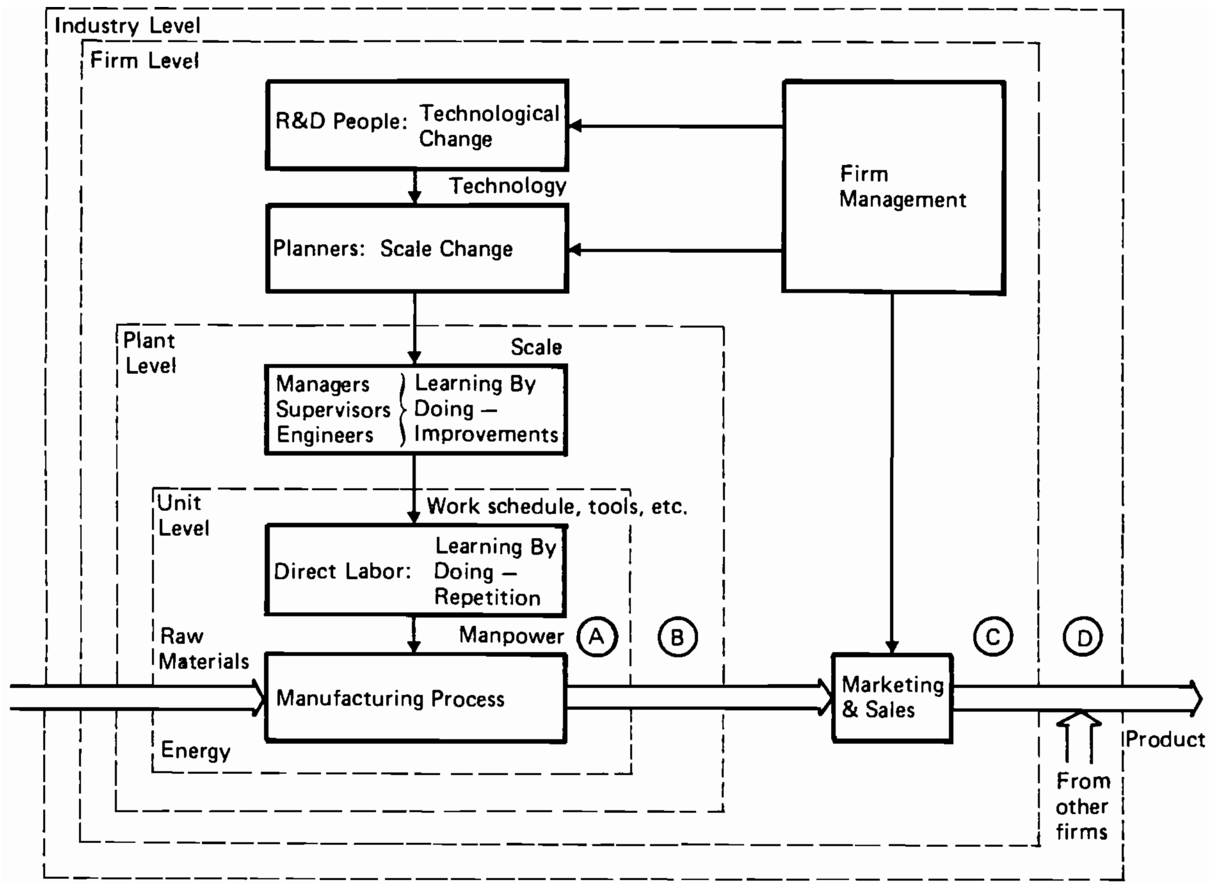


FIGURE 6.3 Conceptual hierarchical structure of manufacturing process: technology, scale, and learning.

based on a literature survey. It is clear from Figure 6.3 that the usefulness and limits depend very much on the purpose of using the learning curve and the level on which it is to be observed (this specifies the range of possible causes of the phenomenon).

In the early days, the learning curve was observed in the unit or plant level. Andress (1954), from his experience in the aircraft industry, finds the learning curve useful for estimating the direct labor cost, which then becomes a basis for price negotiation, make or buy decisions, and others. Since he essentially limits the major cause of the learning curve to the learning of workers from repetition (see the unit level in Figure 6.3), he suggests that every other possible cause must be removed from the raw data in order to get an accurate forecast. In fact, he did not observe similar phenomena in less labor-intensive manufacturing such as basic chemicals and petroleum production. Conway and Schultz (1959) also see the main purpose to be a means of estimating labor costs for pricing, major design changes, make or buy decisions, purchasing, and labor and facility needs for planning. Although they see the necessity for the aggregation of raw data in order to get more accurate forecasts, and hence consider various causes other than the literal sense of learning from repetition (and they do not believe that the learning curve represents a natural law), they conclude that progress can be predicted within some firms by the use of the learning curve, with tolerable amounts of error.

Young (1966) argues that even in the aerospace industry, the learning curve is not always meaningful because of the pessimistic estimate of labor needs, manufacturing methods, and tool changes, and so on, and concludes that there is a distinct possibility that the reduced labor trend may be caused solely by budgetary action.

More recently, as it has been observed in a wider variety of industries, the learning curve has been treated as an index that expresses the combined effect of various factors, including scale changes and technological changes. Abernathy and Wayne (1974) claim that the learning curve is a useful tool for strategic planning in marketing, financial planning, and production. Also, they imply that through exploiting the learning curve,

> . . . a strategy that seeks the largest market share at the earliest possible date can gain not only market penetration but also advantages over competitors who have failed to reach equal volume.

But they emphasize that cost reduction along the learning curve does not happen automatically, and a strict cost minimization strategy is necessary to pursue the benefits of the learning curve. Hedley (1976) points out similar implications of the learning curve phenomenon. However, by using the Ford Model T as an extreme example, they discuss some unfavorable consequences of pursuing the learning curve.

This creates a dilemma: "management must realize that the risk of misjudging the limit rises directly with the successful continuation of the strategy" and "production innovation is the enemy of cost efficiency." Similar arguments can be found in Bodde (1976). He finds the learning curve more useful in long-range strategic planning, especially in the formulation of competitive strategy, and denies its usefulness in operating controls or short-term decision making. And he concludes that

> . . . to apply the learning curve successfully the manager will need an awareness of the multi-dimensional forces behind it and how these forces can be integrated into the total strategy of the organization.

These arguments in the literature can be clarified from Figure 6.3. That is, the usefulness and limits depend on the point at which the learning curve is to be observed. In

Figure 6.3 four possible points (A through D) are shown. If we are observing the learning curve at point A, then curve would represent the literal sense of learning where all other factors that come from the level above the unit level are fixed. If the literal sense of learning can be isolated in the raw data, then the learning curve at this level will be useful to estimate labor requirements in the future for existing products or even for a new product. (Yelle (1979a) has proposed a disaggregation – aggregation approach for estimating the parameters for the learning curve for a new product.) By the nature of the learning curve, its applicability will be restricted to those products that can be produced more efficiently by learning from repetition.

If the learning curve is observed at point B, then it can show the effects of various factors. This is the situation where there is a given plant and the learning curve may be observed as an overall cost reduction effort. If the precise relation between each potential cause and the amount of cost reduction is known, then the curve may be useful for control purposes. However, in practice these relations do not seem to be easy to determine and the learning curve may be useful at the most for monitoring. As long as the danger associated with pursuing the learning curve too far is recognized by the manager, it can be a useful guide or index for his cost reduction effort at his plant.

The learning curve observed at point C may be useful for strategic planning at a company, as was shown in the Boston Consulting Group's study and in other studies (Hedley 1976, Abernathy and Wayne 1974, Bodde 1976). It provides an explanatory tool for the implications of holding a large market share. Again the limits of riding down on the curve must be well understood and care is needed to apply the learning curve in, for example, as Taylor puts it,

> defining the product, the market segment, and the stage in life-cycle, and in ensuring that management have the capacity to realize the potential cost savings, and at the same time maintain a high rate of product innovation.

The main sources of the learning curve on this level are categorized as technology, scaling up, and a broad sense of learning. How much would each category contribute to cost reduction? In some industries scale may be the dominating factor, whereas in others it may not. The recent study by Snow (1975) tried to separate the effects of economies of scale and technical advance in satellite communication. Derkx *et al.* (1978), in their study of the steel industry, which was referred to in the Discussion section of this chapter, were successful in seeing the economies of scale in a dynamic sense; the dynamic economies of scale in turn were used by Cantley (1979) to illustrate the case for protectionism.

Another possibility of using the learning curve on this level is to incorporate it into the model of technical substitution. Recent attempts by Robinson (1979a, b) may expand the usefulness of the learning curve concept.

At the highest level, we can still find a learning curve (at point D). Examples are electricity generation and petroleum (see Marchetti 1975). On this level the relation should be treated as a hypothesis backed up with some evidence. This hypothesized relation may be used in strategic planning in a wider sense, for example, in energy planning (Aronofsky and Blum 1978) or in global modeling (Roberts 1978), or in national strategic planing.

6.5 CONCLUDING REMARKS

The issues raised in the workshop discussion have been identified. These issues have been controversial and have been the subject of much literature in the past. I have tried to put the various points together by using a conceptual hierarchical structure that represents the evolution of an industry in the simplest way.

A general conclusion obtained in this directed literature survey is that the learning curve concept is useful at various levels of industry in somewhat different ways (serving different purposes), but at the same time one must be careful in applying this concept to a particular case.

What is lacking is a concrete model that describes the evolution of an industry and results in a learning curve. One such attempt can be found in the paper by Sahal (1979b), but there must be various types of models that result in a learning curve. Such a theoretical development and a demonstration in a particular industry would help to establish the learning phenomena better and to clarify its usefulness in a production application.

REFERENCES

Abernathy, W.J., and K. Wayne. 1974. Limits of the Learning Curve. Harvard Business Review. September/October: 109 – 119.

Andress, F.J. 1954. The Learning Curve as a Production Tool. Harvard Business Review 32: 87 – 97.

Aronofsky, J.S., and H.A. Blum. 1978. Estimating the Potential of a Solar-to-thermal Collector Industry. TIMS Studies in the Management Science 10 (Energy Policy): 197 – 208.

Arrow, K.J. 1962. The Economic Implications of Learning by Doing. Review of Economic Studies 29: 155 – 173.

Baloff, N. 1966. The Learning Curve – Some Controversial Issues. Journal of Industrial Economics 14 (3): 275 – 282.

Baloff, N. 1971. Extension of the Learning Curve – Some Empirical Results. Operational Research Quarterly 22 (4): 329 – 340.

Bodde, D.L. 1976. Riding the Experience Curve. Technology Review. March/April: 53 – 59.

Boylan, M.G. 1975. Economic Effects of Scale Increases in the Steel Industry. The Case of U.S. Blast Furnances. New York: Praeger.

Cantley, M.F. 1979. Scale, Protectionism and European Integration: The Structure of Strategic Control in a Turbulent Field. Paper prepared for Third European Congress on Operational Research, Amsterdam, 9 – 11 April 1979. WP-79-42. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Cantley, M.F. and V.N. Glagolev. 1978. "Problems of Scale" – The Case for IIASA Research. RM-78-47. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Cantley, M.F., and D. Sahal. 1979. Who Learns what? A Conceptual Description of Capability and Learning in Technological Systems. WP-79-110. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Conway, R.W., and A. Schultz. 1959. The Manufacturing Progress Function. The Journal of Industrial Engineering. January/February: 39 – 54.

Delombre, J., and B. Bruzelius. 1977. Importance of Relative Market Share in Strategic Planning – A Case Study. Long Range Planning 10 (8): 2 – 7.

Department of Prices and Consumer Protection. 1978. Economies of Scale and Learning Effects. A Review of Monopolies and Mergers Policy: A Consultative Document. Presented to British Parliament by the Secretary of State for Prices and Consumer Protection by Command of Her Majesty, May 1978. London: HMSO. Annex C.

Derkx, H.H.J.M., A. Kamerman, and A. van der Rijst. 1978. How Experience and Attitude Affect Steelplant Productivity. Iron and Steel International. October: 319 – 329.

Gold, B. 1974. Evaluating Scale Economies: The Case of Japanese Blast Furnaces. Journal of Industrial Economics 23: 1 – 18 (September).

Gold, B. 1979. Revising Prevailing Approaches to Evaluating Scale Economies in Industry. Paper presented at the workshop "Size and Productive Efficiency: The Wider Implications," 26 – 29 June 1979, at International Institute for Applied Systems Analysis, Laxenburg, Austria (chapter 2 in this volume)

Hedley, B. 1976. A Fundamental Approach to Strategy Development. Long Range Planning. December: 2 – 11.

Joskow, P.L., and G.A. Rozanski. 1979. The Effects of Learning by Doing on Nuclear Plant Operation Reliability. The Review of Economics and Statistics 61 (2): 161 – 168.

Kennedy, W.J., and C.R. Allen. 1979. Learning Curves as Applied in Power Plant Construction. Power Engineering 83 (9): 63 – 65.

Lefkowitz. 1. 1966. Multi Level Approach Applied to Control Systems Design. Transactions of American Society for Mechanical Engineers 8: 392 – 398.

Levy, F.K. 1965. Adaptation in the Production Process. Management Science 11 (6): B/136 – B/154.

Lofthouse, S. 1974. Learning, Costs and Market Share. Rivista Internazionale di Scienze Economiche e Commerciali. 21 (11): 1014 – 1040.

Marchetti, C. 1975. Transport and Storage of Energy. RR-75-38. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mesarovic, M.D., D. Macko, and Y. Takahara. 1970. Theory of Hierarchical, Multilevel Systems. New York and London: Academic Press.

Roberts, P.C. 1978. Learning Processes in Global Models. Prepared for the 6th Global Modelling Conference, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Robinson, J.M. 1979a. Technological Shift: A Graphical Exploration of Progress Functions, Learning Costs and Their Effects on Technological Substitution. WP-79-105. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Robinson, J.M. 1979b. Technological Shift: As related to Technological Learning and Technological Change. WP-79-106. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Sahal, D. 1979a. A Theory of Evolution of Technology. International Journal of System Science 10 (3): 259 – 274.

Sahal, D. 1979b. A Theory of Progress Functions. Transactions of the American Institute of Industrial Engineers 11: 23 – 29.

Sahal, D. 1979c. Recent Advances in a Theory of Technological Change. IIM dp-79/11. Berlin: International Institute of Management. (An abbreviated version of this paper appeared as "Technological Progress and Policy" in D. Sahal, ed., 1980. Research, Development and Technological Innovation, Recent Management Perspectives. Lexington, Mass.: D.C. Heath, pp. 171 – 198).

Sahal, D. 1979d. Laws Which Govern the Maximum Capability of Technology. Berlin: International Institute of Management.

Snow, M.S. 1975. Investment Cost Minimization for Communications Satellite Capacity: Refinement and Application of the Chenery-Manne-Srinivasan Model. The Bell Journal of Economics 6 (2): 621 – 643.

Taylor, B. 1976. Managing the Process of Corporate Development. Long Range Planning 9 (3): 81 – 100.

Wright, T.P. 1936. Factors Affecting the Cost of Airplanes. Journal of the Aeronautical Sciences 3 (4): 122 – 128.

Yelle, L.E. 1979a. Estimating Learning Curves for Potential Products. Industrial Marketing Management 5: 147 – 154.

Yelle, L.E. 1979b. The Learning Curve: Historical Review and Comprehensive Survey. Decision Science 10(2): 302 – 328.

Young, S.L. 1966. Misapplications of the Learning Curve Concept. Journal of Industrial Engineering 17 (8): 410 – 415.

CHAPTER 7 COPING WITH THE UNCERTAIN FUTURE

J.A. Buzacott
International Institute for Applied Systems Analysis,
Laxenburg, Austria

7.1 DISCUSSION SUMMARY

The question of the way in which our uncertainty about the future affects decisions on scale arose during discussion on Fisher's paper, and in Betts's paper and in the subsequent discussion.

We begin with the discussion on the papers by Fisher and Spinrad:

de Hoffmann: One of the key things you have to think about when a power station is going to be built is that you have to predict rates of growth of a system for at least 10 years ahead, and it is getting longer and longer. This is just nonsense. It is just throwing darts. You simply can't predict power usage for say 14 years ahead when it takes that long to build a nuclear plant in the U.S.A. One of the people from the Electric Power Research Institute in the U.S.A. was talking to me recently and he said "It is simply a question of whether you expand power systems adiabatically by small additions or whether you take huge chunks in which case you deal with an unpredictable situation." I refer you to "Defending against uncertainty in the electric utility industry" (Ford and Yabroff 1978).

Fisher: On the proposition that uncertainty of the future provides motivation for incrementally adding small units, I couldn't agree more; but I didn't quantify it because I wanted to show that, even in the absence of uncertainty, which pushes you towards small units, you didn't even need that much of a push to go a lot further towards smaller units than people currently think they ought to go.

In summarizing the discussion, Cantley reverted to this issue:

Cantley: I wouldn't like us to lose sight of . . . de Hoffmann's point, which has been taken up by Lucas at Imperial College in the UK context (Abdulkarim and Lucas 1977). If you are going to have to forecast further ahead, then your forecast error distribution is going to be wider and therefore your planned safety margin has to be larger; therefore you probably finish up wrong on the high side, so that you have an underutilized plant which will even feed back to your operating regimes and your fuel efficiencies. The point is that this is all part of the broader social systemic affects of scale although it arises in this special case from the particular technologies of a particular industry.

Shutler: As a quick comment on that one, I would like to point out that in England the overestimation of demand in previous years has led to the earlier scrapping of plant which is so inefficient that total costs remained very much the same as they would have been, had there been a 100-percent accuracy in the forecast.

Further mention of the effects of uncertainty was made in Betts's paper in a discussion of the pressures acting against increase in plant scale:

> . . . although economies of scale suggest a reduced investment per unit of additional capacity, the commercial risks associated with top-end sales is often higher.

Then later:

> It should be recognized that the design and construction of chemical plants of a size significantly larger than the current norm in itself represents an innovative step . . . many of the problems inherent in plant scale and the potential discontinuities of the economies of scale are difficult to foresee. Once the commitment has been made, corrective action may be late or not completely effective.

and

> . . . one way to avoid the risk of scale is to duplicate or even exactly replicate existing plants, or to introduce only small step changes.

and this was followed by discussion on the reasons why capital cost and construction time is often underestimated.

The discussion on Betts' paper reverted to this point when de Hoffmann made an observation based on the data on the utilization of ethylene plants in Cantley (1979) which show a dramatic decline in plant utilization in Western Europe from around 87 percent in the early 1970s to a forecast of around 70 percent in the early 1980s, partly because some very large units were brought in 1979 – 1980. Had these large units not been built, utilizations of more than 80 percent would be achieved.

de Hoffmann: Is it reasonable to expect that demand will go up or is this always the hope? Will you have a hard time in reaching a utilization of somewhere between 82 and 85 percent or is this a moving target?

Betts: It is true to say that prior to 1973 at least, the growth rate in demand for chemical products did in general come to the rescue of past investment decisions. There was a seemingly never-ending increase in the demand for the products, partly promoted by the cost and price reductions. On the question of plant utilization, one has to look very carefully at the macro view of the situation. Plant underutilization can be caused by many factors such as too optimistic a forecast of total demand, too optimistic a forecast of manufacturer's market share, by some breakdown of the plant, by taking a long time to commission, all sorts of things. The problem I tried to address in my presentation was, to what extent is utilization of plant capacity specifically and directly related to size.

Price returned to this point subsequently:

Price: We must recognize that there is not going to be very much investment in the petrochemical industry in Western Europe for say 10 years ahead because the two main factors which have in the past brought this development, namely a high rate of growth in demand for the products and a reduction of operating costs per unit through increasing scale of production, are no longer very evident.

Betts: I agree with what you say up to a degree, but only up to a degree. I think the point you make about plant scale insofar as chemical commodity products are concerned is valid. However, a new challenge which the chemical industry is about to face is as big, though somewhat different, and undoubtedly will be its principal concern in the next few decades. It is that of feedstock and energy; so it may be that we will find there is less to be gained purely by increasing physical size and capacity. But the motivation is now shifting, its motivation has shifted, into the problems of raw materials by the year 2000. What will they be and in particular will these feedstocks be processable by existing plant and existing investment?

Gold: I just wanted to comment on the point de Hoffmann raised, to which there have been allusions (both in the discussion of Fisher's paper as well as in this discussion) about the importance of correctly estimating demand and the penalties of winding up with plants which are underutilized. I think this is a serious problem, but I think it is also misunderstood. You cannot use aggregate demand or aggregate capacity as the basis for evaluating whether you should build a new plant. Whereas people have argued that the underutilization of existing capacity in various industries implies some sort of defect in the economy, our studies of a number of industries, including steel and cement, suggest quite the opposite. It may represent the degree of dynamism in the economy.

Why do we build new plants if we now underutilize existing plants, if we have more capacity than current demand? Answer: because (a) there are newer, more advantageous technologies available; (b) we want to move into developing markets in new geographical areas; (c) there are changes in product mix or product specifications which can't be effectively met with the equipment in the existing plants; and (d) in some cases, we may undertake preemptive investments – move into a newly growing market with a big plant whose underutilized capacity will keep everbody else out while we develop a dominating distribution and service system. The point I am trying to make is that the decision with respect to the particular plant may have very little to do with the average relationship between total capacity and total demand. You must make a much more precise estimate of what is likely to be the demand for the products of this particular plant with this particular kind of process in the particular market area that we plan to serve, rather than using aggregate comparisons.

Betts: It is difficult to estimate total demand, it is very, very difficult to estimate market share and in a free market economy you have no control over what other manufacturers, whether existing or new, may or may not decide to do. One makes one's best estimates but they are clearly susceptible to a fairly high degree or risk.

Gold: To be concrete, we have gone to the people who build huge new paper mills even when the capacity utilization in the U.S. has fallen below 70 percent and asked why they do so. They say: "Yes, it really is too bad for what will now be the marginal plants but we don't anticipate any difficulty because of the advantage of our new plants."

After reviewing the behavior pattern that characterized the postwar development of the chemical industry (cf. Simmonds 1972), Simmonds then went on to say

Simmonds: One other comment on an issue which has not been raised except briefly by Mr. Price: the future of the industry. This industry runs on tonnage, poundage, gallonage, in other words on a quantity basis, the idea "more is better." The future of the chemical industry is entirely the opposite, better is less, the more skillful use of molecular architecture is the way into the future as, when and if supplies of energy and petroleum hydrocarbons are either actually constricted or become too expensive for people to buy. Lying on the side is the biochemical industry, in which you use low-temperature methods and biological methods to compete; so the question of raw materials is no longer the hydrocarbons, it is also the natural products once again. What I am after here is that when you look at scale you have got to be very careful not to make the assumption that the current behaviour pattern will necessarily extend into the future if the factors producing that behavior pattern are actually changing, and, boy, are they changing at the moment. So you have got two separate issues: what happens if you go on, all other things being equal as in the past; and what happens when those things cease to be equal.

Betts: Yes. One comment which is that insofar as change in feedstocks is concerned there will no doubt be a gradual shift in feedstocks used by chemical industry – possibly a gradual moving over to coal, in certain areas of the world biomass – but I think the point ought to be made that, whatever these new feedstocks are and the technology developed to use them, at least initially, they have to compete with the existing established cost and price of the products currently manufactured from petroleum sources. That is obvious. It is of no interest to you or me to buy a plastic bucket at $40 made from biomass ethanol if we can buy a plastic bucket made from the cracking of petroleum naphtha at $20. So therefore what this means is that there will be a time phase in which the price of petroleum has to rise to a level at which the new technology based on the newer raw materials becomes economically attractive to the consumer, and one of the factors in that is that to some extent the current price of chemical products has obviously been influenced by scale, namely by large-scale plants.

7.2 MODELS OF SCALE DECISIONS UNDER UNCERTAINTY

7.2.1 Demand Uncertainty

The usual way of allowing for the effect on scale decisions of uncertainty in the estimates of future demand is to adjust the forecast to allow for the uncertainty and use the resulting demand estimate in a deterministic model (Anderson 1972). That is, if demand at some future time is estimated as having mean D and the variance of forecast errors is σ^2, then the scale decision will be based on a deterministic demand of $D + k\sigma$ where k is chosen so that the probability of demand greater than $D + k\sigma$ is less than some safety margin α. Effectively, this means that the decisions will be made assuming a faster demand growth than the expectation; this tends to result in installations that are larger than if the decision was based on the expected demand.

However, there are a number of models in the literature that attempt to allow specifically for the effects of uncertain demand. Manne (1961) assumed that demand growth followed a Brownian motion pattern. This is the continuous time analogue of assuming that demand growth per period comes from a set of identically independently distributed random variables with constant mean. He assumed that the time to construct a new installation was zero and showed that the resultant installation was somewhat larger than when the demand growth had no variability.

By contrast Meier (1977) assumed that demand growth up to the time horizon was at a constant rate; however, the value of the growth rate was uncertain at the time the decision on the size of the first installation was made. However, if the first installation proved to be too small, the next installation could be constructed on the basis of the actual demand growth which is then assumed to persist to the planning horizon. His model does not require any assumption on construction lead time (except that construction of the second installation begin after the actual demand growth has been observed.) He found that the optimal size of the first installation is somewhat smaller than that appropriate if the growth rate is known with certainty when the decision is made.

Both Manne's and Meier's models assume that the only relevant costs are the size-dependent construction costs of the installation and those resulting from the time value of deferral of subsequent installations. By contrast, Sharkey (1977) in a one-period model assumed that an installation, although built, would not operate if the realized demand was insufficient to cover its fixed operating costs, which were size-dependent. This can result in a situation where it is optimal to build more than one installation and the installations do not have the same size.

Finally, the two papers referred to in the discussion, Ford and Yabroff (1978) and Abdulkarim and Lucas (1977), both try to model the effect of size-dependent construction lead times with uncertain future demand. Ford compares by simulation the effect of short and long lead times. He assumes that there is no economy of scale and as would be expected this makes the short lead time (which he associates with small plants) preferable. Abdulkarim and Lucas allow for economies of scale in construction and operating costs, deteriorating reliability with size and the longer construction lead times of large units. They arrive at an optimal unit size of around 250 MW, less than the 450 MW that Fisher's formula gives. The smaller optimal unit size is almost certainly due to the effect of the longer construction lead time with increasing size, although consistent overestimation of demand growth over most of the period simulated might also contribute to the optimality of smaller units.

This brief review of the literature on models allowing for uncertainty in demand shows that there do not appear to be any conclusive results that explain how decisions in installation size should be made. There is obviously a need for models that have reasonable assumptions concerning the uncertainty of demand, the forecasting procedure, the size-dependent capital and operating costs, and construction lead times. The most intractable problem is to specify statistically plausible models of demand, including its serial correlation over time and its cross-correlation with other factors relevant to the investment decision; the nonstationary world environment and the gaming aspects of competitive interactions render any simple statistical specification a naive guess.

7.2.2 Other Factors

There were a number of other factors that were mentioned in the discussion as sources of uncertainty that could affect scale decisions, such as technological risk (will the plant perform as expected? will the plant work at all?), strikes and labor unrest, and the ability to forecast and control construction time and cost. Gold particularly stressed that the Japanese plant constructors achieved dramatically shorter times than U.S. constructors by a stringent control system reminiscent of U.S. standardized manufacturing industry. These aspects do not appear to have been discussed in the literature concerning models of scale decisions.

In connection with technological risk, and as partial support for the utilities' decisions on electric generating unit sizes, Figure 7.1 from Krasnodebski and Christians (1977) shows a projection of generating unit equivalent forced outage rate made by FPC (Federal Power Commission) in 1963 at the time when decisions on unit sizes for the early 1970s were being made. On the same figure is also shown the actual average performance of all North American generating units over the period 1966 – 1975 as reported by EEI (Edison Electrical Institute). Also shown is the actual experience of Ontario Hydro with their generating units of various sizes. If we assume that maintenance and planned outages rates are independent of unit size, the FPC projection implies that the capacity factor decreases by 4 percent per 1,000 MW of unit size, while the EEI data give a reduction of capacity factor of 25 percent per 1,000 MW. With a scale factor of 0.16, this implies that:

optimal size based on FPC projection = 1,000 MW
optimal size based on EEI date = 170 MW

It can be concluded that, as far as electric generating units are concerned, utilities chose too large a size because they overestimated the rate of growth of demand and underestimated the deterioration of technical performance with size. These types of consistent biases might be difficult to allow for in a model, although it would be interesting to determine what factors, institutional and other, led to these errors.

Even more critical is whether management in fact recognizes that certain factors may change quite substantially and hence affect the viability of their scale decisions. This underlies Simmonds's comments on industrial behavior patterns and the way in which managers assumed that a particular constellation of factors would continue to exist.

Two further perspectives on attitueds to technological risk were referred to by workshop participants. De Hoffmann, during the discussion on Fisher's paper, read an excerpt from a letter from the president of RAND written in the 1970s concerning experience acquired during the development of a complex system:

> My recommendations for sequential development include austere development, incremental design and time to test. Faced with such a large degree of uncertainty, the prudent decision maker will (a) elect not to make decisions that can't wisely be made now, (b) make today only the decisions that must be made today, (c) plan for the resolution of uncertainty over time.

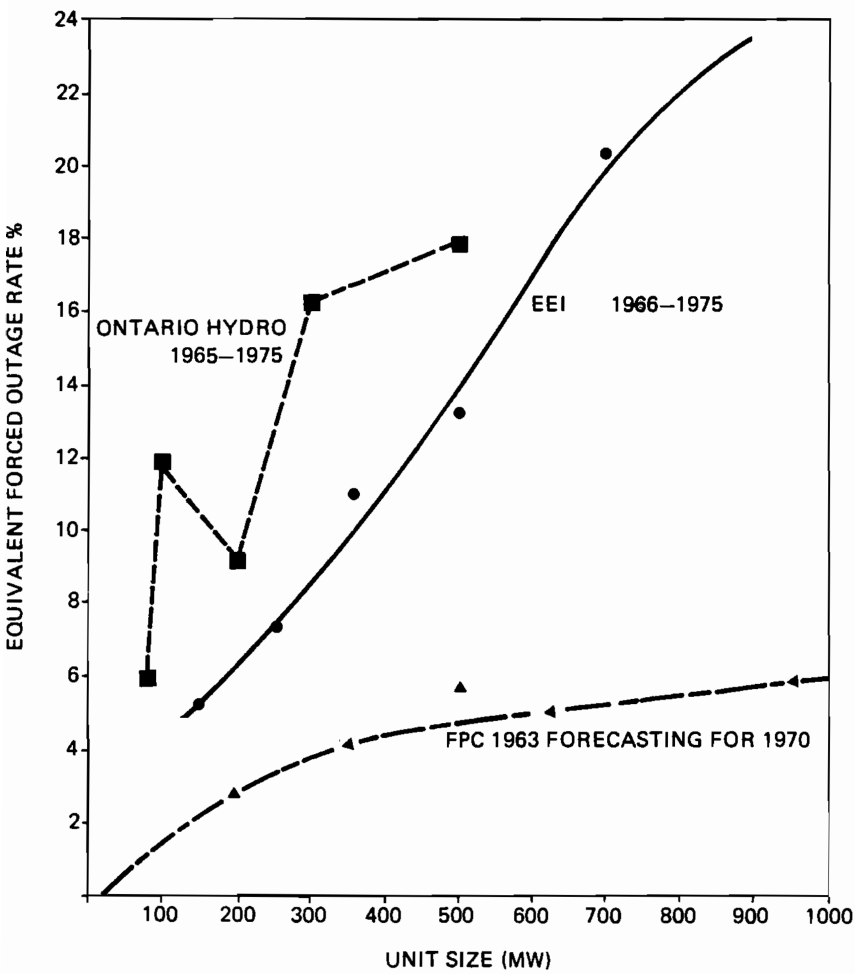


FIGURE 7.1 Equivalent forced outage rate versus unit size for fossil units.

Then Derkx and Kamerman, in reviewing the difference in blast furnace performance between japanese and European steel manufacturers, made the point (van der Rijst *et al.* 1978):

In Europe, planning was based on optimization and the minimizing of risks . . . pursuing minimum costs meant avoiding all possible risks . . . . The Japanese were less worried about the optimum costs, and efforts were directed more via fast technological innovation towards an essential and general lowering of costs. They were less inclined to caution, an attitude which emerges from the following quotation concerning the construction of Nippon Steel's Oita steelplant: "When a number of techniques covering extensive areas of projected work includes many unsolved problems as in this case, one tends to seek solutions on the safer side. But we succeeded in avoiding such a wait-and-see attitute by collecting as many data as were then available and making as many tests and investigations as possible, which enabled us to make definite decisions."

7.2.3 Implications of Uncertainty

Although an uncertain future would appear to have a significant effect on scale decisions, it will be noted that even if a firm recognized that its projections might not be realized, some of the participants were not so certain that its decisions on the investment timing and magnitude would be affected. Gold's and Shutler's comments can be interpreted as implying that the firm is likely to go ahead anyway on the basis of what seem to be "rational" expectations about the future. If the future turns out to be different from what was foreseen, then the firm anticipates that it will be able to respond and exploit the advantages that its investment has realized. For example, in the case of ethylene, Simmonds mentioned that the price reductions that followed from surplus capacity stimulated demand so that the utilization of capacity soon improved. More recently, the surplus capacity has been used to scrap old and obsolete plants (Cantley 1979). Shutler pointed out a similar situation of electric generating units in the UK. Surplus capacity enabled inefficient units to be retired from service and so the penalty of over-estimation of demand was greatly reduced. Nevertheless, it would seem desirable for the firm to consider carefully what its options are if it goes ahead with a particular investment and the projections on which the investment was based are not realized.

Finally, although there is a need for better models of scale decisions with an uncertain future, Gold's remarks would support the point of view that, even with much improved models, what matters most is the overall attitude of the firm to the future, that is, its long-term goals, and to a large extent they will determine its decisions.

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.

Anderson, D. 1972. Models for Determining Least-Cost Investments in Electricity Supply. The Bell Journal of Economics 3: 267 – 299.

Cantley, M.F. 1979. The Scale of Ethylene Plants: Background and Issues. WP-79-43. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ford, A., and I.W. Yabroff. 1978. Defending against Uncertainty in the Electricity Utility Industry. Los Alamos Scientific Laboratory Report. LA-UR-78-3229.

Krasnodebski, J., and J. Christians. 1977. Reliability and Maintainability in the Design of Electric Power Stations. Paper presented at the "Availability Engineering Workshop" sponsored by Electric Power Research Institute, Albuquerque, New Mexico. October 1977.

Manne, A. 1961. Capacity Expansion and Probabilistic Growth. Econometrica 29 (4): 632 – 649.

Meier, P.M. 1977. Game Theory Approach to Design Under Uncertainty. Journal of the Environmental Engineering Division. ASCE. 103 (1): 99 – 111.

Sharkey, W.W. 1977. Efficient production When Demand is Uncertain. Journal of Public Economies 8: 369 – 384.

Simmonds. W.H.C. 1972. The Analysis of Industrial Behaviour and Its Use in Forecasting. Technological Forecasting and Social Change 3: 205 – 224.

Van der Rijst, A., H.H.J.M. Derkx, and A. Kamerman. 1978. How Experience and Attitude Affect Steelplant Productivity. Iron and Steel International. October: 319 – 329.

Part 3

ORGANIZATIONS AND SCALE

OVERVIEW

The focus of this section is on the factors that determine scale at level 3, that of the organization.

The paper by Bendkowski, Stachowicz, and Straszak describes the results of a survey of mine managers in the Polish coal industry in which the difficulty of managing the mine was related to mine size.

Van Hees and den Hertog develop a conceptual model for determining the appropriate size of organizational units in a large diversified manufacturing firm. In particular, they outline the circumstances in which small size units are appropriate.

Shutler's paper is based on his experience with the UK Price Commission. He concludes that for the large companies investigated, a key aspect of an efficient company is the formality of the management system.

Egiazarian and Glagolev discuss the reasons behind the recent reforms in the organizational structure of Soviet industry aimed at greater concentration and specialization of production in large multiplant, multiproduct enterprises.

The last two papers are concerned with more general and formal models of the relationship between the scale and performance of production organizations. Stoyanov and Razvigorova describe an ambitious attempt to use a systems approach to model this relationship. Ansoff suggests a number of general propositions about the appropriate size of organizations and the appropriate size of their management.

The section concludes with reviews of the discussion on two issues relating to organization and scale:

1. Organization size, structure, and environment
2. Innovation and organization – in particular, the relationship between the effectiveness of the innovation process and the size of the organization.

CHAPTER 8 THE SCALE OF COLLIERIES AND THEIR TOP-LEVEL MANAGEMENT PROCESS REQUIREMENTS IN THE POLISH COAL-MINING INDUSTRY

J. Bendkowski
The Computer Institute of the Mining Industry Katowice, Poland
J. Stachowicz
Institute of Organization and Management Problems Bytom, Poland
A. Straszak
Systems Research Institute, Polish Academy of Sciences Warsaw, Poland

8.1 INTRODUCTION

This paper describes the results of the first stage of the research work carried out by Polish research institutions within the framework of the IIASA study "Coal – Issues for the Eighties." The focus of this study is the organizational aspects of the development of coal mining. Our particular research is concerned with the analysis of management systems in different sizes of collieries within the mining industry, the identification of current trends, and the design of improved management processes and structures for an industry that is essential to our national economy.

An important part of the research task is the analysis of the qualitative and quantitative aspects of the management systems that are dependent on the size of the mining divisions and mines. In our opinion, the results of the research will be of key importance to the further improvement of management organization in the coal industry.

8.2 THESIS

Differences of mine size lead to differences in the complexity of the situations in which managers participate and this results in differences in the degree of difficulty of the decisions that confront them.

The recognition of these differences will provide better guidance for colliery designers, since most methods currently applied focus on the technological and economic aspects of mine size and, near the optimum, give a relatively flat relationship between mine performance and mine size. Adding the organizational and managerial variables to the technical and economic ones should give more precise recommendations to the industry.

8.3 THE TASK

To investigate the above thesis, the following tasks were carried out:

- Analysis of the relationship between mine size and performance since 1950 in Poland
- Identification of some relationships between mine size and the complexity of management

In this paper, the results of the first task and some preliminary work on the second are reported.

8.4 DEFINITION OF SIZE AND PERFORMANCE

Strictly, a definition should state the genus and give the attribute which differentiates the thing to be defined from other sub classes. However, the issue of how to define completely and explicitly the size of a system is the subject of many disputes. For example, the meaning and measurement of system size and its influence on other features of the system is described in Melcher (1976), Caplow (1957), and Khandwalla (1977). The three Polish authors Zawiślak (1975), Doktór (1978) and Pańkow (1979) should also be mentioned.

From our point of view, the most convincing definition is one that describes the size of a system in terms of its influence on the environment, although this definition is not perfect and it is not in accordance with the rule quoted above.

Thus, in order to define the size of a system it is necessary

- To identify the boundaries of the system, i.e., the separation between the system and its environment
- To determine the quantitative criteria for evaluation of size

A socialist enterprise (Gliński 1977) is characterized by territorial, legal, financial, and organizational separation from the environment. While this separation is quite clear in the case of a mine, the extent of actual economic independence is a matter of dispute because economic independence is closely associated with decision independence and the extent of decision independence is related to the degree of centralization or decentralization, which is relevant not only to a planned socialist economy but also to a market economy. Within a branch of industry there can exist varying degress of decision independence; this problem is to be a subject for future research.

Nevertheless, for this study it has been assumed that the difference in degree of decision independence between mines is not significant. There is a great similarity, confirmed by many decision makers, between the underground mines in one mining region and such mines have been the main object of our study.

For the purpose of our analysis it has been assumed that the size of mines is determined by the following indicators:

- Output quantity
- Total employment
- Total length of mine headings and number of shafts

The performance is determined by

- Overal output
- Underground output
- General labor intensity
- Total electric power consumption
- Production costs

8.5 RESULTS OF THE COMPARATIVE STUDY

8.5.1 The Development of the Industry

The mining industry of Poland has a dominant position in the national economy and has shown considerable development from 1945 to 1979. There has been an increasing trend in coal output: from 1970 to 1977 coal output increased from 140 million tons to 186 million tons per year, an average annual increase of 6.6 million tons. It is forecast that there will be further increase of coal output to 210 million tons in 1980 and to 240 million tons in 1985.

This considerable increase in output in Poland is being achieved in spite of a continuous deterioration of the geological conditions for mining and it is expected that this deterioration will continue in the future. As the depth of exploitation goes beyond 1,000 m there is increased likelihood of problems of mine air-conditioning, rock burst, and so on. The planned increase in output will be achieved in spite of the fact that it is difficult to increase employment beyond the 380,000 people now employed in the mining industry.

The Polish coal industry has about 65 mines grouped into seven areas. Six groups of ten mines each are located in the Upper Silesian Coal Basin. The seventh group is located in the Lower Silesian Coal Basin.

The average production per mine and the proportion of mines with a daily output of more than 15,000 tons has shown a continuous increase. This has been achieved from new mines, by reconstructing old mines, and by combining old mines into larger production units.

All Polish coal mines use both the advance and retreat longwall mining system. The output per face has increased from about 600 tons per day (with caving) in 1970 to 1,000 tons per day in 1977. The output per production level increased over this period from 2,300 tons per day to 4,000 tons per day. Ninety-five percent of the total coal industry output is achieved from mechanized faces.

The continuous development of output in the Polish coal industry is supported by several services and ancilliary enterprise as well as by the scientific research and design base. There are about 20 factories producing mining machinery and devices; they are grouped into one mining machinery enterprise and they ensure that the modern equipment for achieving our high level of coal output is available. The investment activity for the whole coal industry is undertaken by specialized enterprises, the mine construction enterprise and the coal industry surface facility construction enterprise.

The following activities are performed centrally for the whole industry

- Sale of coal

- Mine timber supply
- Extraction and supply of stowing sand to the mines
- Other materials supply

8.5.2 Performance since 1950

Tables 8.1 and 8.2 and Figure 8.1 show the increase since 1950 of the overall performance of the industry and the change in the size distribution of the mines. Six mines that can yield 12 – 24,000 tons per day (tpd) are under construction (Glanowski 1979).

TABLE 8.1 Productivity and labor intensity in Polish coal mines from 1950 to 1977.

	1950	1955	1960	1965	1970	1975	1977
Total output (10^6 tons)	78.0	94.5	104.4	118.2	140.1	171.6	186.1
Overall productivity (%)	100.0	113.2	126.9	156.1	197.8	326.7	343.4
Underground productivity (%)	100.0	127.1	135.4	161.8	231.5	286.9	300.9
Overall labor intensity (%)	100.0	88.2	79.9	78.8	38.9	30.6	29.1
Underground labor intensity (%)	100.0	78.6	73.8	73.2	43.2	34.8	33.2

Source: Mining Industry Statistical Data.

TABLE 8.2 Size, output, and employment of Polish coal mines from 1950 – 1977.

	1950	1955	1960	1965	1970	1975	1977
Number of mines	81	87	86	80	77	68	65
Number of mines with output over 10,000 tpd				2	7	13	14
Number of mines with output over 18,000 tpd							1
Output in total (10^6 tons)	78.0	94.5	104.4	118.2	140.1	171.6	186.1
Average output of one mine (10^6 tons per year)	963.0	1168.3	1287.3	1459.2	1819.5	2523.5	1863.
Total employment (x 10^3)	296.0	318.1	331.7	350.8	362.3	341.2	343.9
Employment industrial group (x 10^3)	285.6	304.2	312.4	324.2	330.5	321.6	326.0
Face output (%)	38.7	51.6	60.0	69.1	72.1	78.1	84.2

Source: Mining Industry Statistical Data.

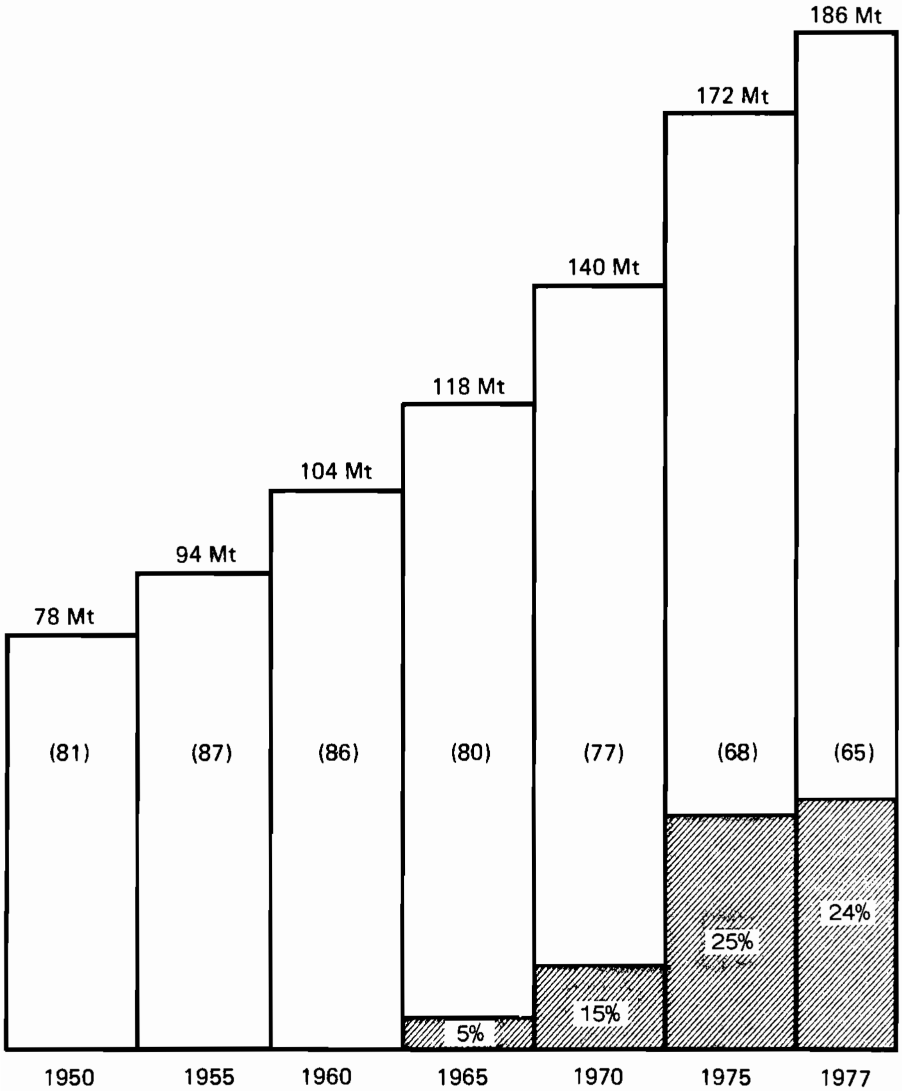


FIGURE 8.1 Changes in output, total number of mines, and proportion of output from large mines. Numbers in parentheses indicate number of mines. The shaded areas are the contributions from mines producing over 10,000 tpd.

The technological progress in the design of mines and the construction of mining machinery and equipment, the organizational progress in terms of our ability to organize production from large mines, and our experience in mine construction have led to the view that in the geological conditions of Polish coal basins it is advisable to design and construct large integrated mines of 15 – 20,000 tpd. Each mine consists of individual mining areas connected by common underground transport, but the mines are independently ventilated.

The optimal size of such a mine is determined by technological – economic criteria (e.g., minimization of coal mining costs) and the assumption that maximum use will be made of the technological means available.

However, while technological progress has made it possible to design and construct big mines, and modern and expensive equipment requires big mines to achieve high efficiency in coal production, new management and organizational issues arise in such mines.

8.5.3 The Relationship between Size and Performance

Existing mines differ in size because of the way in which mines are designed and developed. There are two reasons:

- Even though market needs, and the technological – economic, market, and social potentials are known, the quantity of production cannot be determined when the deposits are unknown
- Even when a sufficient deposit has been located, the output rate cannot be determined until the mining development has been designed

Table 8.3 shows the relationship between various indicators of mine performance and mine size. For clarity, only those indicators have been included that are most closely associated with size and performance. It can be seen that in our mining industry the relationship between mine performance and size is as follows:

- Total productivity of mines with output over 18,000 tpd is 18 percent higher than those with ouput under 10,000 tpd
- Underground output is 14 percent higher
- Energy consumption per ton is considerably lower
- Production costs are 33 percent lower.

Nevertheless, the range of mine sizes that show efficient performance is quite wide, though the range may be quite different if other variables are considered, such as management requirements and complexity. This leads to the need to consider not only the economic effects of mine size but also the requirements on management.

8.5.4 The Effect of Mine Size on Mine Management Process Requirements

The second aspect of our study was to investigate more thoroughly our thesis that differences in mine size lead to differences in the degree of difficulty in the decision-making problems confronting managers, and these differences have a significant effect on their performance as managers.

In thinking about the different types of difficult decision problems that confront managers, we found it desirable to classify them in the following way.

1. By the manager responsible for dealing with the problem (general manager, chief engineer, deputy director for economic matters, deputy director for personnel)

TABLE 8.3 Performance versus size in Polish coal mines (%).

Output (tpd)	Productivity		Electric power consumption	Average employment rate in mine			Production costs				
	Underground	General		General	Industrial group	Nonindustrial group	Total	Labor	Materials	Depreciation	Power
18,000 - 24,000	100	100	100	100	86.56	4.44	100	100	100	100	100
16,001 18,000	89.58	91.8	63.51	100	86.58	4.46	113.18	112.92	114.91	85.31	97.1
14,001 - 16,000	94.43	96.5	67.89	100	86.01	4.98	127.85	120.84	147.70	96.23	111.01
12,001 14,000	94.76	94.91	74.50	100	85.76	5.24	126.52	122.91	123.14	95.42	109.54
10,000 - 12,000	90.05	90.83	84.20	100	85.80	5.30	128.82	128.54	128.35	99.37	113.52
below 10,000	82.06	86.17	152.0	100	86.13	4.90	132.85	137.78	133.20	100.21	116.04

SOURCE: Mining Industry Statistical Data.

2. By whether the problem arose from within the mine or from the external environment
3. By where the source of the problem was in the technical and organizational system

For (3) we found it appropriate to use a classification developed in Dobrov (1978) and further refined in Dobrov *et al.* (1979). These authors view a technological system as three components: hardware, software, and orgware. In the context of a mine:

Hardware consist of the materials, machines, and other equipment required to extract, transport, and process the coal. The structure of mining activities is also considered to be part of the hardware.
Software consists of the plans, procedures, methods, know-how, and skills associated with the coal mining process. We also consider it to include the organizational climate.
Orgware consists of the more formal and prescribed organizational arrangements that enable the technical, human, and institutional factors to interact, not only within the mine but also with the external environment.

8.5.5 Preliminary Survey

First of all, we conducted a preliminary survey of a group of decision makers in order to develop a list of difficult decision problems. We interviewed managers from 25 collieries with wide experience and who were associated with mines of different sizes. We classified the mines into three size groups: group I - small mines (output less than 10,000 tpd.); group II – medium mines (output 10,000 – 18,000 tpd); group III – large mines (output more than 18,000 tpd).

As a result of the interviews we developed a list of 33 different difficult decision problems. We asked the managers to rate each problem by (a) frequency in the last year: once (1), occasionally (2), frequently (3), and particularly frequently (4); and (b) degree of difficulty on a scale of 0 to 4 (0 – not very difficult, 4 – particularly difficult). Tables 8.4a and 8.4b show examples of difficult decision problems and the responses by managers. The problems in table 8.4a arise from the environment and the problems in table 8.4b arise from within the system.

The results of the analysis for decision problems arising from the mine environment were

	Average frequency of occurrence	Average degree of difficulty
Small mines	1.5	1.8
Medium mines	2.0	2.1
Large mines	2.2	2.5

TABLE 8.4a Some decision problems associated with the environment.

	Group of mines					
	I		II		III	
	Frequency	Degree of difficulty	Frequency	Degree of difficulty	Frequency	Degree of difficulty
Change of the production plans during the year (hardware)	3	3	3	4	4	4
Rational utilization of machines and equipment. Increase of breakdown of technical equipment owing to shortage of spare parts and irregular supplies of machines, equipment, and materials (software)	4	4	4	3	4	3
Providing an appropriate social infrastructure (orgware)	2	2	3	3	4	4

TABLE 8.4b Some decision problems connected with the system(s).

	Group of mines					
	I		II		III	
	Frequency	Degree of difficulty	Frequency	Degree of difficulty	Frequency	Degree of difficulty
Serious worsening of mining and geological conditions at the mine level with the highest production capacity (hardware)	2	2	3	3	3	3
Propable failure of qualitative or quantitative plan either as a whole or according to types and assortments (software)	3	3	3	4	3	4
Material procurement (orgware)	2	2	3	3	4	4

For decision problems arising from the system:

	Average frequency of occurrence	Average degree of difficulty
Small mines	1.3	2.0
Medium mines	2.3	2.2
Large mines	2.6	2.3

These results confirm that while the mines used similar technologies, and similar systems of production and organization, the difference in mine size results in a differentiation of decisions. The deputy directors of personnel who were interviewed mentioned that large mines resulted in particularly difficult decisions.

8.5.6 Factors Influencing the Decision Situation

In order to understand the nature of the difficult decision problems it is necessary to determine

- The substance of the problem
- The factors influencing the situation out of which the problem arises
- The strength of influence of the factors
- The ability of the decision maker to modify the influence of the factors

This ability of the decision maker to modify the influence of the factors involved in a particular decision situation depends first of all on the subjective features of the decision maker (see Bartnicki 1979). Furthermore, some factors restrict the decision maker (limiting factors) while other factors help him (facilitating factors). So a list of factors influencing difficult decision-making problems was prepared and the factors classified into limiting factors and facilitating factors.

8.5.7 Determination of the Influence of Factors

The next stage of the research was to determine the decision makers' perceptions of the strength of influence of the factors on difficult decision-making problems and to relate them to the size of the mine. Prepared interviews were used in which the decision maker was asked to rate the strength of influence of each factor on a scale from 0 (slight influence) to 4 (particularly strong influence)

The analyzed results are shown in Tables 8.5a – c and 8.6 a – c. Table 8.5. shows the limiting factors, subdivided into (a) hardware, (b) software, and (c) orgware, and Table 8.6 shows the facilitating factors.

The average influence of limiting factors was:

TABLE 8.5 Degree of influence of some limiting factors associated with (a) hardware, (b) software, and (c) orgware.[a]

		Group of mines		
(a)	Hardware	I	II	III
	Shortage of necessary material supplies	2	3	4
	Increase in proportion of total output from safety pillars	4	2	1
	Quantity and quality of hardware	1	3	4
	Mechanization and automation of production process	1	2	3
	Waste utilization	1	2	3
(b)	Software			
	Inefficiency of computerized information systems	1	2	3
	Inflexible organization of the unit imposed by the superior authorities	1	3	3
	Lack of experienced staff sections	1	3	4
	Failure to adjust control systems	2	3	3
	Frequent modifications of regulations and standards	1	1	1
(c)	Orgware			
	Low skills of personnel	1	2	3
	Short time horizon of the evaluation system	1	2	3
	Lack of responsibility initiative of:			
	– staff	1	2	2
	– workers	2	2	2
	Operation of dispatching centers	2	3	4
	Varying expectations of external authorities	3	4	4

a 0: slight influence; 4: particularly significant influence.

TABLE 8.6 Degree of influence of some facilitating factors concerning (a) hardware, (b) software, and (c) orgware.[a]

		Group of mines		
(a)	Hardware	I	II	III
	Relatively easy acquisition of additional material supplies	2	1	1
	Relatively easy increase in production capacity of the working front	2	3	3
	Quantity and quality of resources	1	2	3
	Relatively easy acquisition of additional technical resources	2	1	1
(b)	Software			
	Legality of activities	1	1	1
	Engagement of the environment to accomplish. the program to which the decision is pertinent	1	2	2
	Computerization of management	1	2	3
	Proper measures for work evaluation	3	3	4
	Relatively easy distribution of resources and planning	2	3	4
(c)	Orgware			
	Liability of the environment to requirements of mine management	1	2	3
	Engagement of persons taking part in the decision-making process	3	3	3
	Appreciation of the preliminary stage of activities	2	3	4
	Competence of decision-makers	2	3	4
	Organization climate	1	2	3

a 0: Slight influence; 4: particularly significant influence.

	Hardware	Software	Orgware
Small mines	1.7	1.9	1.3
Medium mines	2.3	2.5	2.5
Large mines	2.5	2.7	2.8

The average influence of facilitating factors was:

	Hardware	Software	Orgware
Small mines	2.0	1.9	1.7
Medium mines	2.0	2.5	2.1
Large mines	2.0	3.0	2.3

It can be seen that as the size of the mine increases the influence of the limiting factors increases substantially. Combined with the increasing frquency of difficult decision problems, this means that further increase of colliery size could result in a reduction of their efficiency (see Figure 8.2).

A comprehensive study of technical and technological aspects of scale in the mining industry, combined with an investigation of the organizational, managerial, and environmental aspects should help answer the questions: What size of colliery ought one to design and build in the 80s? How do we design organizations in the mining industry that are appropriate to the size of the mine?

In the future we plant to examine more closely the strength of influence of particular factors in a variety of different decision-making situations. We intend to analyze how decision makers behave and how they modify the effect of the key factors. This should lead to guidelines for planning and modernizing the management system in the mining industry.

The results obtained so far must be considered as preliminary and incomplete although they do confirm our thesis and provide a basis for further studies.

REFERENCES

Bartnicki, M. 1979. Metoda Analizy pola sił w Badaniach Procesów Decyzyjnych (The Method of Strength Field Analysis in the Decision Making Process Investigation). Problemy Organizacji. Nr. 1 (in Polish).

Caplow, T. 1957. Organizational Size. Administrative Science Quarterly 1.

Dobrov, G.M. 1978. A Strategy for Organized Technology. In Dobrov, G.M., R.H. Randolph, and W. D. Rauch, eds., Systems Assessment of New Technology: International Perspectives. CP-78-8. Laxenburg, Austria: International Institute for Applied Systems Analysis. pp. 13 – 30.

Dobrov, G.M., M. McManus, and A. Straszak. 1979. Management of Technological Innovations Toward Systems-Integrated Organized Technology. CP-79-6. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Doktór, J. 1978. Struktury Organizacyjne a Efktywność w Przemyśle Chemicznym (Organizational Structures and the Efficiency in the Chemical Industry). Gliwice: Instytut Ekonomiki Przemysłu Chemicznego i Naczelna Organizacja Techniczna.

Glanowski, M. 1979. XXX Lat GBSiPG w Służbie Polskiego Pzemysłu – Projekty-Problemy Budow-

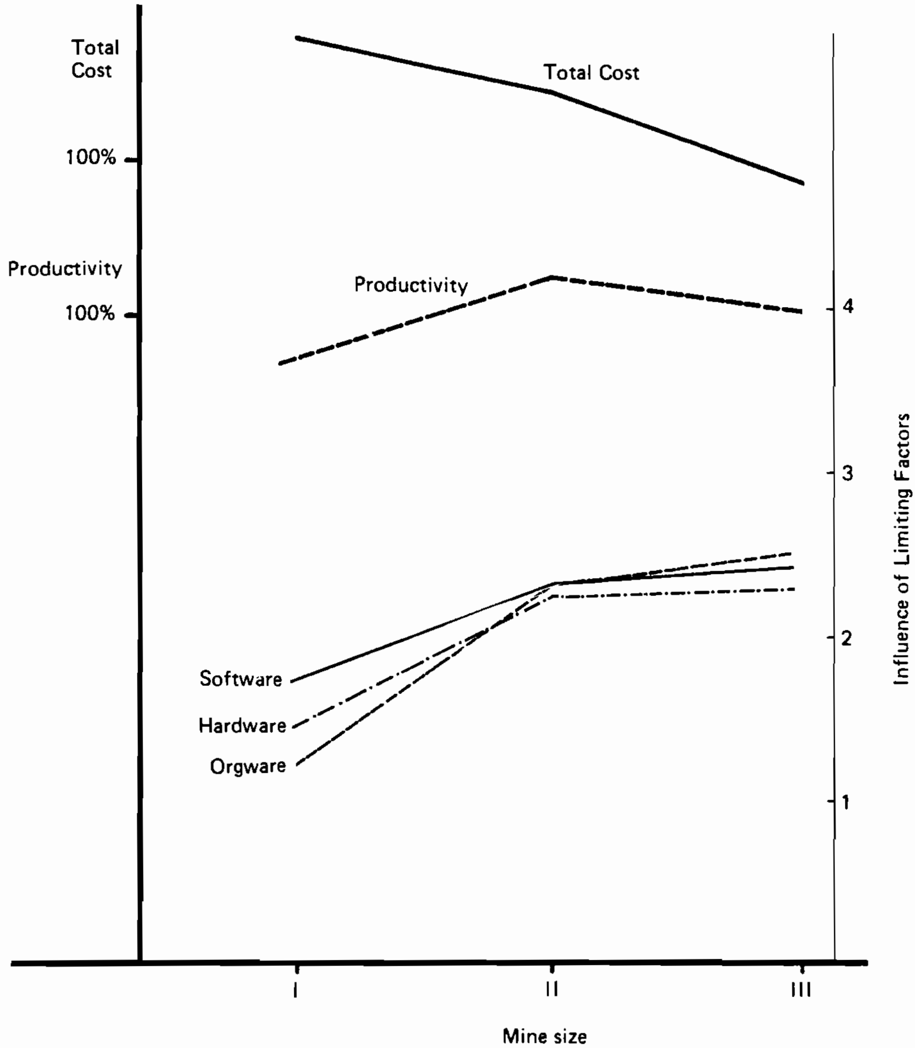


FIGURE 8.2 Effect of mine size on productivity, cost, and influence of limiting factors.

nictwa Węglowego (Thirty Years of the Main Study and Design Office for the Mining Industry Activity in the Polish Mining Industry). No. 1. Katowice. (in Polish).

Gliński, G. 1971. Socjalistyczne Przedsiębiorstwo Przemysłowe (Socialistic Industrial Enterprise). In: Ekonomika i Organizacja Przedsiębiorstwa Przemysłowego (Economics and Organization of Industrial Enterprise), Chapter i. Warsaw: PWE.

Khandwalla, P.N. 1977. Design of Organizations. New York: Harcourt Brace & Jovanovich, JNC.

Melcher, A.J. 1976. Structure and Process of Organizations. A System Approach. Englewood Cliffs, N.J.: Prentice-Hall.

Pańkow, W. 1979. Uwarunkowania Struktur Organizacyjnych a Podejście Sytuacyjne (Organizational Structure Conditions on the Basis of the Contingency Approach). Problemy Organizacji. 1.

Zawiślak, A.M. 1975. Szkice o Zarzadzaniu (Essay on Management). Warsaw: PWN.

CHAPTER 9 MAIN ASPECTS DETERMINING THE SCALE OF AN ORGANIZATION – A FIRST TENTATIVE PROBLEM ORIENTATION

Rudy van Hees and Friso den Hertog
Philips Gloeilampenfabrieken,
Eindhoven, The Netherlands

9.1 INTRODUCTION

The aim of the paper is to consider those criteria and conditions according to which the scale of an organization unit can be determined and assessed. The paper is a result of an analysis done by an interdisciplinary task force within Philips.

The status of the report as presented in Laxenburg in June is an orientative one and was therefore written as a discussion paper; it may not be considered as a well-defined policy of our company. Now, half a year later, the analysis has been worked out further so that field tests can be made based on an analytical procedure.

R.N. van Hees (Department of Organization and Efficiency) and J.F. den Hertog (Department of Social Research) summarized the discussion in the task force. Other members were: R. Williams and C. Buitenhuis (Department of Organization and Efficiency), J. van Ham (Department of Social Research) and J. van Dam (Department of Planning and Resource Allocation).

Organization size is a mystifying factor in organization design. For years big meant better and big and business were associated with each other. These phrases had a strong appeal for decades. Economy of scale was the basic idea behind this way of thinking Big. Schmenner (1976), however, points out: "The phrase draws nods of recognition, yet few managers can define it successfully and fewer still have critically thought about it."

The bigness thesis seems recently to have lost its power to its antithesis: Small is beautiful. Smallness is associated with (Schumacher 1973):

- Flexibility
- Creativity
- Innovation
- Care for people

Both slogans add little or nothing to the management process (Cason 1978). The discussions about big and small often result in a cacaphony. A realistic synthesis is needed.

9.1.1 Finding the right size

Big is not necessarily better, nor is small always beautiful. Our aim in this respect is to help management in its task to take the scale question in consideration again. In the concrete situation, the right mix has to be found: the incorporation of the advantages of small scale within the context of a larger system.

9.2 THE PROBLEMS OF LARGE DIVERSIFIED FIRMS

Economy of scale means that an activity's required input per unit output is lower for a big activity than for a small one. This is often expressed graphically as in Figure 9.1 (Cason 1978) where unit cost refers to labor cost, machine cost, and material cost.

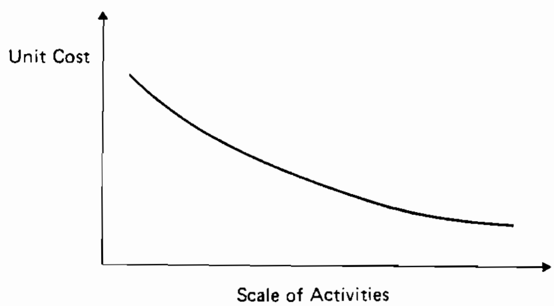


FIGURE 9.1 Relationship between unit cost and scale of activities. The unit cost is input cost per unit output.

The increase of size however may confront organizations with a more than linear increase in *complexity*. (Complexity refers to differentiation into units, diversity of products and processes, interactions required, interdependencies, and so on.) This increase in complexity causes diseconomies of scale, or, as we like to call it: organization cost. Both tendencies can be put graphically in one picture (Figure 9.2).

The problems of complexity as associated with size become manifest in connection with:

- The infra-structure at levels above the unit under consideration
- Coordination of activities
- Logistics (networks with more products)
- Ability to make decisions
- Motivation of workers (commitment, identification, understanding of the total operation)

The problem is especially relevant for large diversified firms. New answers and new organization philosophies have to be developed to deal with the problem of size. In General Electric the new approach has been labeled "The Strategical Business Unit-approach (SBU)." Hall (1978) states in this respect: "Diversified firms suffering from profitless growth are moving away from traditional planning and adopting the strategic business unit approach to overall corporate strategy." Shell gives another example of a new approach: the development of smaller decentralized, ROUND (relatively self-supporting and self-steering) units.

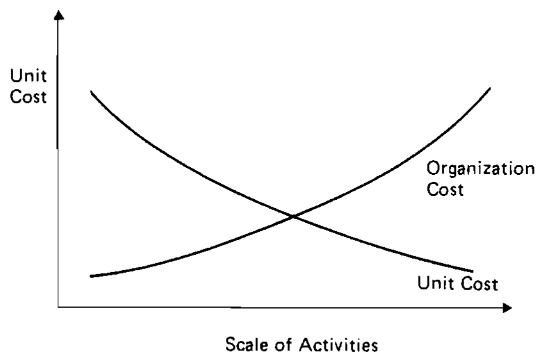


FIGURE 9.2 Organization cost and unit cost as a function of scale of activities.

Recently, in addition to "organization control," another criterion has been added to the question of scale: the quality of working life (QWL). Studies from Sweden (Job Reform in Sweden 1975) make clear that the idea of "the factory in the factory" is one of the basic conditions to enhance the QWL. Philips finds itself in this situation.

Managers do not deal with the problem in abstract discussions about organization scale or complexity. They are faced with concrete questions like:

1. What to do with small units that are altogether strong and well functioning internally, but that are so hard to coordinate on a next higher level in the organization?
2. What to do with big units suffering from high organization costs owing to complexity? (Make them smaller, or factory in the factory.)
3. What to do in this respect with small units that are becoming larger and big ones that by technological developments become smaller?

9.3 LOOKING FOR ANSWERS

In the search for a more adequate frame of reference in dealing with these questions, two areas of experience within the company were relevant and useful:

The efforts in the field of job and organization design work structuring. These efforts show how organization scale is related to the internal organization.

The experience within Philips with a wide range of products, markets, and technologies, showing that the question of scale is strongly related to the (partly external) conditions under which the organization operates: the organization domain as Galbraith (1977) calls it.

We would like to elaborate on both.

9.3.1 Work Structuring and Scale

In dealing with questions of scale and complexity much can be learned from the experience with work structuring at the shop floor and factory level of the organization. Work structuring (or job redesign) refers to the efforts within Philips to create a match between the needs and capacities of workers on one hand and the content and organization of work on the other.

The development of small product-oriented semiautonomous groups proved to be especially effective (and economically justified) in terms of flexibility, commitment, and job satisfaction.

Our experience and field research in this area indicates, however, that:

1. Fostering autonomy on the shop floor level also demands on a higher level in the organization a certain degree of autonomy and integration of functions. Redesign of jobs often requires also a redesign of organization structures at a higher level.
2. The development of decentralized and semiautonomous organization units demands that certain functions be well taken care of at a higher level (the information system, internal traffic, and so on. So organization cost on the next higher level is also affected).
3. Principles followed at shop floor level are also often applicable at higher levels of the organization.

9.3.2 Conditions

The management in a firm with such a large variety of markets, products, and technologies learns that under certain conditions the advantages of small (or large) scale are greater than in others.

Translated into more abstract terms, this means that we have to find out under which organizational conditions the curve of organization cost tends to be steep or relatively flat. That is the basic question in our analysis: What are the organizational and environmental contingencies that determine the right (or optimal) size? This question can be illustrated as in Figure 9.3.

Our basic proposition in this respect is that we can find conditions (A), where the process, product, and market are such that organization costs are rising fast when the scale of activities increases and that we can find conditions (C) where this rise in organization size causes a less drastic increase in organization cost.

On the basis of our own experience and a survey of the literature we developed a tentative list of conditions relevant to the question of size and complexity (Table 9.1). In other words, the table shows which conditions make the organization cost curve steeper or flatter.

The relations between conditions (or domain), the internal organization structure, organization cost, and scale can be visualized as in Figure 9.4. The procedure for field

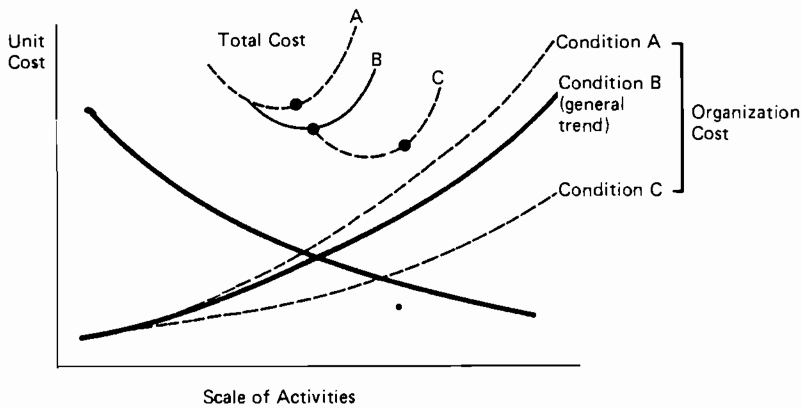


FIGURE 9.3 Relationship between total cost and scale of activities for different organizational conditions.

TABLE 9.1 Conditions favoring small-scale organization units.

Aspect	Differentiation	Condition in favor of small scale	
Process	Man/machine intensive	Man intensive process	
	New/grown up	New process (technology not yet controlled)	
	Critical mass of innovation potential	Low critical mass	
Product	Catalogue/to customer spec's	Product to customer spec's	Quick reaction to market required
	Begin/end life cycle	Product at beginning of life cycle	
	Fashion/nonfashion	Fashionable product	
Market	Turbulent/stable	Turbulent market	
	Marketing cost high/low	Low marketing cost	Flexibility required
	Importance of market share	Little importance of market share	
	Transportation cost	High transportation cost	

tests consists of the three steps described in Figure 9.4.

1. Fixing of the technical economic optimum (1) within a set of boundary conditions (2), of market, process, and product, which gives solution (A).
2. Design of a motivating structure (3) within (A) giving solution (B).
3. Test of the consistency of (B) under the conditions of (1). If not, then restart the cycle with a new or modified (A).

The argument behind this scheme is that in discussions about scale in practice, scale is often only related to an isolated and very limited set of parameters. Scale is, for example, only related to unit cost or to market considerations. It is our opinion that the question of scale has to be answered in an integrated way, taking into account the internal structure and the conditions as well as the total cost.

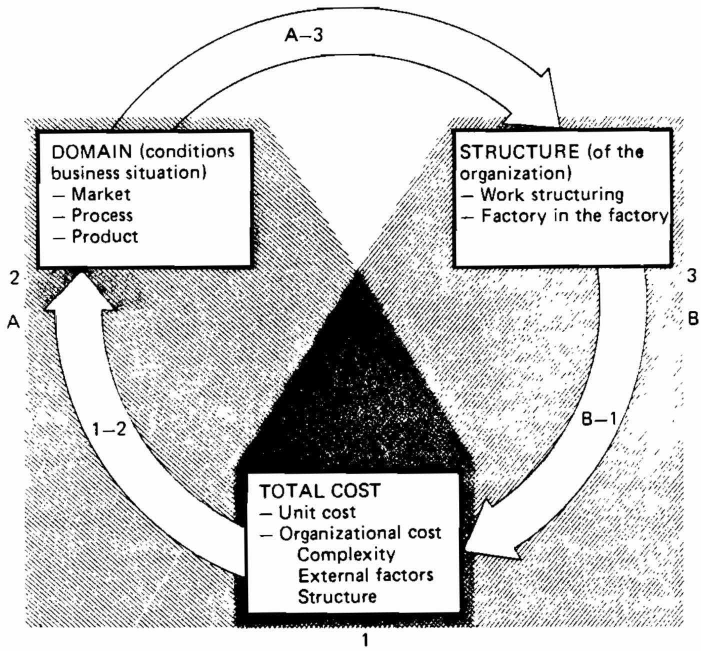


FIGURE 9.4 Unravelling of the problem by solution in stages.

REFERENCES

Cason, R.L. 1978. The Right Size: An Organizational Dilemma. Management Review. April: 24 – 39.

den Hertog, J.F. 1978. Arbeitsstrukturierung. Bern: Huber.

Galbraith, J.R. 1977. Organization Design. Reading: Addison Wesley.

Hall, N.K. 1978. SBU's: Hot, a new topic in the Management of Diversification. Business Horizons. February.

Job Reform in Sweden. 1975. Stockholm: SAF:

Schmenner, R.W. 1976. Before You Build a Big Factory. Harvard Business Review. July – August: 100 – 104.

Schumacher, E.F. 1973. Small is Beautiful. London: Blond & Briggs.

CHAPTER 10 SCALE ECONOMIES – THE EVIDENCE FROM PUBLISHED REPORTS OF THE BRITISH PRICE COMMISSION

M.F. Shutler
Monopolies and Merger Commission,
*London, UK**

One of the workshop aims is to identify from the practical experiences of participants "what can be learned about, e.g., technology, organization, management, and control and the influence of size" and another is to inquire whether there are "lessons that one industry can learn from another." The aim of this paper is to derive evidence from the reports of the British Price Commission published up to the date of the workshop (June 1979).

This evidence is introduced diffidently because it was not obtained as a result of directed research, but selected for this workshop from studies already carried out. The aim is of course to aid other researchers. Full copies of reports can be obtained from Her Majesty's Stationery Office, P.O. Box 569, London SE 1, but it would be dangerous to generalize beyond the industries studied.

The British Price Commission was established in 1973 to undertake, according to fairly detailed rules, control of the prices charged by major companies in the UK. In July 1977 its role became more investigative and many detailed "allowable cost" formulae for price rises were discontinued. Instead, it was given power to delay prenotified price increases for up to 3 months, provided that it carried out an investigation of them. It could also be directed by the Secretary of State for Prices and Consumer Protection to examine prices, costs, and margins in complete industries. Such investigations and examinations were carried out in accordance with criteria that implied that the commission had to judge the efficiency with which goods and services were supplied; a difficult enough task to carry out in a 3-month period. All the companies investigated were "large", measured by either turnover capital employed or labor force size. The average turnover was about 1.5 billion dollars per annum. They ranged from the engineering sector to retail distribution and public utilities. None, however, could be classified "advanced technology."

This paper gives my own personal views, not the results of deep specific research. It will probably, therefore raise more questions than it answers. The first topic concerns risk.

The papers submitted by IIASA in advance of the workshop listed risk as a possible diseconomy of "level 2" scale. In many industries there is a risk that industrial action at one or more key sites may disrupt the whole of a firm's output. Electricity generation, automotive spare parts, and brewing are cases in point.

---

*Formerly with the Price Commission.

The decision not to decrease the number of plants, giving up the scale economies of one large plant, is a serious consideration, but in the European car industry as a whole, duplicate sources of components are now the rule in order to keep production of the main assembly plants going in the event of industrial action at one supplier.

I would like now to turn to the evidence from published Price Commission reports for the existence of positive plant scale economies in the UK, either achieved in the recent past or remaining to be achieved now. In assessing the conclusions it should be noted that inevitably subjective judgements have been made.

Figure 10.1 lists the industries in which some or all the firms were studied and reports published. Since the Car Spare Parts Report had not yet been published when this paper was prepared, it was not included in the analysis.

Animal Feeding Stuffs	Hotels and Catering (2)
Banking	Industrial Gases (2)
Batteries	Metal Cans
Beer	Metal Windows (8)
Buses	Natural Gas Supply
Car Spaces (50+)	Oil (2)
Cement (2)	Pharmaceuticals (10+)
Chemicals	Plasterboard
China	Power Tools (3)
Diesel Engines	Printing and Publishing (2)
Electric Power Generation & Distribution	Road Haulage (100+)
Estate Agents (50+)	Spectacles
Fertilizers	Taxis (10+)
Food Products (8)	Tobacco (2)
Glass (2)	Water Supply

FIGURE 10.1 Industries studied by the British Price Commission. Figures in parentheses give the number of firms examined or investigated in each industry up to July 1979.

Nine industries were examined as a whole. In only one were potential economies of plant scale reported as still to be achieved. This conclusion is nevertheless somewhat negative; it cannot be positively asserted that more were possible since in some cases questioning of firms on this point was oblique rather than direct.

In the case of investigations of individual companies where full efficiency studies were undertaken, unexploited plant scale economies were reported in only 3 out of 35 cases, but in 15 cases such economies had been achieved in recent years. It should be remembered that the British market is small, serving some 50 million inhabitants, but even so, increased transport costs tended to dominate potential production economies from increased plant scale. This was of course the case particularly where the product had a low value-to-weight or value-to-volume ratio, e.g., cement, plasterboard, glass, metal cans. Hence, in these cases, even in the UK market small plants are replicated and placed near to customers' or suppliers' works.

This has happened despite the considerable rationalization of distribution during the late 60s and early 70s, resulting in decreased numbers and increased size of depots. It was brought about by the development of motorways and by the increasing size of lorries, which itself led to a 75 percent saving in costs per ton/mile. This process is, however, probably at an end for environmental reasons. Furthermore, the current increase in oil prices will reinforce the domination of marginal transport cost increases over marginal production cost savings.

The potential for exporting would of course alter the argument, but in my experience even firms expanding into overseas markets tend to replicate their plants into these markets, rather than expand their UK plants.

Exploiting plant scale economies is not the sole means of improving productive efficiency, and indeed we found that 62 percent of the firms had a history of steady self-improvement in labor productivity both in manufacturing and distribution or in capital productivity in distribution, or in reduced raw material consumption per unit of output. These are perhaps "level 3" scale economies as defined in the introduction to the workshop.

To explain this fully would need much more research, but there appear to be some indications set out in Figure 10.2. A history of steady self-improvement by a firm appears to be associated with the existence of a management control system based on:

Formal cost reduction targets or profit improvement plants
A firm's being part of a foreign-owned or British-owned international group
A firm's being judged efficient by the Price Commission

By "British- or foreign-owned international group" in this context, I mean what is often referred to as a "multinational." Shell Petroleum is an example of a member of a British-owned international group. Esso is a member company trading in the UK of the foreign-owned international group Exxon Corporation. In the case of both types of group, only the UK member company was investigated, although in many cases, through the good offices of the parent company, the Price Commission was able to obtain data on members of the group in other countries in order to assess efficiency by international comparison.

This argument could of course be circular if the Price Commission had judged efficiency solely by the existence of a system of target setting. I would argue that this was not so. In many of these multinational groups there appeared to be effective competition or rivalry between the different national units. However, when such units are big enough to dominate a national market, it does raise the question of whether competition between managers in multinationals can be an effective substitute for price competition in the marketplace in promoting productive efficiency.

Figure 10.3 shows the division in another way and brings out more clearly the association of membership of a multinational group, either British- or foreign-owned, the existence of cost reduction programs, and the commission's assessments. In some cases for the purposes of this analysis, the weight of the evidence contained in the commission's published report had to be judged so that the scale does in fact tip one way or another. It will be seen from Figure 10.3, however, that a large number of mistaken judgements would have to be made to affect the hypothesis.

It may be the case of course that it is not a corporate structure effect which we are observing here but one of sheer size. I do not have the turnover figures for the foreign parents of the multinationals, bit the UK turnover figures for the companies judged efficient have a mean of 1,575 million dollars and a standard deviation of 2,128 million dollars. The other companies have a mean and standard deviation of 1,344 million dollars and 1,925 million dollars.

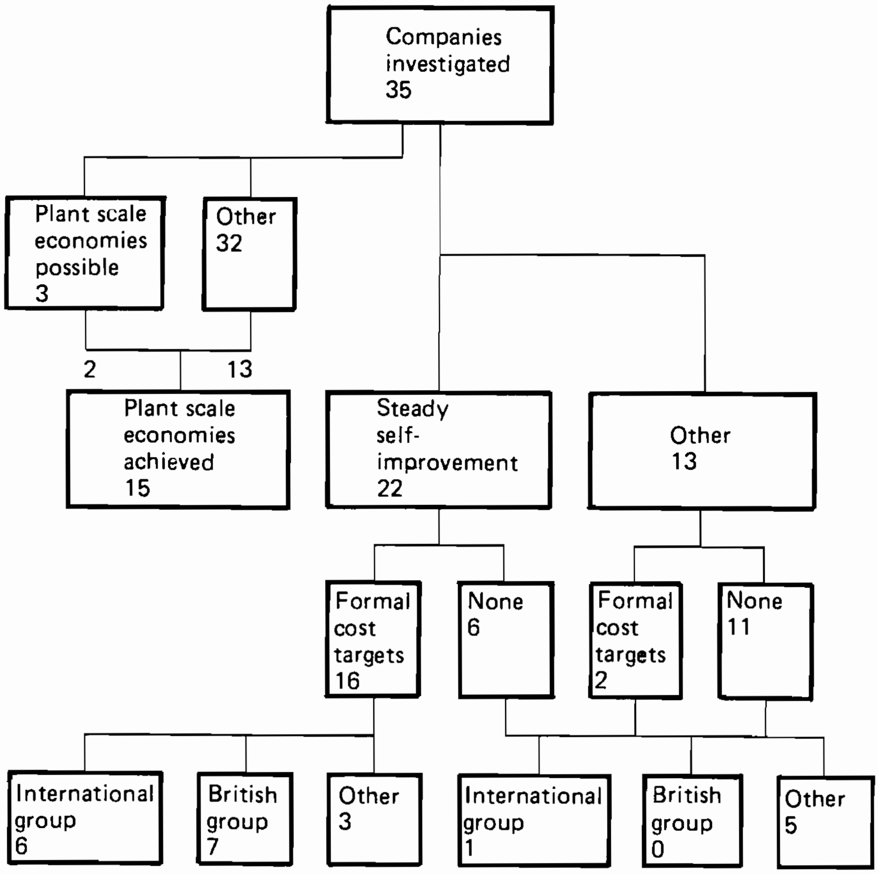


FIGURE 10.2 Distribution of companies according to methods used for improving productive efficiency.

Finally, one result of this conference that I should like to see is that we should cease to talk about economies of scale *per se,* since to do so leaves out of account too many other factors. Should we not instead talk of optimal plant scale and the factors that determine it?

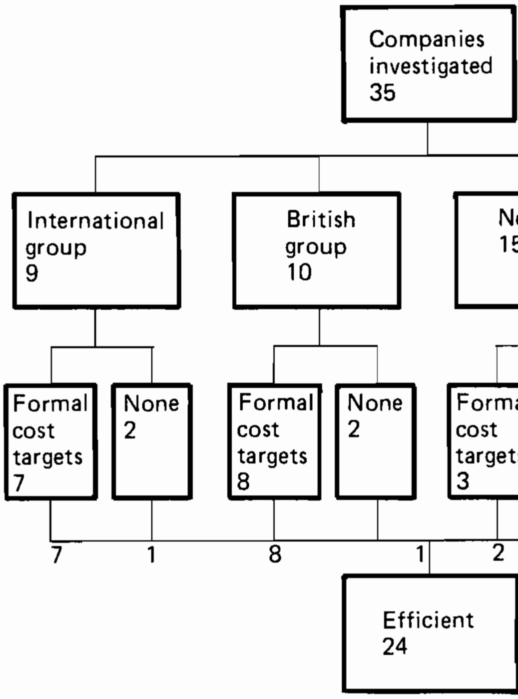


FIGURE 10.3 Effect of belonging to a multinational group on the existence of formal cost targets and efficiency. Distribution of companies with regard to formal cost targets and efficiency.

CHAPTER 11 PROBLEMS OF DETERMINING PRODUCTION SCALE IN SOVIET INDUSTRY

A. Egiazarian
Faculty of Economics,
Moscow University, USSR
V. Glagolev
International Institute for Applied Systems Analysis,
Laxenburg, Austria

11.1 INTRODUCTION

It is very significant that problems of scale have been studied by the Management and Technology Area at IIASA. There is such a strong interconnection and interdependence between scale, production management and organization, and technological progress that no problem in one area can be solved without taking into account the others.

In the USSR particular attention is being paid to the problems of the scale of enterprises and the improvement of their efficiency. This is being done in the context of improving the management system and organizational structure of industry on the basis of the development of production concentration, specialization, cooperation, and the use of achievements in science and technology.

Advances in technology have influenced production concentration, specialization, and cooperation in the following ways:

The development of specialized machinery, instrumentation, and methods of automatic control that yield high productivity requires concentration of production of homogeneous items so that the high capacity and productivity of such machines can be utilized optimally

The increase of the complexity of products and the growth of the number of their components requires close interconnection between the groups of plants producing components and subassemblies for such products.

But it is necessary to stress that the extent to which technological advance and production concentration create possibilities for performance improvement is determined by whether the production system has the organization, specialization, cooperation, and scale appropriate to the level of economic and technological development.

The main problem is to develop and establish organizational forms, structures, and management systems in industry that permit the fullest use of the possibilities. The system of management, organization, and scale determination must be in accordance with developments in production and process, equipment, science and technology, and the level of education and professional skills of employees.

Prior to the 1970s, the traditional structure of Soviet industry had as its basic unit the plant. However, the plant was separated by up to five intermediate administrative levels from the ministry where overall policy was determined. To match the present stage of social and economic development of industry the following forms of industrial concentration and specialization have been developed:

1. Production has been integrated by combining plants into multiplant complexes or production associations, which has become the basic unit of industry instead of plants. A two-level system of management (ministry–production associations/enterprises) was established on this basis, with production associations directly responsible to the ministry.

2. As intermediate administrative levels (glavk, kombinat, and so on) have been abolished on the middle level of ministerial management, in some cases industrial associations were established which are responsible for coordination of a subsection or sector of industry. This results in a three-tier system of industrial management (ministry – industrial associations – production associations/enterprises). In either case production association is now responsible for both ongoing operations and strategic planning.

3. Regional concentration of production has been achieved by forming production complexes (such as the Bratsk-Ilimsk territorial complex (Knop 1977, Knop and Straszak 1978)) and production centers. A production center or a small production complex can be defined as a set of industrial enterprises that are located on a compact site (50 – 100 ha) and have a common infrastructure (water and power supply, transport, communications, and so on). This method of industrial location has been widely used in the USSR. Of the total investment in industrial construction, 10 percent goes into such production centers in the USSR overall and 30 percent in the Lithuanian, Byelorussian and Moldavian SSRs (Economics of Construction 1976).

All three forms of production concentration and specialization have an effect on the scale of production but the first form has the greatest influence. Thus, particular attention will be paid to it in this paper.

11.2 HISTORICAL BACKGROUND FOR THE DEVELOPMENT OF THE PRODUCTION ASSOCIATION

In order to understand why production associations emerged it is necessary to review the history of production concentration, specialization, and scale in Soviet industry. Until 1973 the process of production concentration and growth in the size of enterprises occurred within the traditional Soviet framework of the basic industrial unit: the plants. The data in Table 11.1 show that between 1964 and 1973 the contribution of big enterprises increased substantially, irrespective of whether their contribution is measured in terms of number of enterprises, value of gross output, number of employees, or value of fixed assets employed.

In Table 10.1 enterprises are classified in terms of their gross output. This is the usual classification adopted in the USSR. An alternative classification of interest to scientists from other countries is a classification by number of employees. Table 11.2 shows that, while the distribution of enterprises by size class remained relatively stable, the

TABLE 11.1 Distribution of production enterprises in Soviet industry by value of gross output (1964 and 1973).

Size class	Value of gross output of enterprise (10^3 roubles)	Percentage of total							
		Number of enterprises		Value of gross output		Production personnel		Cost of fixed assets employed	
		1964	1973	1964	1973	1964	1973	1964	1973
I	up to 500	27.1	20.7	1.3	0.5	3.3	1.8	2.4	1.0
II	501 – 5,000	55.4	50.5	22.5	11.9	32.3	20.5	25.6	12.8
III	5,001 – 10,000	8.5	12.6	12.8	10.3	13.7	13.4	13.7	10.6
IV	10,001 – 50,000	7.6	13.0	33.7	31.7	30.7	32.4	29.2	28.2
V	50,000 and more	1.4	3.2	29.7	45.6	20.6	31.9	29.1	47.4

Data Source: National Economy of the USSR (1964 - 1965, and 1973 – 1974).

TABLE 11.2 Distribution of production enterprises in Soviet industry by average number of workers (1964 and 1973).

Average number of workers	Percentage of total							
	Number of enterprises		Value of gross output		Production personnel		Cost of fixed assets employed	
	1964	1973	1964	1973	1964	1973	1964	1973
up to 100	33.9	35.0	4.6	4.2	3.5	3.4	3.4	2.9
101 – 500	43.7	42.5	21.6	19.9	21.7	19.4	16.4	15.2
501 – 1,000	11.8	11.3	15.0	14.4	16.2	14.9	15.0	13.2
1,001 – 10,000	10.3	10.9	48.1	49.9	48.3	50.7	50.8	52.3
10,000 and more	0.3	0.3	10.7	11.6	11.3	11.6	14.4	16.4

Data Source: National Economy of the USSR (1964 – 1965, and 1973 – 1974).

share of big enterprise in value of gross output, number of personnel employed, and value of fixed assets increased. Note that over 75 percent of the total value of gross output was produced by enterprises employing more than 500 workers.

Of particular significance is the relationship between labor productivity and size of the enterprise. Figure 11.1 shows that labor productivity increases with the size of the enterprise. Also, over the period 1964 to 1973 the difference in labor productivity between large and small enterprises increased.

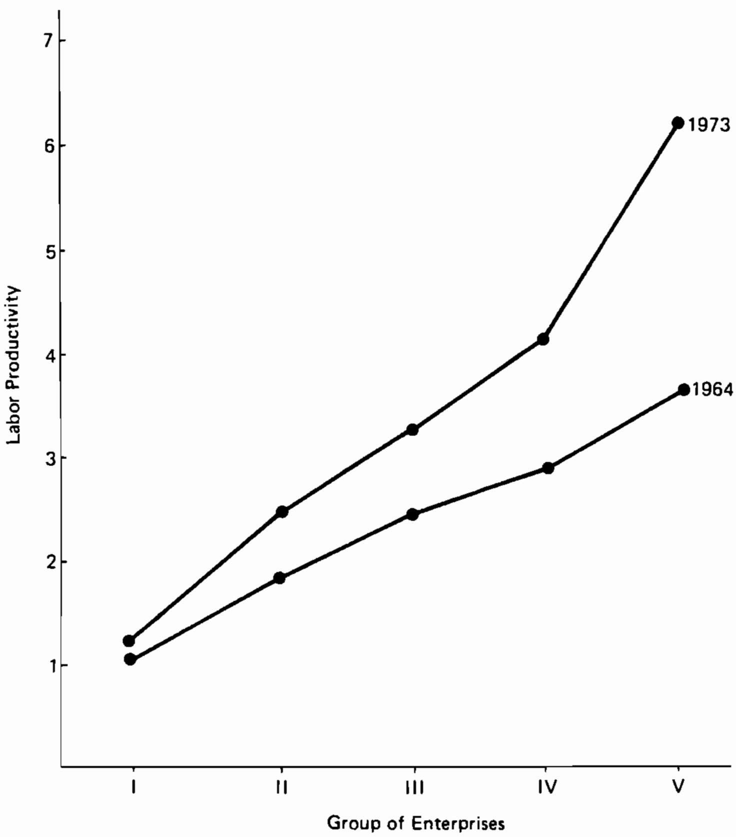


FIGURE 11.1 Relationship between labor productivity and the size of production enterprises in Soviet industry in 1964 and 1973. Data Source: National Economy of the USSR, 1964 – 1965 and 1973 – 1974.

Another performance index that is widely used in the USSR national economy is the ratio of value of sales to value of fixed assets. Table 11.3 shows that this index also icnreases with the size of the enterprise.

From this analysis it can be concluded that in the USSR big enterprises perform better than small enterprises. This statement is not a result of speculation; it is the result of economic analysis. Thus in order to improve the overall performance of the national economy it was concluded that the smaller enterprises should be combined to form larger enterprises. In 1973, 10,000 enterprises (21 percent of the total) had a gross output of less than 500,000 roubles and 3,000 of these had a gross output less than 100,000 roubles (National Economy of the USSR in 1973, 1974).

However, further analysis indicated that even if the small enterprises were combined into large enterprises the labor productivity would not necessarily be at the level achieved in other countries. There are a number of reasons for this but one in particular should be singled out: if production concentration is not accompanied by a high degree of specialization it is not possible to make use of the advantages of large-scale production, in particular the opportunity to use specialized high-production equipment, production mechanization, and automation.

In some branches of industry the proportion of general purpose plants that produce the whole range of components and subassemblies for products is relatively high. For example, in terms of the mechanical engineering industry, 80 percent of the enterprises are characterized by a broad range of products that are similar from the point of view of the customer's description but are different from the point of view of the design and other technological characteristics. Many enterprises have their own preparatory workshops. As a result, 70 percent of the workshops or production units in this industry have a size that is less than optimal (Problems of Economics 1977).

A further problem is that many industrial enterprises contained the full range of auxiliary services and facilities such as tool making, maintenance and repair of equipment, production of containers and packaging material, and spare part manufacture. This was recognized as an obstacle to the use of high-productivity equipment, production mechanization, and automation and hence to increasing labor productivity. Given that in Soviet industry the ratio of main to auxiliary workers is 1 : 1 there is an obvious benefit in developing concentration and specialization in auxiliary services and facilities (Glagolev 1975). Such a concentration was not possible within the traditional structure of industry. However, it was shown that by merging small and medium-size enterprises and then changing the organizational structure of production, fuller use could be made of the economic advantages of a socialist planned industry.

So, after a detailed analysis of the state of the art in industry, the Soviet government and Party adopted a resolution in 1973 under the title "On Certain Measures for the Further Improvement of Industrial Management." This resolution called upon the industrial ministries to work out general schemes of industrial organization and management for the whole of industry and for each particular sector (or ministry) with the goals of reducing the number of linkages in the management structures and establishing a system in which production associations would be the basic unit of industry.

It is necessary to stress that the need to seek out new methods of management and forms of organizational structure had been created by the rapid rate of development and the increasing complexity of industry. The dynamic growth of production systems and the increase in the scale of the national economy must lead to changes in the system of management and the organizational structure of industry. Without such adjustments the production system cannot function efficiently and effectively.

11.3 PRODUCTION ASSOCIATIONS

Production associations are a new development in the organizational structure of industrial production in the USSR. They are not just an agglomeration of enterprises but a coherent production--economic complex in which specialization and cooperation are strongly developed and science, technology, and production are closely coordinated and integrated.

The basis for establishing a production association is technological and production similarity combined with geographical proximity. They join together independent enterprises or plants connected by the vertical processing of raw materials to end products, by the similarity of product range, or in other ways. Associations may be created on an industry or an interindustry basis. They can include enterprises and organizations that fully retain their independence or enterprises and organizations that are turned into subsidiaries. Guidelines exist for determining the validity of establishing associations and the inclusion in them of enterprises and organizations. These guidelines take account of production management, organization, and economic points of view. Among the objectives of production associations one can single out some that have a strong influence on the determination of production scale; in particular

Concentration on homogeneous products
Development of a higher degree of specialization
Development of cooperation within and between associations.

In Soviet industry one can distinguish three main forms of production specialization based on production concentration:

Product specialization – concentrating on the production of similar products
Component specialization – manufacturing particular components or subassemblies as a distinct product line
Technological specialization – singling out a specific technological process or production technique as an independent enterprise

Increasing the degree of production specialization is inseparably linked with the development of cooperation between enterprises to produce a certain output. Cooperation makes it possible to better utilize the available production facilities of each enterprise and establish the most economical production links among all the enterprises that manufacture a certain type of output. As a result, in the operations of production associations special emphasis has been put on the scrupulous fulfillment of delivery plans because this is essential for the smooth operation of the whole industry and the efficient use of production facilities. Those associations and their enterprises that are engaged in mass and large batch production have switched to direct and long-term links. Furthermore, the size of economic incentive funds and bonuses depends in large measure on the fulfillment of delivery plans in accordance with contracts and orders.

At the beginning of 1979 there were 4,000 production and scientific-production associations and they accounted for 46 percent of the industrial output (Pravda 1979). The establishment of production associations in industry is to be completed in the tenth 5 year period (1980) (Kosygin 1976).

The combination of enterprises into production associations results in an increase in the total production capacity because fullest possible use is made of the production machinery and other facilities. However, the development of production associations does not mean that their production needs become the basis for optimizing the scale of plants. The main criterion for determining the scale of enterprises is national-economic: achievement in the interests of society of the best results at the least cost. In practice, optimization with respects to this criterion consists in the minimization of total annual costs or the present worth of total costs as described below:

Total annual cost

$C + E_N K \rightarrow \text{Min}$

Total cost on present basis

$K + T_N C \rightarrow \text{Min}$

where

C = Production costs of annual production output (in roubles)
K = Investment (in roubles)
$E_N = 0.15$ = Normative coefficient of efficiency

$T = \frac{1}{E_N} = 6.66$ (in years)

After establishment in accordance with a general scheme of the managing ministry for the industry, the production associations are initially only associations in name. The plants and production units forming the association operated in the same manner as prior to the creation of the association. However, the advantages from association cannot be realized until concentration and specialization of main and auxiliary production have been developed, that is, until there has been centralization and reallocation of managerial functions, technological and design services, planning and accounting, transport and communications, construction and maintenance, supply, and marketing.

To achieve this it is necessary to develop a detailed reorganization proposal. This requires redesigning the production – economic complex: developing a new production structure and organization and creating a new management system. Sometimes this is connected with a change of the mix of products of the production units belonging to the association.

It must be understood that alternative solutions exist to solve this problem. Several variants of production association must be developed and evaluated. One of them must then be selected taking into account not only economic efficiency but also managerial and social aspects and the capability for innovation. In developing alternative solutions the following characteristics must be taken into account:

1. Type of production: mass, large batch, medium batch, small batch, one-off
2. Method of production organization: flow, batch, or unit

TABLE 11.3 Ratio of value of output to cost of fixed assets in relation to the size distribution of production enterprises in Soviet industry in 1974.

Size class	Value of gross output of enterprise (10^9 roubles)	Total value of gross output of enterprises in size class (10^9 roubles)	Total cost of fixed assets employed by enterprises in size class (10^9 roubles)	Ratio of value of output to cost of fixed assets
I	up to 500	2.24	5.9	0.38
II	501 – 5,000	53.2	60.4	0.88
III	5,001 – 10,000	46.0	50.1	0.92
IV	10,001 – 50,000	141.8	104.9	1.35
V	50,000 and more	203.8	102.7	1.99

Data Source: National Economy of the USSR (1973 – 1974).

3. Complexity of products
4. Type of production process: continuous or discontinuous
5. Technical level of production: progressiveness of machines and equipment, degree of labor mechanization and automation
6. Economic level of production: value of fixed assets, number of production personnel, production capacity
7. Innovation capability

We would propose that future research on scale at IIASA focus on the interconnections and relationships between the scale, organization, and innovativeness of production systems.

REFERENCES

Economics of Construction. 1976. N9, p. 76.

Glagolev, V.N. 1975. Mechanizacja i avtomatizacija truda v promyshlennosti Litovskoi SSR (Labor mechanization and automation in the Lithuanian SSR's industry). Intis: Vilnius.

Knop, H., ed. 1977. The Bratsk-Ilimsk Territorial Production Complex. Proceedings of the Second IIASA Conference on Case Studies of Large-Scale Planning Projects, 22 – 25 March 1976. CP-77-3. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Knop, H., and A. Straszak, eds. 1978. The Bratsk-Ilimsk Territorial Production Complexes. A Field Study Report. RR-78-2. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kosygin, A.N. 1976. Guidelines for the Development of the National Economy of the USSR for 1976 – 1980. Moscow: Novosti Press Agency Publishing House. p. 39.

National Economy of the USSR in 1964. 1965. Moscow: Statistika.

National Economy of the USSR in 1973. 1974. Moscow: Statistika.

Pravda. 1979. July 3, N 184 (2249).

Problems of Economics. 1977. Nt, p. 60. Moscow: Pravda Publishing House.

CHAPTER 12 THE FACTOR "MANAGEMENT" AND THE PROBLEM OF THE SIZE OF ECONOMIC ORGANIZATIONS

Vesselin Stoyanov and Evka Razvigorova
Institute for Social Management,
Sofia, Bulgaria

12.1 INTRODUCTION

One of the basic problems of the organization of social production at its current stage is the development of the economic organization as a structural unit of the national economy. This question acquired particular urgency from the moment when the individual production unit (a plant, a factory, a mine, etc.), differentiated by production and territorial principle, ceased to be the only form of economic organization. Economic organizations appeared that included in their structure and individual enterprise only as a subdivision, as one of their elements. This process confronted science and technology with a number of new problems, the most important of which are the problems of the functions, the forms, the scope, and the size of the various kinds of economic organizations.

The experience acquired in Bulgaria, and in the other socialist countries, serves as a basis for an evaluation of the various forms of economic organization and makes possible the analysis of the results of their activity under various conditions of work. Such an evaluation on the basis of the experimental and statistical data is currently useful and necessary. But this evaluation is not sufficient and it does not reveal the form, the scope, and the size of a particular economic organization under specific conditions necessary for the realization of the highest efficiency and quality of its work. We are of the opinion that, for this purpose, of particular importance is the investigation of the various kinds of factors and of the way that they influence the scope and the size of the economic organization.

The problem of the size and the scope of economic organizations has been discussed hitherto from various points of view and has been a subject of investigation, carried out by a number of specialists from various scientific fields. But it has not yet been solved. In the course of the theoretical and applied investigations, as a rule, the size of the economic organization has been determined only as a result of the effect of one or several factors, and first of all of the production – technical factor. Considerably less attention has been paid to the connection and interdependence between the various factors, and also the totality of the factors has not hitherto been clarified. The way management affects the size of the economic organization is a problem that has also not been investigated very profoundly till now.

In this study the authors attempt to differentiate and to make a general presentation of the influence of the factor "management" upon the determination and change of the size of the economic organization. This report has predominantly an interim character, and the conclusions are drawn on the basis of the experience of our country and on the basis of observations and studies carried out in the Institute for Social Management.

12.2 ECONOMIC ORGANIZATION – NATURE AND FORMS

At the present stage, by "economic organization" we understand every organizational system, formed and established on the basis of socialist property, which through the efforts of its staff exercises an economic activity in accordance with the General plan for social – economic development of the country, works on the principle of self-support, and is a separate legal entity.

In the process of development of the national economy there emerged various forms of economic organizations. At present, in this country the following organizations are considered to be economic organizations: an economic association, an economic combine, an incorporated economic enterprise, a foreign trade organization, an agrarian-industrial complex, an industrial-agrarian complex, an economic agency, and an economic enterprise. Each of these forms is applicable and may be accepted according to the specific conditions of organization and management of the national economy.

In addition to the unique characteristics of each of the above-mentioned forms, these economic organizations have the following features in common: (a) the economic organization is a basic unit accomplishing the initial distribution of income within the framework of the national economy; (b) the economic organization is characterized by:

- Production-technical unity
- Economic unity
- Organizational-management unity
- Social unity

The specific features of the various kinds of economic organizations depend on the extent that the common features manifest themselves in each of them, and also they may be regarded as a result of a systematic combination of the above-mentioned characteristics.

12.3 THE SIZE OF THE ECONOMIC ORGANIZATION

One of the basic problems of the current operation of each of the above-mentioned organizations is their size. In the theoretical and applied investigations and developments in the field of economics and management, the size of the economic organization is determined in different ways. Thus, for example, the size of the economic organization depends on the number of its basic (direct) workers or on its total staff, on the total volume of the used products, etc.

The observations and investigations prove that the size of the economic organization is one of its complex characteristic features that cannot be defined by way of only one parameter. For example, an economic organization can be "big" in terms of its production output as a result of the use of modern technology and a higher level of automation, and "small" in terms of its direct and indirect personnel, in terms of used resources, or in terms of low social effect (satisfying limited social needs, etc.). In this sense the size cannot be defined by the goals set in the process of the organization's creation, analysis, and development. To this aim we should use a system of qualitative and quantitative parameters. It was here that the first problem appeared – the problem about the principle of development of this system of parameters.

The authors are of the opinion that the economic organization may be regarded as a system of processes and the resources necessary for their accomplishment, oriented to the realization of definite aims. That is why, in order to describe an economic organization, we have to use parameters characterizing its objectives, processes, acquired results, and resources used.

The objectives of economic organizations may be divided into three main groups – production – technical objectives, economic objectives, and social objectives. Each of these groups of objectives may be subdivided into subobjectives and tasks, which may be defined by way of specific quantitative characteristics. Thus, for example, the economic objectives may be characterized by the gross income, net income, net production, etc., and the social objectives – through the amount of funds for social, living, and cultural arrangements, etc.

To characterize the resources such parameters may be used as the size of the basic stock, measured in value or in kind, of raw materials used, the number of the total staff, the ratio between direct and indirect workers, etc. The results of the activity of the economic organization are defined by parameters of the type: volume and kind of production, volume and kind of services rendered, the degree of satisfaction of the social needs, etc. The processes running in the economic organization (economic, social, production, and management) are described by their nature (for example, continuous or discrete production), by the mode of their realization (simultaneous, successive, or both), or by their variety (production activity, and/or scientific research activity, distribution-supply activity, etc.) It may be considered that such a definition of the economic organization as a system presents it exhaustively in time and space.

Within this general system of parameters there are parameters that may be used as characteristics of the size of the economic organization. The complete and exact totality of these parameters requires profound investigations and theoretical generalizations. At the present stage it may be considered that this totality is not determined in total and in a simple way. The totality must include parameters describing the effect of the environment (outside factors), parameters describing technical progress (equipment and technology) and the level and frequency of technological changes, parameters describing the market impact, environmental protection, etc. Another important problem is the determining of the qualitative parameters describing the different aspects of size subject to management process and decisions (social development of the working team, social and public effect of production, personnel qualification, etc.). This is the first important problem that should be solved in future investigations.

But in the course of the construction of this totality we should have in mind that each of these parameters characterizes the size of the economic organization only in one or several (but not all) of the aspects discussed so far. Thus, for example, the characterization of the economic organization in its production-technical aspect may be accomplished on the basis of the volume of the production in kind; in its economic aspect, through the volume of production in value, etc. Without claiming for exhaustiveness, the authors present a model totality of parameters, referring to the production-technical, economic, organizational management, and social nature of the economic organizations. These parameters and their relation to the various aspects of the economic organization are systematized in Table 12.1.

Table 12.1 reveals the size of the economic organization in various aspects as a complex totality of parameters. The parameters change under the influence of various factors, and also under the influence of the factor "management." From a methodological point of view this means that the effect of the factor "management" should be studied separately with each of the above-mentioned parameters and at the same time we should investigate their effect upon the totality of parameters.

12.4 THE FACTOR "MANAGEMENT"

Management is an innate characteristic of each economic organization. For the purpose of its description and characterization, however, a system of parameters should be used.

The application of the systems approach to the investigation and clarification of the management of the economic organization (and to the task dealing with the effect of the factor "management" on the size of the economic organization requires the inclusion in the totality of parameters of two basic groups of characteristics analyzed in their inseparable unity – management as information process and management as human activity. From the point of view of cybernetics, the management of each system (and of the economic organization as well), is an information process. At the same time, however, the economic organization of this information process is carried out through the working activity of the people and represents the totality of their purposeful human activities.

Management as an information process manifests itself through the information system, whose organization may be described by:

- The information needs (input) and the sources satisfying them
- The information results (output) and the direction of their orientation
- The information flows within the organization
- The algorithm of the information processing
- The volume of the processed information
- The variety of information and of its carriers

Management as a human activity may be characterized by:

- The means and methods used in the fulfillment of the various operations
- The man as a performer of the various operations using his knowledge, skills, and specialization in the working process

- The working operations, connected with information processing, combined in the most efficient way for the fulfillment of the various operations and functions (a system of procedures)
- The personal qualities and informal actions (behavioral aspects) of the participants in the management process.

The parameters described so far give the most general characteristic of each economic organization. But here also, as was the case with the determination of the size of the economic organization, there appears the problem about the exhaustive and precise clarification of the totality of parameters, describing the factor "management" as such.

12.5 THE CONNECTION BETWEEN THE FACTOR "MANAGEMENT" AND THE SIZE OF THE ECONOMIC ORGANIZATION

For the investigation of the connection between management and the size of the economic organization, it is suitable for us to use factorial analysis. In accordance with the above approach, the argument "management" is described by a set of parameters $\bar{Y}(y_1, y_2, y_3, \ldots, y_n)$, where y_i is the designation of a specific parameter; the function "size of the economic organization" is determined by a great number of parameters $\bar{P}(p_1, p_2, p_3, \ldots, p_s)$, p_j being a parameter characterizing the size of the economic organization.

In the application of factorial analysis, we should have in mind the following principles.

1. Practical observations and some special investigations show that the connection between the factor "management" and the size of the economic organization does not manifest itself as a functional dependence between the sets $\bar{Y}$ and $\bar{P}$ but only between y_i and p_j, that is, the dependence is between the various elements of the sets, but not between the sets themselves.
2. On the basis of (1), the connection between y_i and p_j should be revealed in two aspects: as a direct influence of y_i upon p_j or as an indirect influence of y_i upon p_j. By indirect influence we understand that a given y_i influences a given p_j through other y_i and p_j, for example, $y_2 \to y_7 \to p_3$, or $y_2 \to p_3 \to p_4$, or $y_3 \to p_5 \to p_3$. The indirect influence is due to the connection and interdependence between the elements of the set $\bar{Y}$ and the set $\bar{P}$ taken separately. The main point of this conclusion is that the connection between y_i and p_j is of a complex character and that in the course of the investigation the direct as well as the indirect connections should be accounted for.
3. In the investigation of the relations between the factor "management" and the size of the economic organization, management appeared in two aspects: once as a characteristic of the size of the economic organization and secondly as a factor determining this size. This complicates considerably the clarification of the investigated connection.
4. At present, the effect of the majority of y_i upon p_j may be determined only qualitatively. It is very difficult to find measures for the quantitative determination of the nature of this phenomenon, as a result of which the influence of y_i upon p_j usually is measured by the degree of fulfillment of some definite requirements proceeding from the effect of management upon the size of the economic organization.

5. The interdependence between the factor "management" and the size of the economic organization is of a complex matrix character. This interdependence is shown in Figure 12.1.

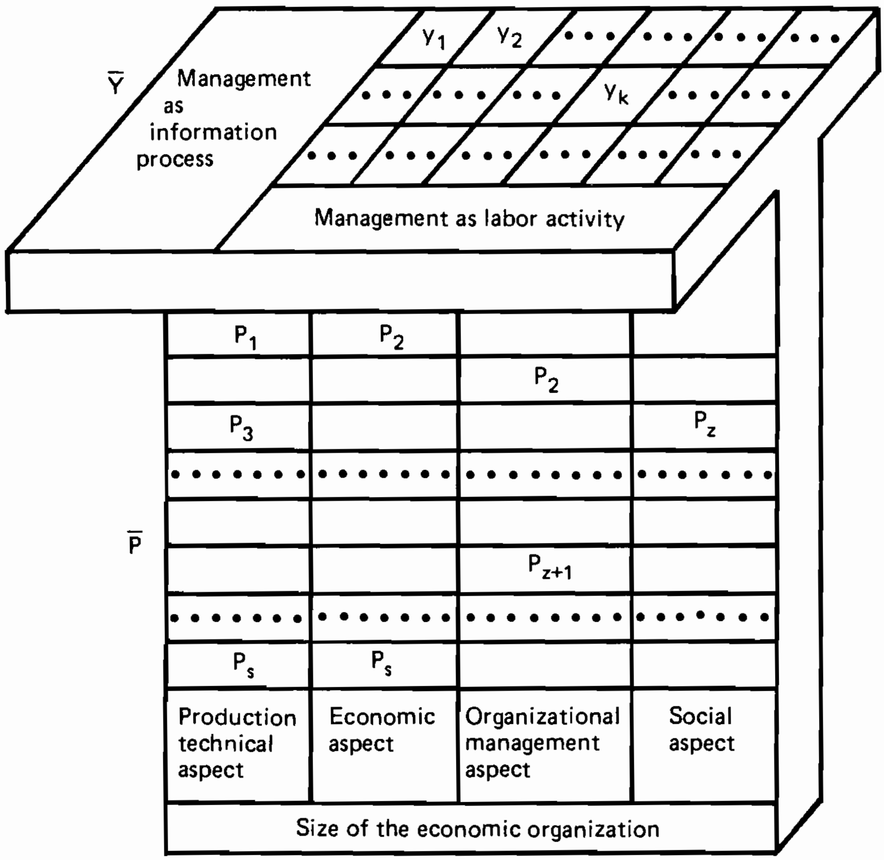


FIGURE 12.1 The complex relationship between "management" and the size of an economic organization.

These considerations lead to the conclusion that the development of a model reflecting the dependence between the factor "management" and the size of the economic organization is a complex task because (a) the clarification of the argument "management" and of the function "size of the economic organization" needs additional profound investigations, and (b) the development of a model, reflecting the complex interconnection between management and size of the economic organization, needs profound developments and investigations in methods of scientific measurement.

Because, on the one hand, the parameters of the factor "management" determine the size, and on the other, the parameters of the production – technical, economic, and social characteristics are determined and changed on the basis of management decisions, it may be accepted that the size of the economic organization is first of all a management problem. This points to the necessity to create instruments for analysis, description, observation, and construction of the management process in a way that ensures timely and efficient control of the process of decision making concerning the parameters determining the size of the economic organization (see the scheme in Figure 12.2).

At the present stage of development of economic organizations (not only in Bulgaria but all over the world), in our view, what is basic for the determination of the organization's size are the decisions about the level of concentration and specialization of production and the level of centralization and decentralization of functions and activities and responsibilities. The size of each socioeconomic system in terms of the production-technical factor is determined by decisions concerning the specialization and concentration of production. The latter are subject to management decisions on different system levels and different management activities. For example, centralization of research and development leads to the creation of large research units with all the ensuing changes; decentralization compels each economic organization or factory to develop this type of activity with due consequence to the size. The size is also influenced by the decisions on centralization and decentralization of management functions (for example, decentralization of the supply function). Therefore, a conclusion can be reached that the concentration/specialization and centralization/decentralization ratios give us a possibility to regulate and determine the size of every economic organization within any big economic system. The building of a simulation model of the management system (like the one presented on Figure 12.2) will make possible efficient control, and fast centralization and decentralization of activities on different levels, creation of a supervision and control system, and correction of the parameters describing all aspects of the size. It is necessary to note that this simulation model can solve the problem about "size" only as one of the problems of management of large systems. Above all, it is oriented to the problems of structuring and functioning (technology and mechanisms) of the system and it develops the set of tools of organizational design.

The above arguments are sufficient reason for applying systems analysis and modeling methods in the solution of the problem of the effect of the factor "management" upon the size of the economic organization.

12.6 CONCLUSIONS

1. The size of the economic organization is one of its complex characteristic features, and is revealed in four aspects by a system of parameters.
2. Management is a phenomenon that is revealed at least in two aspects and is described by a system of parameters.
3. The influence of the factor "management" on the size of the economic organization cannot be investigated through the influence of the totality of indicators $\bar{Y}$ upon the totality of indicators $\bar{P}$, but requires investigation of the effect every y_i exercises upon every p_j; the effects may be direct or indirect.
4. We should develop the necessary instruments for analysis, description, and observation of the change of parameters, clarifying the factor "management" and its influence on the size of the economic organization.

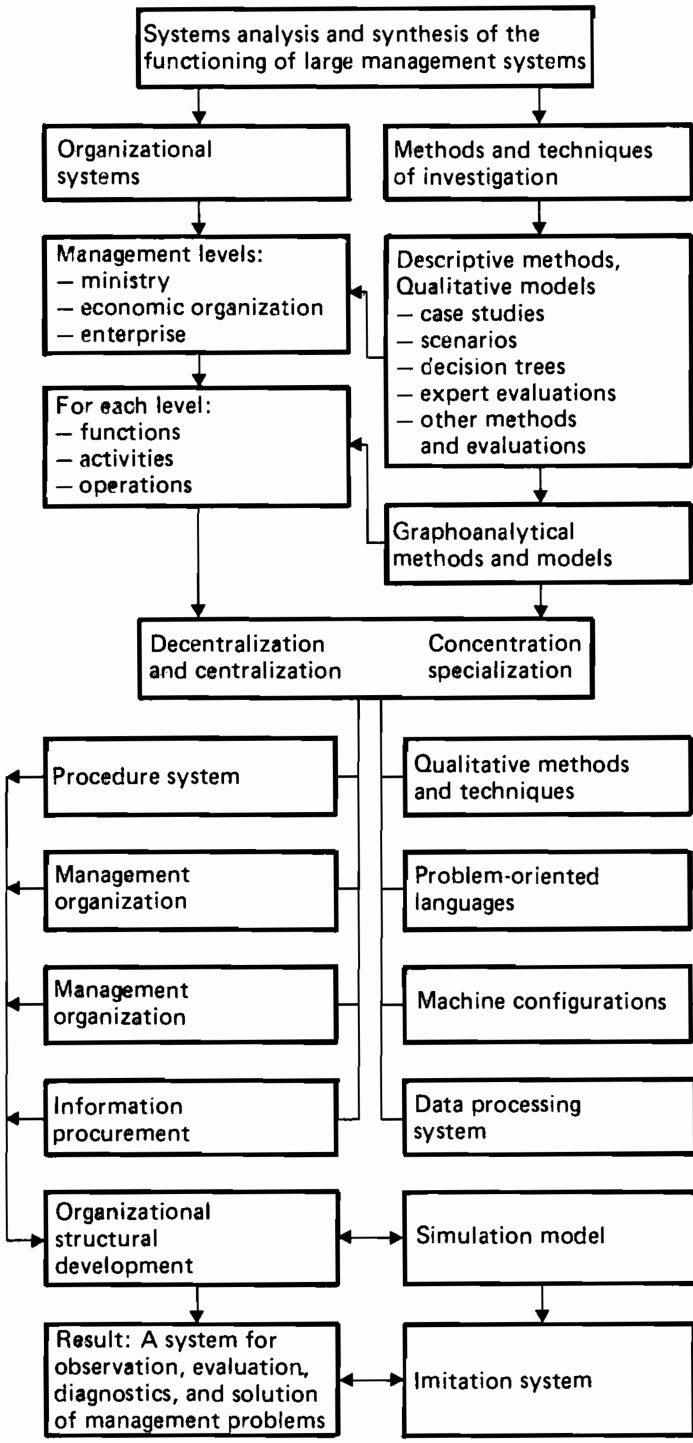


FIGURE 12.2 Systems approach for understanding large management systems

TABLE 12.1 Indicators characterizing the size of the economic organization.

	Production technical aspect	Economic aspect	Organizational management aspect	Social aspect
Volume of production in kind	x			
Volume of production in value		x		
Basic production funds in kind	x			
Basic production funds in value		x		
Number of workers in the economic organization	x		x	x
Working capital in kind	x			
Working capital in value		x		
Volume of the processed raw materials	x			
Volume of production (including cooperative supplies and commodity funds)	x			
Territorial proportions of the economic organization	x			x
Variety of production (specialization of the economic organization)	x		x	
Amount of the funds for social development of the working staff in kind				x
Amount of the funds for social development of the working staff in value		x		
Degree of economic independence		x		
Organizational structure of the economic organization (including number) of managed units and number of management levels)			x	
Type and scope of management functions			x	
Management staff			x	
Management subsystem costs in value		x	x	

CHAPTER 13 THE MANAGEMENT OF MANAGEMENT AND THE SIZE OF MANAGEMENT

*H.I. Ansoff**
European Institute for Advanced Studies in Management, Brussels, Belgium

The typical enterprise takes in inputs from the environment and has as outputs goods and services. My view is that the enterprise consists of two processes: the logistic process which converts what goes in, such as materials, ideas, information, or money, into identifiable goods and services delivered to the environment; and the management process which provides control and guidance.

In thinking about the size of management, there are a number of factors that one must consider:

- The qualifications of managers
- Their cultural attitudes, e.g., value systems, rewards, and their view of the world or model of reality
- The way power is distributed between managers and other constituents
- The formal process for doing organizational work
- The kind of logistic technology
- The technology of management, such as formal models and methods for strategy formulation

It is obvious that as the scale of operations of the enterprise increases so does the size and complexity of the management function. Thus it is necessary to ask ourselves "If we are concerned with enterprise efficiency, how does what we gain by increasing the size of factories compare with what we may gain or lose as management becomes more complex?"

To answer this question I would like to suggest the following two propositions:

Proposition 1: In a given industry there is a size of enterprise beyond which its effectiveness declines owing to over-complexity of management (the dinosaur effect).

Proposition 2: In a given industry there are several local optimal sizes (Figure 13.1).

* This summary of Professor Ansoff's presentation at the workshop was written by J.A. Buzacott.

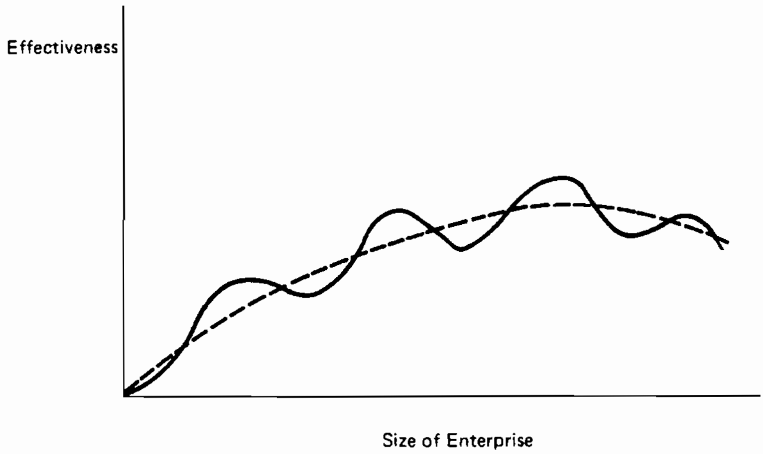


FIGURE 13.1 Relationship between effectiveness and size of enterprise.

The explanation of proposition 2 is that management is not infinitely subdividable – it is lumpy. For example, up to a certain size, an enterprise can depend totally on oral communication. Once a management communication system is introduced, the complexity and size of management immediately increases and it is not until the size of the enterprise is further increased that full advantage of the efficiency gained can be realized.

Thus far it has been assumed that it is the scale of productive activities that is the prime determinant of the scale of management. While that may have been true in the first half of this century, it has since come to be recognized that there is another important function of management – strategic management: the enterprise-defining and direction-giving function.

To understand the nature of this problem it is necessary to ask what are the characteristics of the responsiveness that is expected from the organization.

1. Operating responsiveness, measured by efficiency – the ability of the firm to produce at lowest cost
2. Market responsiveness, measured by near-term effectivensss – the ability of the firm to get an adequate return on sales or investment and achieve growth.
3. Strategic responsiveness, measured by long-term effectiveness or flexibility – the ability to determine what business the firm wants to be in
4. Social responsiveness, measured by the satisfaction of workers and the minimization of negative externalities such as pollution
5. Political responsiveness, indicated by the enterprise being viewed as having a viable legitimacy

As the respective dimensions of responsiveness become important to the success of the firm, both the size and qualifications of management become an important factor not only in success but also in cost.

Increasing attention must be paid to management so it is necessary to identify the major factor that determines the relative importance of the management of the logistic process and the management of all the other aspects. As suggested in Figure 13.2, the relative importance of management of items 2 through 5 above varies with the level of environmental turbulence. At high turbulence level the bulk of management attention shiftes away from the production plan.

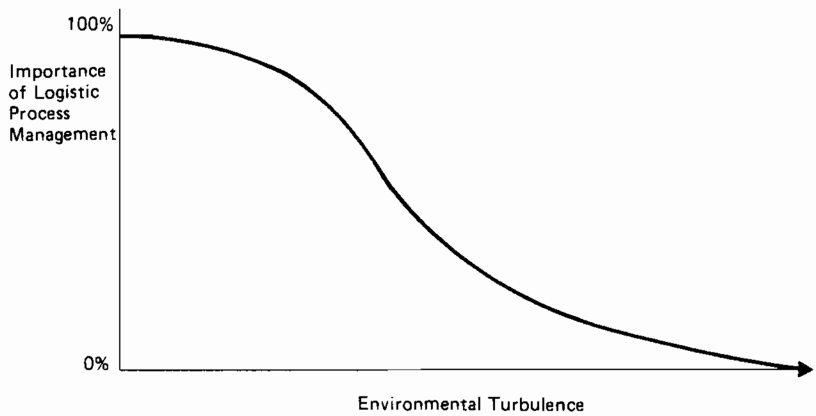


FIGURE 13.2 Relative importance of logistic process management as a function of environmental turbulance.

In the United States, in particular in the early years of this century, environmental turbulence was low. As a result, the "minimum management principle" developed, and managers were judged by the smallness of their indirect to direct employee ratio. This was appropriate because if management was not restricted, all kinds of bureaucreatic tendencies set in in large organizations – Parkinson's and Peter's principle. But once the environmental turbulence increased the minimum management principle was no longer appropriate and the question of the optimal management size appeared.

Instead of the assumption of the minimum management principle (Figure 13.3a) it is necessary to develop an optimal management principle (Figure 13.3b). In different environments more responsiveness is required. One can summarize the underlying assumptions by the following propositions:

Proposition 3: In an environment with a given level of turbulence there is a minimal size of management necessary for continued viability of the enterprise (the critical mass)

Proposition 4: Above the critical mass level there is an optimal size beyond which the potential effectiveness declines (Figure 13.4).

However, it is important to note that for a given level of turbulence the optimal size dedepends on the sophistication of management and its use of appropriate management technology.

That is, the size of management is the sum of the size of the "lower" management who are directly responsible for the logistic system and the size of the "upper" management who are concerned with the relation between the enterprise and the environment. Thus, the optimal size of management will be determined primarily by the

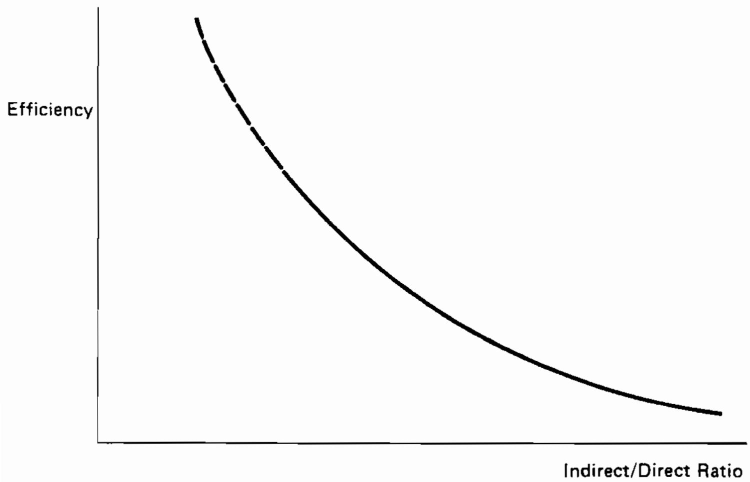


FIGURE 13.3a Minimum management principle.

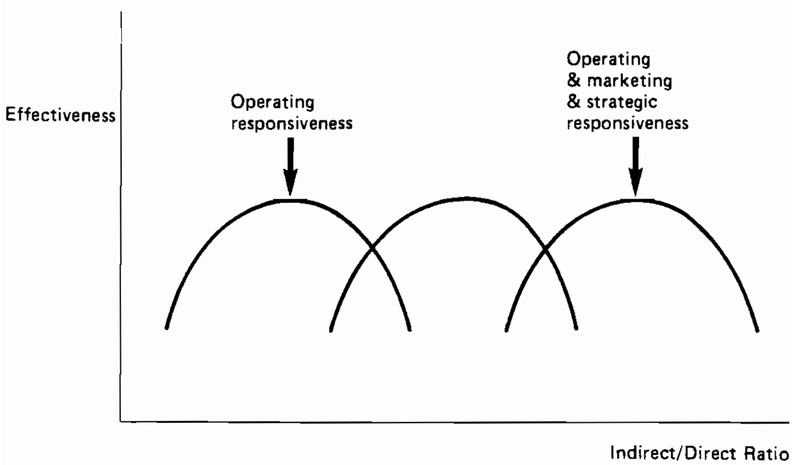


FIGURE 13.3b Optimum management principle.

factors affecting the size of lower management: the size of the logistic system and its technology, and by the factors affecting the size of upper management: the environmental turbulence, the diversification of the enterprise, and the management technology.

To summarize:

1. Economic productivity is only one dimension of organizational effectiveness.
2. The relevant dimensions of effectiveness that need to be included in the study of the design of the enterprise are determined by the turbulence and the complexity of the demands in the enterprise.

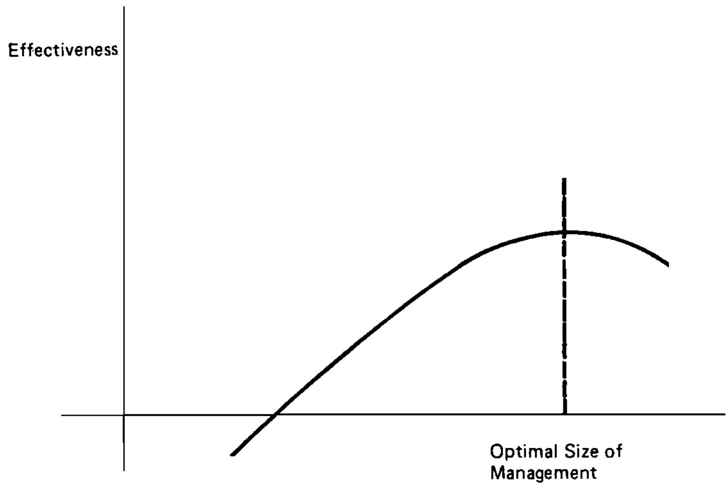


FIGURE 13.4 Effectiveness of enterprise as a function of size of management. The parameters are level of turbulence and size of the logistic system.

3. Scale, as well as other characteristics of the enterprise, are determined not by logistic work alone, not by managerial work alone, but by influences and imperatives of the two. It is the increase in managerial imperatives that is raising problems at present.

4. The demands of logistic responsiveness (may) conflict with managerial responsiveness. To think about and react successfully to the environment may require a different managerial structure from that required to think about and react to the problems of the logistic process. There is beginning to emerge a duality of structure. An organization may go into one form for the purpose of dealing with the reality of the environment, sort it out, and then flip into another form for the purpose of making money out of that reality.

5. In highly turbulent environments management imperatives (may) dominate logistic imperatives.

CHAPTER 14 ORGANIZATIONAL SCALE: SIZE, STRUCTURE, AND ENVIRONMENT

J.A. Buzacott and M.F. Cantley
International Institute for Applied Systems Analysis,
Laxenburg, Austria

The papers and discussion on the scale of organizations were concerned with three major topics

1. The relationship between the size, cost, or difficulty of management and the scale of the productive activities
2. The appropriate structure of the organization and the degree of centralization or decentralization of management activities
3. The effect of the environment on the size and structure of management

Ansoff's paper was concerned with all three topics, although with less of an emphasis on the second. However, a number of other papers focused on just one of these topics.

14.1 DISCUSSION SUMMARY

14.1.1 Management and the Scale of the Productive Activities

Empirical evidence of the way in which the difficulty of management increases with the scale of productive activities is given in the paper by Bendkowski, Stachowicz, and Straszak (Chapter 8). They found that in large coal mines certain classes of decision problems occurred more frequently and were more difficult, particularly personnel problems.

Van Hees and den Hertog (Chapter 9) in their Figure 9.2 also postulated the same relationship: organization cost increases with the scale of activities. They stressed the management problems of coordinating efficient small units, overcoming the difficult problems of big units and coping with the dynamics of the change of scale of activities. They also proposed (Table 9.1) a taxonomy of the situations where small-scale units are appropriate.

In the discussion of the van Hees and den Hertog paper some further aspects of organizational scale were brought up by LaPorte in a question.

La Porte: . . . in large technological organizations in both the public and private sector in the U.S., at some point the complexity of the decision elements that upper-level management faces forces them to radically simplify their situation in order to bear it. There is a high degree of stress attached particularly to changing conditions that are surprising and haven't been thought about both from inside and outside. Have you had enough experience with different product lines to be able to say something about the social properties of the technology which change the internal complexity of management requirements? Could this element be used to relate the scale of persons or scale of activities to different degrees of complexity in the situation?

Van Hees agreed that the complexity of management was very dependent on the production situation, the technology, the market, and the amount of coordination required. Apter followed up this question by asking

Apter: In your presentation you mentioned that in the 1960s you were concerned with the human organization and what to do about the behavioral environment of work. Does this wind up with the kind of problem La Porte was mentioning – additional coordination needs, greater bureaucratization costs or managerial costs – so that you really are further behind after these human behavioral factors have been taken into consideration?

den Hertog: We have about 20 years' experience in the field of job design. In the first 10 – 15 years what we did was to try and rearrange all kinds of things and bring in the human side to the shop floor of the organization. But after a while we discovered that the output in the social sense, in job satisfaction, and in the economic sense was very limited by the structures at the next highest level. So at this moment we are trying to connect views from a higher level with those that work on the shop floor level. We are trying the idea of factory development and factory design. We have moved over from job design practice to organizational design and so we have to develop a far more integrated view of the function and effectiveness of the organization. When you just try to implement isolated changes in some corner of the factory you can get the impression that you have discovered the door of industrial paradise. However, the only way to do it effectively is to redesign the whole organization so that you can take into account at the managerial level the same arguments and the same needs that you are considering at the shop floor level. Managers are, in their own way, also alienated.

Razvigorova and Stoyanov (Chapter 12) were concerned with developing a model by means of which the appropriate size of management could be determined. In the discussion, a number of other factors were suggested:

Plug: In linking the size parameters of an organization with the management ones you have many parameters which are purely quantitative and which would not contribute very much to managerial load if some of them were large and some of them were small. Some of these factors are things which would be dealt with by accountants and would not really be managerial load.

Wouldn't it be good to go somewhat deeper into qualitative aspects, or to select a few that have real elements of variety, of unforecastability. Also wouldn't it be good to include a few external factors because many of the parameters you mention are parameters describing the organization itself, internally, whereas things like market phenomena or developments in technology which are largely confined to the environment of the enterprise do not enter into this list of parameters.

Razvigorova agreed that more factors describing the environment were required. Also it would be desirable to include the impact of technology on management.

14.1.2 The Appropriate Structure of the Organization

In the discussion on Razvigorova and Stoyanov's paper this question was raised:

Buzacott: I am starting to get the feeling that as far as management is concerned size is not really the issue. Maybe the amount of management you need is pretty well determined by the logistic system and, to some extent, the environment. What really matters is the structure of management, that is, who is going to be assigned responsibility to react and who is going to be assigned power to be able to make other people do things. I am wondering whether, in your model, you just look at simple centralization and decentralization which seems to imply just two extremes. This may be too simplistic an approach to this problem.

Razvigorova: The problem of centralization and decentralization and concentration and specialization is just a small box in our model but in fact we think it is the core of the problem. To some extent we can solve the problem of scale by centralizing and decentralizing management activities and to some extent we can solve the problem of scale by concentrating and specializing some working activities in our economic organizations. So everything is directed to solving the optimum or efficient ratio between the level of centralization or decentralization and the level of concentration or spezialization.

Ansoff: The Western experience is that the original notion of centralization or decentralization is being replaced by a somewhat more sophisticated set of notions – at what level does a particular activity optimally take place. It is not a question of whether we put everything down or push everything up but how to distribute the various functions or responsibilities. I presume that is the sense in which you are talking about it.

Razvigorova: Yes.

Other papers that considered the question of organization structure were those of Egiazarian and Glagolev (Chapter 11) and Haustein and Wittich (Chapter 18). Both papers are concerned with the development of improved industrial structures in socialist countries. The objectivies of the industrial reorganization have been to decentralize much decision making to a level where the management is responsible for a number of plants making a specific range of products. This enables the advantages of concentration and specialization to be exploited. It gives management the responsibility for ensuring on the one hand efficiency in the production of existing products and on the other hand the development of new products and processes.

A presentation at the workshop that considered similar issues from a Western perspective was that by Een (1980). He described the results of a study of the likely structure of the food industry in the year 2000. He projected a structure in which there will be

> a small number of national or multinational corporations with centralized resources, for among other things, R & D. In many countries the food industry is in a defensive situation regarding the food and health authorities. The industry has to spend the larger part of its R & D budget compying with regulations and proving that its products are nutritious and harmless. Small organizations cannot afford this type of work.
>
> On the other hand the production units will become more decentralized than at the present time. These satellite production units will be standardized, however, thus achieving a certain economy of scale in design, building and maintenance. It is also possible that the production units will be more flexible in the sense that certain unit operations can be reprogrammed for changes in raw material and end-products, thus avoiding the worst effect of short season crops.

Among the reasons for the small production units are high transport costs, decentralized sources of energy, diminishing world trade, a higher degree of national self-reliance, a desire for complete utilization of agricultural primary products, technological developments permitting more flexible automated processes and increasing demands for reduced vulnerability of the production system to accidental disturbances or intentional disturbances such as strikes.

In discussing this paper, Rochlin raised some interesting issues:

Rochlin: Two points that I would like to make. (1) Threshold vs. saturation . . . if things are small enough you stay below some threshold of trouble, either environmental or social-political. However, if the distributed effects everywhere are such that they exceed the threshold where problems occur there is an interesting argument that you should centralize massively and concentrate all the damage in one place. Once you get past some saturation point it doesn't matter how much more than that you pile on. Once you get above a certain threshold of labor difficulties you can make the task progressively worse and worse and not have much more in the way of outages. This also applies for environmental effects. I think this needs to be discussed more in order to see which approach is applicable in the future. (2) The problem of networking. When you have a large distributed system it is not right to say that system is more flexible. It is entirely possible that, in order to get the decentralized system coordinated, the political and managerial complexity which is necessary to keep this operation running in a large-scale society is a higher cost than nominal centralization. Again, the question of the costs of networking is something which needs to be discussed more.

14.1.3 The Effect of the Environment

Ansoff's propositions in the effect of turbulence in the environment on the scale of management attracted considerable discussion, particularly on the operational definition of turbulence.

Cantley: You mentioned that you could put scales on degree of turbulence. I wonder whether you would be prepared to offer some sort of insights on that. At what point do you run into the need for qualitative changes because of a quantitative increase in turbulence? How do you know? How do you measure it?

Ansoff: The concept that I find useful both theoretically and practically is a relative concept and it is made of two issues.

One is that one has to learn not to confuse changeability with turbulence. The concept of (discrete) levels of turbulence has to be elaborated and on each level there may be a great deal of changeability . . ..

The second issue may not be right but certainly for managerial purposes it determines what an enterprise must or must not think. One of the most important characteristics of an existing level of turbulence is the extent to which it calls from the enterprise behaviour that is discontinuous from prior learning. Another way of saying this, which arises out of Chandler's historical work (Chandler 1962), is that for every level of turbulence in the environment . . . there is a type or range of behaviour by organizations which are successful in that environment. It is a very narrow range and outside that range you will not succeed. There is an internal pattern of competence, which is sufficiently documented by Chandler, that is needed to support this behaviour. If the environment had been at a particular level of turbulence it evoked a particular type of behavior from the organization and the organization at one time adapted its capability. If the environment moves to another level of turbulence then the distance (between levels) is the measure of the difficulty. So discontinuity from prior experience and from prior learning is one way in which this can be approached. It works in practice and it is reasonably well elaborated theoretically.

Tomlinson: The point I raise may seem to be a quibble of words but I am getting worried about the way turbulence is being used and I don't think it is due to the fact that I once did research in aerodynamics. But it seems to me that what you are talking about is a mixture of complexity and instability. You are essentially dealing with things that change state and the problem of management in this environment is also its complexity that you can't structure and therefore is one of the things that gives rise to size.

La Porte: It strikes me that the major unexplored and unsystematized aspect of thinking with regard to scale and organizational behaviour is an attempt to clarify and empirically examine the character of environments, so that we can specify that they are turbulent, the degree to which their turbulence exists and so forth. In the organizational development literature over the last 10 years there has been a lot of language on the importance of the environment, the importance of turbulence, without almost any conceptual or theoretical refinements of those terms. There is a kind of intuitively nice feel to the notion of turbulence but in any systematic sense we cannot say very much more than things are becoming more turbulent or changeful . . .

One area of research with regard to scale should be to seek a better understanding of the changes in technology and in organizational structure that occur when an organization increases in scale and in size. An important aspect of this research, in addition, would be to more clearly specify the relationship of an organization undergoing the process of increasing scale with the organizations in its immediate environment. These studies would give us a basis for understanding the degree to which changes in both internal and external relationships effect the sense executives have of "turbulence in their surroundings." Turbulence is not predominantly an objective condition; it is a perceived condition, relative to an executive's perception of conditions both within and outside the organization and his or her ability to understand them. Turbulence doesn't mean change alone – change is everywhere present. Rather it means surprising change – things we didn't predict and cannot take into account once we have perceived them. Reduction of that sense has to do with improving our causal understanding of the social relations between large enterprise and other large-scale organizations in that environment. We have come a very short distance in providing conceptual language, and certainly empirical follow-up is necessary to give a more systematic meaning to the notion of turbulence.

Ansoff: . . . In terms of empirical work we are not all that naked. You are suggesting we are at ground zero but we are very far from it. We are off the ground and in terms of empirical work there is a fair amount of work being done . . .. I think really that in a sense the most seminal important properly empirical work is that done by Chandler. It has shed an enormous amount of light in this area so while I agree with your imperatives I really don't agree with the state of knowledge as you describe it.

14.2 COMMENT

There is a very large literature on the size and structure of organizations and one of the problems of a workshop with participation drawn from different countries and a wide variety of disciplinary backgrounds is that the participants do not have common terminology, assumptions, or familiarity with the same literature. Apart from the references cited in Chapter 1 (Dewar and Hage 1978), two review papers that evaluate and discuss the literature on organizational size and complexity are Kiberly (1976) and Caplow (1957). Thus the major value of the workshop papers and discussion on the relationship between management size and structure and the scale of productive activities is probably the insight they provide into the perspectives of scientists from East and West.

However, Ansoff's paper and the discussion on the meaning of turbulence and its effect on management size and structure indicate that there is still considerable disagreement on this aspect, so further comment on the different types of environments and the behavior of the combined system of the organization under study and its environment seems appropriate. What is required is the development of conceptual models of the nature of environments and the way in which changes in environments affect the change in size and structure of organizations.

There are fundamental differences between the environments in planned economies and market economies and it is essential that consideration be given to both (a) the environment perceived by the management decision maker, and (b) the general goals and objectives of economic development.

In a socialist economy the basic criterion of effectiveness is the achievement of the best results in the interests of society. Private ownership in a market economy is not directly committed to the maximum satisfaction of social needs and there can be a basic conflict between effectiveness from the point of view of the total society and the objectives of corporate organizations.

In a socialist economy there is an interlocking system of plants at different levels in the economy (Berry 1977). This means that to the decision-maker the environment seems relatively clearly described but on the other hand the system of plans must be carefully coordinated in order to ensure that the decision-maker's perceived environment leads him to act in a way that is in the best interests of society.

In market economies the environment of the organization is less certain, first because of uncertainty about changes in the total market, secondly because of uncertainty about the behavior of competitors.

Emery and Trist (1965) developed a now well-known categorization of different ideal types of environment and the characteristics of organizations appropriate to each type:

Type 1: *Placid randomized* environment – benefits and dangers randomly distributed. "Organizations can exist adaptively as single, quite small units."

Type 2: *Placid clustered* environment – a detectable structure in the distribution of benefits and dangers. "Organizations grow in size."

Type 3: ***Disturbed, reactive*** environment – the placid clustered environment is complicated by other organizations sharing the same environment. Size becomes perceived as an instrument of competitive strength, as in military contexts (e.g., Lanchester's laws on a tactical level), or in terms of market share, as in a strategy analysis based on "dominant market share."

Type 4: *Turbulent field* – the behavior developed to cope with the conditions of the type 3 environment leads to changes in what were previously its constant factors (e.g., climate, fundamentals of social custom, and behavior). The interactions between organizations contribute to effects beyond their control or expectations. This often leads to attempts by the participants to establish and reinforce commonly held values as a constraint to their behavior, such as the development of cartels in oligopolistic industries, or the establishment of new organizational entites with specifically mediating or coordinating roles.

In a paper that builds on Emery and Trist's concepts towards a more rigorous analytical examination of modes of regulation in turbulent fields, Metcalfe (1974) expresses the need as follows:

> It calls for the creation of a body of empirically based macro organization theory to explain the dynamics of complex pluralistic systems. Macro organization theory in this context does not just mean theory applicable to large scale as opposed to small scale phenomena. The term is used to denote an important new theoretical distinction between adaptive behavior at the system level as distinct from the organizational level. The same principles need not apply at both levels. What is individually rational may be collectively irrational and vice versa.

Metcalfe argues that

> coordination of pluralistic systems depends on voluntary interorganizational negotiation. But when the causal texture of the environment becomes very richly joined, this method of achieving cooperation becomes both more difficult and more important to sustain.

The same phenomena are illustrated by Simmonds (1975) with specific reference to the petrochemical industry as an example of an industry in transition from a disturbed reactive phase:

> The industry is thus no longer the self-generating, self-propagating, on-going petrochemical system of the 1950s and 1960s. It is evolving into a key part of a number of larger systems. The factors which increasingly control its actions now reflect the growing impact and interaction of society with the industry, e.g.,
>
> – Significant political intervention/control over raw materials
> – Growing competition between natural and synthetic technologies
> – Further ecological and energy constraints
> – Increasing responsibility for the effects of products and services
> – More competition from more major competitors in a wider range of industries within one market area.

This type of general analysis and discussion tends to appear over-philosophical and speculative, as soon as it leaves concrete operational realities. But in our opinion it may provide a route to greater general understanding, and to the development of more appropriate methodologies, in many of the contexts in which problems of large-scale organization appear. Some of these contexts have been illustrated in the chapters of this book.

REFERENCES

Berry, Ya. ed. 1977. Planning a Socialist Economy, Vols. 1 and 2. Moscow: Progress Publishers.

Caplow, T. 1957. Organizational Size. Administrative Science Quarterly 1 : 484 – 505.

Chandler, A.D. 1962. Strategy and Structure. Chapters in the history of the American industrial enterprise. Cambridge Mass.: MIT Press.

Dewar, R., and J. Hage. 1978. Size, Technology Complexity and Structural Differentiation: Toward a Theoretical Synthesis. Administrative Science Quarterly 23 : 1

Een, F. 1980. Food Industry in the Year 2000. CP-80-XX. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Emery, F.E., and E.L. Trist. 1965. The Causal Texture of Organizational Environments. Human Relations 18, reproduced in Emery, F.E., ed. 1969. Systems Thinking. London: Penguin.

Kimberly, J.R. 1976. Organizational Size and the Structuralist Perspective: A Review, Critique and Proposal. Administrative Science Quarterly 21: 571 – 597.

Metcalfe, J.L. 1974. Systems Models, Economic Models and the Causal Texture of Organizational Environments: An Approach to Macro-organization theory. Human Relations 27: 639 – 663.

Simmonds, W.H.C. 1975. Industrial Behavior Patterns: A New Dimension for Planners. Futures. 7: 284 – 292.

CHAPTER 15 INNOVATION AND ORGANIZATION SCALE

J.A. Buzacott
International Institute for Applied Systems Analysis,
Laxenburg, Austria

The purpose of this chapter is to highlight the issues concerning innovation that were mentioned in papers or discussion at the workshop and that specifically concerned the size or structure of the organization in which innovation was occurring.

15.1 DISCUSSION SUMMARY

Innovation and organization size arose in the discussion on the paper by van Hees and den Hertog. Innovation was the main topic of the presentation by Utterback and it also arose in the discussion on Gold's presentation. The significant issues that arose are described, followed by some general comments that attempt to summarize the major insights of the workshop into the relationship between innovation and the size or structure of organizations.

15.1.1 Van Hees and den Hertog

Van Hees and den Hertog suggested in their paper (Table 9.1) that if the mass of innovation potential for the process was low then a small-scale organization unit was appropriate. This point was taken up by a number of speakers in the discussions.

Ansoff: You suggested that there has to be a minimum critical size to be able to innovate, yet the general societal argument today is that it is the small firm that innovates, that small is more beautiful than large for innovation and you want to increase the size of the unit in order to make it innovative.

van Hees: I have not mentioned any absolute size. You must at least have a certain size of organization to be able to participate in innovation. For example, one important thing today is computer-aided manufacturing. We have seen, by comparing plants, which of the plants were able to fix the signals from their environment and to go, on their own, into using computer-aided manufacturing. There is more or less a correlation between the size of the plant and the willingness or ability to absorb this kind of thing. You have to have a certain level of management, not only one capable manager but a corps of managers who are able to converse about these new developments and pick it up and develop it.

Tomlinson: Is there not an indication that this may not be a linear effect - that beyond a certain size it may drop again?

van Hees: First it climbs more or less but above a certain size you arrive at a bureaucracy where you don't have this fear any more to be innovative.

A somewhat different aspect was raised by Price.

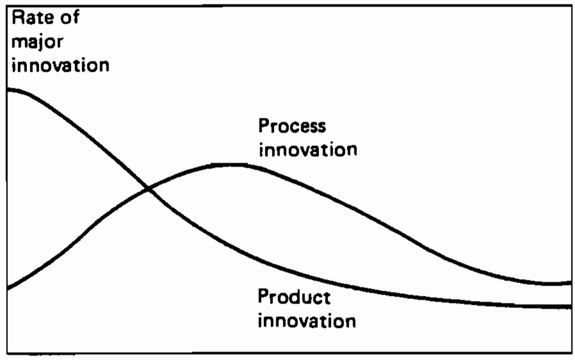
Price: In the way ideas have been developed so far one has been looking at the conditions within one country, or one culture. Now I think it is true that work forces in some counties are more disciplined than in others. For instance, one imagines that the Japanese, the South Korean, and the Asian work forces generally are more disciplined than those in other industrialized countries. This implies that the threshold in terms of optimum plant size must vary thoughout the world. This has certain implications for the industrial life-span of a particular product. When the product's production methods are small scale, when it is new, when the type of R & D that is needed is centered on improving the product (rather than the process type of innovation which gets you on to a larger scale with low cost and as a standard product), in this first stage it is probable that countries such as those in the developed Western World have a certain relative competitive advantage, while in the later stages of life-span of a product (i.e., where the product is mature) its production may go to the newly industrialized countries, say in the Far East. There is this dimension to the subject of your paper which at present isn't brought out. It tends to fit in with Haufbower's international trade theory and it is interesting for that reason.

van Hees: First it climbs more or less but above a certian size you arrive at a bureaucracy a very complex one because not only do you have to decide on the size of an operation but also the location. You can arrive at the situation where you have to pick it up and take it to some other place in the world which is very hard to do. It is easy to move someone from a little place in Germany to Paris but the other way round is much more difficult.

Cantley: Isn't one of the interesting questions in this that of identifying the extent to which you can control your own future and the extent to which it is in fact going to be controlled for you by factors in your own environment? Are you describing a historically inevitable process in which, to put it crudely, "don't pee into the wind" or are you in fact in control and if so to what extent?

van Hees: From my presentation you can derive that we are being influenced more by factors we don't have under control any more. You can't simply pick up this factory here and make it bigger or smaller or move it. We believe that we are in a much more difficult position to develop the whole structure than we were 10 or 15 years ago.

The changing character of innovation, and its changing role in corporate advance. Seeking to understand the variables that determine successful strategies for innovation, the authors focus on three stages in the evolution of a successful enterprise: its period of flexibility, in which the enterprise seeks to capitalize on its advantages where they offer greatest advantages; its intermediate years, in which major products are used more widely; and its full maturity, when prosperity is assured by leadership in several principal products and technologies.



	Fluid pattern	Transitional pattern	Specific pattern
Competitive emphasis on	Functional product performance	Product variation	Cost reduction
Innovation stimulated by	Information on users' needs and users' technical inputs	Opportunities created by expanding internal technical capability	Pressure to reduce cost and improve quality
Predominant type of innovation	Frequent major changes in products	Major process changes required by rising volume	Incremental for product and process, with cumulative improvement in productivity and quality
Product line	Diverse, often including custom designs	Includes at least one product design stable enough to have significant production volume	Mostly undifferentiated standard products
Production processes	Flexible and inefficient; major changes easily accommodated	Becoming more rigid, with changes occurring in major steps	Efficient, capital-intensive, and rigid; cost of change is high
Equipment	General-purpose, requiring highly skilled labor	Some subprocesses automated, creating "islands of automation"	Special-purpose, mostly automatic with labor tasks mainly monitoring and control
Materials	Inputs are limited to generally-available materials	Specialized materials may be demanded from some suppliers	Specialized materials will be demanded, if not available, vertical integration will be extensive
Plant	Small-scale, located near user or source of technology	General-purpose with specialized sections	Large-scale, highly specific to particular products
Organizational control is	Informal and entrepreneurial	Through liaison relationships, project and task groups	Through emphasis on structure, goals, and rules

FIGURE 15.1 The pattern of innovation. (From Abernathy and Utterback, Technology Review, June/July 1978, © 1978 Alumni Association of the Massachusetts Institute of Technology.)

15.1.2 Utterback

The major presentation on innovation was by Utterback. He reviewed his work on patterns of industrial innovation (Abernathy and Utterback 1978, Utterback 1978), putting particular emphasis on the way the pattern changes over the life cycle of the product (Figure 15.1). Early in the life cycle, major innovations tend to come from small entrepreneurial organizations outside the "formal" production system, while the relatively minor process innovations are performed in large organizations with formal management structures and clearly set goals. That is, a given organization size and structure is characteristic of the type of innovation and as the product matures there should be an accompanying change in the organization so that the innovation potential can be exploited.

Various aspects of these ideas were explored in the subsequent discussion. First of all, the requirements for successful entry:

Utterback: If you are trying to get into a business at the lefthand end of the spectrum, high-performance new products and capital goods, you tend to get in as an entrant, as an invader, stressing high performance and not cost. As the business develops, innovation shifts to process innovation and integration within the growing firms in the industry and a very high rate of productivity improvement occurs. What then happens later is a period of more incremental innovation, and focus on cost and quality. So there are different ways of getting into a business. One way is to be very innovative with products; another is to come in and imitate the product innovation, build up the technical capacity and be very good at process innovation; and another is to be very good at finance and come in late in the business at a very good location, with very good access to material supplies, and the most large-scale kind of plant. What I a saying is that I believe that, as competition grows from Third World countries like Korea in areas like fibers and inexpensive automobiles, it is going to be more and more difficult to follow that kind of a strategy and countries like ours are going to be more and more involved in the kinds of innovations which tend to occur in more highly demanding and more affluent markets and production settings, as opposed to these kinds of innovations.

The question of the relationship between organization size and innovativeness was raised by Glagolev.

Glagolev: In your opinion, is it possible to combine efficiency and size with innovativeness and flexibility?

Utterback: Many corporations are successful in combining both efficiency and innovatieness by organizing differently for the two functions. Many companies in the U.S. have 10 to 50 or even 100 strategic business units and they tend to organize the very efficient things in one way and the more innovative things in another way. Now, at least in our social context, there are a number of reasons why that tends to fail. One of the reasons is that corporations often apply rules uniformly across the entire corporation for whom you can hire, how much you can pay them, or what ownership they can have in the business and so forth. Why many of the major electronics companies failed in the semiconductor business is that they didn't appreciate the enormous value to them of a few individuals who knew about the art, and those individuals tended to get dissatisfied and leave and start new companies that were the ones that were so successful in the market place. I think that is one reason why it is often difficult in a major company. Another reason why it is difficult in a major company is that sometimes people divide things up. You know, I go and have a beer with you and I say "Vladimir, I know this semiconductor business is great stuff but I have to make my living in the tube business." So OK, we are heads of two different divisions in the company; I say "Look why don't we come to an informal agreement. You go this way and I'll go this way and we won't fight one another and life will be a lot more pleasant." I think that tends to happen too when you are trying to compete internally. So there are many reasons why it is difficult. I think it is often also the case that the vice-president of a division in the old technology can fight very hard and effectively against the new thing, and I've seen this happen in a few instances. I think one of the most courageous and difficult things for a company to do is take its best resources from something it is doing very well and put them into something that is very new and uncertain. It it very clear that you can make major gains from taking small steps that give quick returns spread over a large volume of production in businesses like the ones at the right-hand extreme, and so you can continue to innovate incrementally in the old technology until you go bankrupt. When you really begin to pay attention to it you can make more major gains, when you really start putting resources into it, you see it going very well and becoming even more profitable for a while until it dies. We call this kind of business a "cash cow" in the U.S. The idea is from a strategy point of view you are supposed to milk the cow to get the resources to put into a new business - you are not supposed to let the cow at the same time consume all the food and resources you ought to be investing into something better.

Later, Utterback commented

Large companies tend to fail by repeating their pattern of past success, of too much capital and organization too soon, while small companies tend to fail by underpricing.

The difficulty of having organization structures that accommodate innovation was also referred to by Gold:

Gold: One of the most difficult problems to deal with in an advanced technology organization is the tendency for it to be dominated by scientists and engineers of a certain type of specialty who will not let the new kinds of technical specialists in. United Shoe Machinery, for example, was dominated by mechanical engineers. You could not hire an electrical engineer in that firm. Finally the vice-president told us that in order to match competitors they had to set up a new research laboratory in a different city so that they could hire electrical engineers. The mechanical engineers then performed unbelievable feats of ingenuity in trying to match by mechanical means what the electrical engineers were turning out with ease.

A further problem is the timing and manner by which an organization's structure changes over the life cycle of the product:

Gold: But you can't stay at the left-hand side indefinitely, so when do you switch over? If you have an advantage and are doing well in the market with a new product, one of the ways to attract more capital and begin to increase earnings substantially is by moving towards the other side. You don't want to stay on the left, you want to exploit its benefits by moving towards the right.

The discussion then went on to the wider implications of this process.

Gold: Not only is this relevant for individual company policy but it bears on our discussion about Britain's problem because big employment is not on the left side of the model. Contrary to much of the chatter about industrial policy, no major industrialized country can afford to have everybody in the exciting but embryonic industries developing new technologies. You can't replace an automobile industry with 500,000 employees by 25 to 20 firms employing 50 to 500 persons in exploring exotic technologies. So once again, the essence of what you are saying is that one must utilize the full range of developmental stages. You can't stay on the left, but if you concentrate solely on the right you will not survive; as with a living population, you must have dynamic baby firms as well as an array of progressively more mature firms, including some approaching the infirmity of old age.

Utterback: I would only like to add one or two things to that. One is that I would want to see a climate where many new small things could start and sort themselves out and experiment and try to move along. That is very important; because if you don't move along you get eaten up by some competitor, perhaps the Japanese, that can imitate and produce and adapt and put technical resources behind the development of the product to push it along that way. The other thing is that I wouldn't protect the established companies from rivalry. To add to your comment on employment: I think the real gains in employment are created by the companies moving across the middle. Betts (Chapter 5) showed fascinating data where production was growing at 15 or 20 percent a year and employment stable. That is what happens on the right. What happens in the middle is that, as production grows, the process is starting out in a rather disorderly way and becoming more and more organized. That is when you get the big gains in employment and economic growth. I don't know how you feel about Burton Klein's idea in *Dynamic Economics.* Klein (1977) contends that it is when products and firms and market shares in your economy are rapidly changing that your overall economy tends to be rapidly growing, unemployment tends to be low, inflation low, and productivity advance high. When participation in the economy, products, and market share are stable, then you have rising unemployment, rising inflation, poor levels of productivity advance, and stagnating economic growth. I think that is a fascinating hypothesis.

Finally, the discussion returned to the question of the effects of increasing scale and its effect on innovativeness within the company.

La Porte: As we think about increasing the size of organizations there is a tendency to do things which reduce the uncertainty of managers at the top as the scale of the organization grows. The thing that happens is something we are all familiar with; the larger the organization the more tendency for the overhead to standardize. Why they do that is not really for efficiency but to reduce their uncertainty and diversity and it has to go back to this notion of the management problem of facing complexity when you haven't got a way of thinking about that from your past experience. (When innovation occurs) a manager who has more or less got his job together will experience increasing anxiety, so it all adds increasing uncertainty regarding the consequences of change for that person. Gold's example of the specialist who wouldn't let somebody in falls into that kind of a situation. It seems to me that (given the way we think about management) increasing scale forces us to do all the things which make it hard others to try change things. We punish people inside organizations as they get larger for attempting to do the sort of things Utterback was talking about – in business organizations (the same is true in political organizations) it is a very uniform phenomenon, mainly connected, I think, with the way we think managers should behave. It is a product partly of the way we believe we should behave as managers and the kind of reward structure that we lay on people in organizations as they grow so as to make it hard, nearly impossible, to reward internally the kind of behavior that Utterback has been asking for; I think it is a major cost of scale insofar as we think about scale in the way we have in the past.

15.1.3 Gold

Questions of innovation were also raised in the discussion on Gold's paper.

Gold: One of the interesting things about Japanese companies is that they have been more successful in developing and applying new technology. One of the reasons is that there has been a strong commitent to innovations, to the new technologies, from the top management on down. A second reason is that this is a longrun commitment.

Let's take computerization as an example. Many Japanese top managements seem convinced that pervasive computerization is the basis for eventual centralized planning, control, and integration of all operations. Consequently instead of trying to retrofit computerization and computer experts into the organization of existing plants, the computer people were included from the outset in the organizations which designed the giant Kimitsu and Kashima steel mills. Their thinking and operating capabilities were embedded in the very structure of plant design, equipment planning, and managerial arrangements. There was no question of whether they would or would not be accepted. They were part of the managerial machinery from the beginning.

Moreover, instead of treating computer specialists as we do – as a separate group who will ask questions of the people whose operations they are supposed to program – the Japanese often require computer specialists to acquire some actual experience in running such operations. And the operating people are often put through intensive training programs in systems analysis and computerization. Then the two groups jointly program the computerizing of a new operation on the basis of the overlapping insights of each participant instead of the disparate background and interests of two separate subgroups. This obviously ensures a fuller grasp of realities, a tendency which is further reinforced by awareness that the resulting operation may be turned over to one of the computer specialists instead of the former operating staff.

This represents a promising approach to introducing new kinds of specializations and innovations. But it contrasts strongly with how it is done in most places.

Finally, moving to the wider implications of innovation.

Apter: . . . The social consequences of these kinds of innovations, the changing character of the working force, the gradual erosion of the historic industrial working labor component, the enlargement of the technical and managerial side is, I suppose, an extension of the very old problem of technological unemployment. But the question then does come up: How does one handle this as a permanent problem? That is to say, it is quite clear that no plant manager or someone responsible for making investment decisions can really take on the burdens of total consequences in a society. So just as a question: How do you handle that? If you do, how should one really begin to think about this, and, in a setting like this, are there comparative experiences which suggest alternative ways for dealing with the problem?

Gold: There are two levels to the question you have asked. The first level is how does labor in the affected plant react at the prospect of an innovation and how can you get it accepted. Then comes the larger social problem: if this results in decreased employment and the development of pools of unemployed labor, how do we prevent the trade unions from putting a stop to further technological advances. Level one has been marvelously handled in Japan It results not from top management's active pressure for the development of new technology, but also the readiness, and in many cases even the eagerness, of labor to accept new innovations – an enormous contrast with common experience in Western economies. I have been through a variety of Japanese plants in automobiles, electronics, steels, shipbuilding, computers without encountering any evidences of labor resistance to technological innovations. One of the fascinating questions that is posed by such observation is: why does Japanese labor readily accept technological change? And its parallel, of course, is: what are the major threats that prevent labor in Western industries from doing it?

In answer to the latter, there are two key fears: the threat of losing jobs, and the threat of losing skills, current wage lavels, and seniority. Well, neither of these operates in Japan's large industrial firms. They have a guaranteed lifetime employment in the plant, so job loss is not feared. Secondly, they don't have jurisdictional boundaries among different skill categories, so they needn't fear being demoted to a lower wage category. Indeed one of the most fascinating you discover in Japan is what are they doing with the labor in a recession. Labor not needed for plant operations is shifted to maintenance, rebuilding, and even to building additional capacity. Hence, when the next upturn comes, they are going to have more and better capacity than before. So, they not only take care of their labor, they utilize it. Currently, with 70 percent capacity utilization in the steel industry, the companies have loaned out some of their labor to other firms with whom they have associations, thus easing problems for both sets of firms.

Turning to the question of the effects of technological advances on national unemployment, Gold commented:

This is a very serious problem in the U.S. as well as in England and Western Europe. What do we do about it? I think we have no alternative to what you have said. This is a governmental responsibility and my feeling about the matter is that the goverment must make provisions for transitional support, training, relocation, etc. But it has got to be of an order of magnitude big enough to dissolve the widespread resistance to accepting the technological advances necessary to safeguard competitiveness, and hence jobs and incomes.

So the point you have raised is obviously most important. We must make it possible for labor to recognize and accept the inevitability of technological change without excessively burdening the people directly affected. It is going to take a tremendous amount of money, of course, but not as much as it will cost if you don't do it.

15.2 COMMENT

Utterback's presentation and the discussion by the workshop participants indicate that there is a complex relationship between scale and innovation. On the whole there seems to be general agreement that, as far as incremental process innovations are concerned, large firms with clearly set managerial targets and well-defined organizational structures are the most effective in identifying and using such innovations. It is of interest to note, and it was pointed out by Betts in discussion, that just increasing the size of a plant is an innovation and it is generally accompanied by increasing specialization of function, the coordination of which requires clearly defined organization structures and rules.

For major product innovation there appears to be no reason why large firms should have any particular advantage. Neither the workshop participants nor the literature seems to have clear evidence that there is a significant difference in innovativeness between large firms and small firms. Although some writers have advanced evidence that indicates that small firms are the source of more key innovations than large firms (e.g., Roberts 1975) there are many more small firms than large firms so, when the number of innovations is related to either employment or value of output, small firms do not appear to be more innovative (Freeman 1971). Some small firms in some industries are very innovative while in other industries their role is minimal.

However, as far as major process innovations are concerned, there were a number of interesting aspects discussed at the workshop. Gold's comments on the attitude to computerizaton in Japanese industry (cf. Gold 1978) suggest that it was being treated as a major process innovation, to be handled by task forces. On the other hand, the American steel industry left computerization to staff specialists, indicating that they considered it to be a minor innovation. The difference in approach shows that it is not at all easy for management to recognize a particular innovation as major and adopt the appropriate organizational measures. One can contrast Gold's example with the experience of the many companies in the 1960s who treated the introduction of computers for business data processing as a major process innovation, yet all that could be achieved economiclly with the then batch processing oriented technology were payroll and some accounting functions. As a result, much of the managerial time devoted to computer task forces was wasted and could have impaired profitability and performance.

Van Hess's comments on the need for an organization to have a certain critical size in order to exploit a major process innovation such as microprocessors in manufacturing are very significant. Combined with Utterback's later comments on the dynamic effects of such innovations, the rapid change in relationships within the organization, it suggests that such innovations are best exploited by medium-sized organizations where there is still flexibility in managerial functions, good communication between different areas, and close identification with overall company goals.

Since a process innovation in one industral sector is often a product innovation to another sector (e.g., a machine tool supplier to the automobile industry) there is an implication that the optimal structure of industry for an innovative economy will be a mixture of small, inventive product-oriented firms; closely knit medium-sized firms with sufficient competent well-trained managers; and large capital-intensive firms with effective cost control and management systems. The crucial problem seems to be: What happens to the large firms as they become dinosaurs? Utterback mentioned that too close a contact between them and the rapidly growing medium-sized firm is fatal for the successful development of the medium-sized firm, so mergers are not advisable. On the other hand, their accumulation of commitments to large-size old technology means that breaking them up into smaller "dynamic" units is not likely to be generally feasible, although it is interesting in this context to note counter-examples of the General Electric Company in the UK. This group, built up over the years by takeover battles and amalgamations under the aggressive leadership of Arnold Weinstock, and still of modest size in the world league of such companies, is now reported (Smaller is Beautiful for Britain's Giant GEC) to be seeking legislation enabling it to disband itself, turning the operating divisions into legally independent companies.

However the dynamics of innovation are handled, their inevitable consequences tend to be left as extremely difficult social problems in the hands of goverments who have then to choose between preserving the old, but stifling or distorting the dynamics of innovation, and promoting innovation, but having to confront local unemployment or social hardship. In the United Kingdom, for example, the innovative and fast-growing electronics firms are clustered in the southeast, in such "silicon valley" regions as the town of Worthing, while the traditional industries now in decline, such as textiles, shipbuilding, and steel, are concentrated in the north and west.

It is apparent that the scale implications of Utterback's model of the innovative process over the life cycle of a product is a topic warranting a significant research effort.

REFERENCES

Abernathy, W.J., and J. Utterback. 1978. Patterns of Industrial Innovation. Technology Review 80 (7):1 - 9.

Freeman. C. 1971. The Role of Small Firms in Innovation in the United Kingdom since 1945. Committee of Inquiry on Small Firms. Research Report No. 6. London: HMSO.

Gold, B. 1978. Factors Stimulating Technical Progress in Japanese Industries: The Case of Computerization in Steel. The Quarterly Review of Economics and Business 18 (4): 7 - 22.

Klein, B.H. 1977. Dynamic Economies. Cambridge: Harvard University Press.

Roberts, E.B. 1975. Technology Strategy for the European Firm. Industrial Marketing Management 4:193 - 198.

Smaller is Beautiful for Britain's Giant GEC. 1980. The Economist. 274 (7119):75 – 76.
Utterback, J. 1978. Business Invasion by Innovation. CPA/WP-78-13. Cambridge: Center for Policy Alternatives, Massachusetts Institute of Technology.

Part 4

SCALE AND NATIONAL INDUSTRY POLICIES

CHAPTER 16 INDUSTRY SCALE, FREE TRADE, AND PROTECTION

M.F. Cantley and J.A. Buzacott
International Institute for Applied Systems Analysis,
Laxenburg, Austria

16.1 INTRODUCTION

A strong commitment to free trade has been the hallmark of "liberalism" for well over a century, as passionately expressed by The Economist in 1843 (quoted by Calleo and Rowland 1973):

> Free trade is itself a good, like virtue, holiness and righteousness, to be loved, admired, honored and steadfastly adopted, for its own sake, though all the rest of the world should love restrictions and prohibitions, which are of themselves evils, like vice and crime, to be hated and abhorred under all circumstances and at all times.

This philosophy has been reiterated in the Treaty of Rome which founded the European Economic Community, and in the founding articles of the General Agreement on Tariffs and Trade (GATT), the International Monetary Fund (IMF) and the Organization for Economic Co-operation and Development (OECD).

However, the subject of scale provides a direct route to the perception of some glaring deficiencies in the conventional wisdom of liberal market economics on the subject of international trade theory. In this chapter we present briefly as our starting point the theory of "comparative advantage," as enunciated by Ricardo in the early nineteenth century. The ideological convenience of this argument to Britain in the nineteenth century and to other industrially powerful nations in more recent years may have helped to mask its technical deficiencies; but the interaction of scale effects with unconstrained trade leads to situations that call for fresh considerations of policy and basic assumptions, at both national and regional levels.

16.2 THE THEORY OF COMPARATIVE ADVANTAGE AND ITS ASSUMPTIONS

Ricardo pointed out that trade between countries was always beneficial, even where one country was superior in all production sectors to the other, because of inevitable differences within each country in the relative efficiencies of producing different goods. Suppose, for example, that country A can produce 10 cars or 20 tons of potatoes per man-year, and country B can produce only 8 cars or 12 tons. In country A, a car trades for 2 tons of potatoes, and in country B it can be had for one-and-a-half tons: country B's car industry will certainly make sales in country A, if allowed. Conversely, a ton of potatoes costs two-thirds of a car in country B, but only half a car in country A: country A can export its potatoes to country B. It is the traditional definition of a bargain: an agreement from which both sides gain. Country A specializes in potatoes, country B in cars, a global welfare is maximized.

An assumption usually left implicit in economists' presentations of the case for comparative advantage is that the world is peaceful, trusting, and unchanging. Thus in our simple example, country B can eat their seed potatoes and concentrate on cars, trusting that the need of country A for cars will be as steady and continuing as B's need for potatoes.

Meanwhile country A, specializing in potatoes, can abandon their engineering skills, unworried by the prospect that their recovery of these skills might be problematical if future changes in demand or technology ever made such skills desirable again.

16.3 DEFICIENCIES IN THE THEORY OF COMPARATIVE ADVANTAGE

Of the many deficiencies in this simple theory, two are particularly relevant in the current context: first, the linear assumption of constant returns to scale; second, the assumptions of a static (or pseudostatic) world in which the dynamics of changes in markets and technology and of the accumulation of experience are ignored.

16.3.1 The Assumption of Constant Returns to Scale

Consider first the implications and validity of the assumption of constant returns to scale. In practice, there are diminishing returns to scale in agriculture (i.e., as total output expands, marginal land is brought into production, and expansion by increases of labor and fertilizer inputs is similarly limited). In industry, there are increasing returns to scale over a wide range. Thus, far from the general international equalization of wages, interest rates, and living standards to which the standard theorems of trade theory lead, consideration of scale effects suggests that as trade increases so does the asymmetry between the economic strengths of the trading partners, with the balance of advantage lying with the industrialized countries; this model seems to fit better the empirical data of the last two centuries.

Calleo and Rowland (1973) have given an excellent history of the evolution of free trade ideology on both sides of the Atlantic, and as they point out:

> Ricardo's ideas had a rather special application to Britain's politics in the early and middle nineteenth century. The British had gained a formidable lead over other nations in industry and commerce; British manufacturers were more than capable of competing favorably in any open market. British industrialists naturally hoped to extend and consolidate their position as "the workshop of the world."

Internally, the rapid and forced transformation of British agriculture caused social distress, and criticism by humanist conservatives. Coleridge, Disraeli, and others attacked the disruption of rural society for the sake of cheap food and low industrial wages; but they lost the argument. Externally, the pragmatic Americans remained protected not only by geography and high tariffs, but by an unregenerate nationalist mercantilism. The American economist Henry Carey denounced free trade as Britain's policy to perpetuate her supremacy and reduce all agricultural states to permanent tributaries. Gradually, the European states returned to mercantilism, emboldened by the works of the Swabian-American Friedrich List. List argued that while free trade was the natural view of a powerful developed nation, for a nation seeking to develop its industries, it was a short-sighted policy which sacrificed long-range national interests and productive power. He emphasized productive power rather than increased consumption because such power is "infinitely more important than wealth itself."

This history remains of central relevance to the continuing debate over liberalism and protectionism of various forms. For the debate is not only about the mathematical models of theoretical economics, but also about the nature of the international trading environment, the desirable scale and form of free trade groups, and the validity of assumptions about production capability.

On the nature of the international environment, Keynes, writing in 1933, was clear on the need to revise his strategic fundamentals:

> The policy of an increased national self-sufficiency is to be considered, not as an ideal in itself, but as directed to the creation of an environment in which other ideals can be safely and conveniently pursued. . . . we have until recently conceived it a moral duty to ruin the tillers of the soil and to destroy the age-long human traditions attendant on husbandry, if we could get a loaf of bread thereby a tenth of a penny cheaper.

Echoes of Disraeli!

The classical international free trade model ignores the possibility of any disturbances to production, transportation, and trade. That is, as List pointed out, it assumes international and domestic peace as a given condition of its analysis. In the real world, national security obviously could not be taken for granted. List argued that, for reasons of security, states should not be overly dependent on other states, and should therefore strive for a balanced and relatively self-sufficient economy. Domestically, List also saw economic activity as a great "collaboration" of labor within a smoothly functioning social system. Essential to this collaboration was the existence of a peaceful, protected community, which contained not only the necessary skills, but also the capacity to put them to work.

Notwithstanding the circumstances of his time, List foresaw a stage in European development when free trade would become beneficial; but according to Calleo, List's view was that "a plural system can remain open only insofar as it does not unduly threaten the cohesion and self-determination of the national units."

To operational researchers, the liberal economist's model is familiar in the guise of the prisoners' dilemma, translated into the terms of Table 16.1. On this model, GATT negotiations, Common Markets, and similar activities represent an attempt to build the trust and cooperation to bring all parties into the top left-hand corner, and prevent defections into the adjacent boxes. Each partner is tempted to defect from the agreement, but is inhibited by the general fear of the bottom right. It is a crude, simple model, whose validity depends on assumptions increasingly questionable; in particular it ignores the dynamic aspects of scale economies in the growth of industrial capability at all stages, and at all levels from individual products to general social infrastructure.

TABLE 16.1 "Prisoners' dilemma" model of international trade protection/liberalization issue.

		Country A	
		Liberalization	Protection
Country B	Liberalization	Satisfactory for both	Better for A, worse for B
	Protection	Better for B, worse for A	Worst case for both

16.3.2 The Assumption of a Static World

This leads to consideration of the second restriction of the classical theory of comparative advantage – its neglect of dynamic effects. There are two dynamic effects of importance. One is the potential cost reductions achieved with cumulative experience (cf. Chapter 6). The other is the effect of continuing high annual growth in demand on the extent to which economies of scale in plant construction costs in capital intensive industries can be exploited (cf. Chapter 5).

If early entrants in an industry are able to exploit the learning curve phenomenon, a natural tendency towards concentration or monopoly is the result, since the entry barrier becomes progressively higher. This tendency is limited by natural barriers of transport cost, and other advantages of proximity to customers, possibly supplemented by protective measures ranging from low tariffs to total prohibition. But where the economies of scale or cumulative advantage are great enough, a single global monopoly is a conceivable outcome, e.g., in such high-technology areas as very-large-scale-integrated circuits, aero-engines, or some sophisticated sectors of pharmaceuticals. One might note in passing the important role of patent laws in defending the acquired information likely to have resulted from the greatest cumulative experience.

For developing countries, even the liberal economists accept the case for some protection to build up domestic industry. Moreover, few countries that have struggled to build up a domestic manufacturing capability will tolerate its subsequent elimination by foreign competition, even once its infancy is over. Vietorisz (1974) gives a good illustration of this in his description of the Mexican electric motor industry, in a paper that also has much to say about the inadequacies of "comparative advantage" in allowing for the dynamic and structural aspects of industrial development:

> . . . technology transfer . . . appears as a means of perpetuating dependency, or con-

> trariwise, breaking out of it . . . . Something crucial is evidently left out of the comparative advantage model – namely, the analysis of the development of the institutional structure supporting industrialization and technological progress, which has a sequential character, reflected in the sequence of introduction of electric motors of progressively larger size and progressively greater technical complexity.

Vietorisz also quotes the nineteenth century U.S. and Japanese experience; then, returning to his theme of Third World development problems, on which he was working for the U.N., he comments

> The enormous weight of economic tradition supporting the comparative advantage principle makes it difficult to define effective criteria . . . . Advice given under the traditional point of view is likely to transfer technology in such a way as to perpetuate dependency rather than help break out of it. Yet the protagonists of these views have the most impeccable academic credentials from some of the world's leading institutions of higher education.

More recently, the word *dependencia* has come into prominence as a central theme in the literature of Latin American economists, as they reiterate the substance of Vietorisz's argument. The work of Chichilnisky and Cole (1979) demonstrates by formal algebraic models the perverse effects that can result from the dynamics of trade – perverse, that is, with respect to the welfare gains predicted by the classical trade theory. The effect of domestic income distribution, and the availability and elasticity of factors of production (skilled and unskilled labor, and capital) interact with domestic technologies and with North-South trade (particularly if an export-led growth strategy is adopted) to create a pattern in which inequality is reinforced and the absolute welfare of the poorer groups may be reduced.

Kaldor (1978) has usefully distinguished between three types of trade:

1. Trade between manufacturing countries and primary producers
2. Trade of manufacturing countries with one another
3. Trade of primary producers with one another

In the third case, the variations of climate and topography make the concept of specialization and comparative advantage appropriate, and trade will be beneficial. The first case we have already discussed.

The second type of trade is the most interesting case: it has greatly increased since World War II. Kaldor points out that this type of trade is characterized by significant imbalance in favor of countries with fast growth:

> After the second World War the successful countries like Germany and Japan acquired a cumulative advantage through their fast growth, whereas the slow-growing countries such as Britain (and to some extent also the Unites States of America) faced an increasing handicap due to their slow growth. On account of the dynamic effects of fast growth it is possible that a particular country's products become qualitatively superior, and hence preferred to those of another country, in all

branches of industry. It can be argued that Japan's growth (and to some extent also Germany's growth) in the post-World War II period was enhanced at the expense of the two trade-losing countries, the U.S. and the U.K. These latter countries suffered from increasing import penetration in their domestic market of manufactures which was not offset by higher exports, which meant in effect that their national output was reduced in consequence.

In capital intensive industries where there are economies of scale in plant construction costs, Manne's model of optimal plant size (Manne 1961) provides quantitative justification for this. He shows that the key variable is the amount of demand growth per year. Hence low costs can be achieved in large economies with moderate growth (such as the U.S. chemical industry) or in medium economies with high growth (such as Japan). Medium-sized economies with low growth and small economies are inherently going to be high-cost producers of the commodities where there are significant economies of scale in plant investment unless they have unique advantages in terms of either availability or cost of raw materials.

16.4 POLICY IMPLICATIONS AND LEVELS

16.4.1 Implications for Industry Policy

At the level of the firm and the industry, the mounting evidence of the advantages of cumulative experience and dominant market share is becoming more widely accepted. For example, in the UK goverment's consultative document on Monopolies and Mergers policy (Secretary of State for Prices and Consumer Protection 1978), one finds extensive citation of the Boston Consulting Group's evidence on learning curves, with discussion of its implications for policy:

> . . . a consideration of the combined effects of learning, scale and technology, their apparently systematic operation, their possible links with market share and their predictable consequences for profitability via cost advantage, underlines the critical roles assigned to market share and concentration in competition policy literature but suggests a rather different interpretation of their significance. The Boston Consulting Group argues that, not only does this incentive drive industries towards concentrated structures; it has highly beneficial results. Thus, "there is an implication that the consumer is best served by letting the dominant producer emerge, or even encourage his development and the concentration of production." The divergence of this interpretation from the traditional economic analysis of concentration is fundamental: according to the latter, concentration confers opportunities to exploit consumers via higher prices; according to the experience curve approach, concentration is the outcome of a process which confers on the leading producer a real cost advantage and it is this real cost advantage which maintains its superior profitability rather than exploitative behavior in the market or improper restraints on competition. To the extent that monopoly references and investigations hinge on the profitability of dominant suppliers, they may need to give due weight to the latter interpretation.

The implications for mergers are less clear. The accumulation of experience and the achievement of cost reductions have been assumed to occur within companies. In fact, industry concentration and gains in market share often occur through merger. Under these circumstances, it would not be valid to assume that the greater combined accumulated experience achieved through merger can be translated effectively into lower costs.

Simliarly, considering the relationship between market growth and scale decisions, one of the purposes of a strict competition policy in some Western countries is to ensure that firms make investment decisions independently. The total demand growth then has to be split among a number of competitors resulting in smaller plant sizes. Unless compensated by the marketing efficiencies resulting from competition or by the effects of distribution costs, this may result in higher costs in comparison with those countries with effective coordination of investment decisions (cf. Gold's description of decisions on blast furnace size in Japan (Gold 1974 p 11)).

As Calleo and Rowlands point out, all modern industrial states have governments that in practice are "mercantilist" or "interventionist" in the sense that they accept an obligation to interfere constructively in the unfettered operation of the market. The "infant industry" argument of transient protection or other deliberate intervention for a new industry is exemplified by the UK National Enterprise Board's £50 million funding of Inmos Limited in June 1978, in an attempt to secure a place for the UK in what is seen as an industry of major strategic importance in the future.

16.4.2 Implications for Trade Policy

At the same time, the UK has not made any government-level attempt to arrest the rapidly rising proportion of UK demand for finished manufactured goods which is met by imports; although the balance of payments deficits caused increasingly by the growth of such imports have acted as a brake on every spell of reflationary policy in which postwar goverments engaged from the early 1950s.

Thus the dynamic effects of the scale argument are perceived at levels 1, 2, and 3 (unit of equipment, plants, and the company), but lost sight of at the critical levels 4 and 5 (industry and society). This is the background to the contemporary debate in the UK on the subject of import controls, which have been advocated for several years by Godley and colleagues at the Cambridge University Department of Applied Economics (Economic Policy Review, 1976, and in Beckerman (1979)).

Godley and Cripps (1978) argue that certain countries in the world economy are intrinsically "balance of payments constrained," in the sense that attempts to increase the rate of growth of the economy lead to such rapid growth of imports that the growing trade deficit becomes unsustainable, and expansion has to be slowed. (Devaluation of the currency requires such a change to achieve balanced payments that it leads to unacceptable domestic price increases and a consequent inflationary spiral.) Thus the level of activity is constrained below the full productive potential of the economy – in other words, there is unemployment. Their prescription is that, rather than control the level of imports by deflation and unemployment, it should be deliberately controlled directly (at a level not less than it would otherwise have been); and then domestic reflation resumed by traditional Keynesian means.

This debate touches the questions of scale at two points. First, Godley's argument rests upon certain parameters reflecting a country's propensity to spend marginal income upon imported rather than home-produced goods. In the case of the UK, on which Godley's arguments were developed, the fact that demand expansion stimulates imports of manufactures rather than domestic production is related to a cumulative, long-term cycle of manufacturing decline. High-growth industries abroad exploit scale advantages and produce at lower cost. An industry that perceives a low-growth future cannot build new capacity that will be competitive with imports. This leads to inadequate investment, inadequate profit margins, and a continuing vicious circle of decline. The cumulative dynamic effects at the national level can be seen; for a deeper historical analysis of the origins of this cycle, the reader may be referred to Barnett (1972).

The second scale-related point is that the effect sought by Godley's proposal is, in system terms such as those used in Chapter 14, the creation of a less tightly connected environment, diminished interdependence, and thereby the amplification of possibilities for a domestically determined economic strategy. For a fuller discussion, see Cantley (1979).

Generalizing from this discussion, we may observe again the "overshoot" phenomenon. The basic arguments in favor of trade, i.e., comparative advantage and specialization, clearly have some validity; free trade implies division of functions and specialization of role, a key component of Bonner's morphogenesis (Chapter 1, section 3.1); however, eventually this leads to a situation in which the degree of interdependence is such that diseconomies of insecurity and inability to cope with disturbances and control one's system outweigh the further advantages. Given the cumulative dynamic effects on beliefs, and on the institutional expressions of those beliefs, it may be very difficult to create the will for the necessary major policy changes. A high cost in unemployment and social problems may have to be paid, or may be seen as inevitable, before the necessary system changes are seen as acceptable.

In the long run, a change of direction, a retreat from the extreme of technical possibilities in scale, specialization, and complexity may come to be seen as desirable in the design and operation of social or national systems as it has become in the design and operation of large-scale industrial plants (see Dathe (Chapter 3), Fisher (Chapter 4), Betts (Chapter 5)).

REFERENCES

Barnett, C. 1972. The Collapse of British Power. New York: William Morrow.

Beckerman, W., ed. 1979. Slow Growth in Britain. Oxford: Clarendon Press.

Calleo, D.P., and B.M. Rowland. 1973. America and the World Political Economy: Atlantic Dreams and National Realities. Bloomington and London: Indiana University Press.

Cantley, M.F. 1979. Scale, Protectionism and European Integration: The Structural Dynamics of Strategic Control in a Turbulent Field. WP-79-42. Laxenburg, Austria: International Isntitute for Applied Systems Analysis.

Chichilinsky, G., and S. Cole. 1979. A Model of Technology, Domestic Distribution, and North-South Relations. Technological Forecasting and Social Change. 13: 294 – 320.

Economic Policy. Review. 1976, 1977, 1978. Department of Applied Economics, University of Cambridge. (Especially 1976, chapter 4, The relative decline of the U.K. manufacturing sector.)

Gold, B. 1974. Evaluating Scale Economies: The Case of Japanese Blast Furnaces. Journal of Industrial Economics 23: 1 – 18.

Godley, W.A.H., and F. Cripps 1978. Control of Imports as a Means to Full Employment and the Expansion of World Trade and Output: The U.K.'s Case. Cambridge Journal of Economics. September.

Kaldor, L.N. 1978. The Foundations of Free Trade Theory and its Implications to the Current World Recession. Paper presented at International Economic Association conference, Bischenberg, July 1978, and to be published in proceedings thereof.

Keynes, J.M. 1933. National Self-Sufficiency. Yale Review 22: 755 – 768.

Manne, A. 1961. Capacity Expansion and Probabilistic Growth. Econometrica 29 (4): 632 – 649.

Secretary of State for Prices and Consumer Protection. 1978. A Review of Monopolies and Mergers Policy: A Consultative Document. London: HMSO.

Vietorisz, T. 1974. Diversification, Linkage and Integration Focus in the Technology Policies of Developing Countries. In: Transfer of Technology for Small Industries. Paris: OECD. pp. 57 – 77.

CHAPTER 17 SCALE ECONOMIES AND THE OPTIONS FOR A SMALL COUNTRY

Donald J. Daly
York University,
Downsview, Ontario, Canada

This paper was prepared to raise at the workshop some of the issues and choices that small countries have in making strategic choices about industrial strategy and commercial policy. In writing it up more formally for the proceedings volume on the conference, it has been revised and extended to relate it to other papers and the discussion at the workshop, but not in a way to destroy the appropriateness of the lively discussion that took place in response to the oral presentation at the workshop.

The paper will discuss four areas. Initially, the terminology will be clarified, and then there will be a discussion of the importance of market size in achieving the full potential of scale economies in large markets. The third and longest section will deal with the choices open to small countries, and examples of countries following the various options. The concluding section will consider some of the problems that a small country must recognize as being involved in choosing the international specialization route.

17.1 INTRODUCTION

It is important to be specific about the concepts of scale economies (as there are a number that have been distinguished in the literature and in the workshop). In the commodity-producing industries (which dominated the papers and discussion at the workshop), a major part of value-added costs typically are incurred on the production side. Recent literature distinguishes between product-specific economies of scale and plant-specific economies of scale. Product-specific economies of scale typically relate to the reduction in average costs per unit with longer runs and higher volumes. These can occur from the spreading of overhead costs (including set-up costs of adjusting machines for changing product specifications, etc.) learning by doing, etc. This concept is particularly relevant for some of the modern consumer goods involving product diversity, styling, etc., produced in a multiproduct plant. Plant-specific economies of scale relate to the variation in average costs per unit with alternative sizes of plants producing a standardized product, a concept that has had a much longer discussion in the literature.

In addition to these economies of scale on the production side, there are also non-production economies of scale open to larger firms that operate a number of plants in different locations, and even different countries. These economies can take place in advertising, research and development, and financing (cf. Scherer et al. 1975; Gold (Chapter 2) also emphasizes the importance of being explicit and specific about these concepts.

In industrial practice and experience, these economies of scale are inevitably intertwined with two different but related phenomena. One is technological change – new ways of producing the same product, or new products that can compete with existing products. Another dimension is management – the initiating and coordinating function within the firm. The training, experience, and ability to work with others are important attributes of management and their effectiveness can be important in the openness to change and the speed with which new technology is adopted. These are frequently interrelated with economies of scale in the historical experience of products, plants, and firms in unit costs (as portrayed in progress cost curves or experience curves), and these other dimensions came into papers and oral discussion at the conference on a number of occasions.

Product-specific economies of scale need special emphasis in this proceedings volume for two reasons. For one thing, product diversity is a central feature of the modern high-income, market-oriented economies, especially in the areas of monopolistic competition and oligopoly. In addition, the effects of short production runs on higher costs are frequently greater than the cost effect of plants that are smaller than the minimum efficient scale. As many of the papers dealt with standardized products (electricity, ethylene plants, coal, etc.), this topic did not receive as much emphasis in the workshop as a whole as it warranted, in my opinion.

17.2 SCALE IN LARGE MARKETS

Economies of scale refer to the experience of costs per unit under alternative conditions on the supply side. However, it is only possible for these economies of scale to be achieved in practice if the related demand is also present. This is a major advantage of large economies or markets over small ones. It is much easier to achieve economies of scale (without necessarily high concentration ratios in firms in individual industries) in large markets (or large economies) than in small ones.

The United States is the largest economy, on the basis of market size. Its population approached 214 million in 1975, with a high level of real income per capita. The Japanese economy had about 110 million persons in the same year, with a real income per capita of about 65 percent of the U.S. level. For northwest Europe, one can no longer talk about markets on the basis of individual countries, as the enlarged European Common Market has now achieved complete free trade in industrial products. The European countries in the European Common Market have more than 290 million people. The levels of real GNP per capita in the individual countries range from almost 80 percent of U.S. levels (for the FRG, France, and Belgium) down to about 50 percent for Italy. The larger total population in Europe is about enough to offset the lower levels of real GNP per capita, making the size of the European economy very similar that of the United States in terms of real GNP (data from Morawetz (1977, p. 93) and Kravis *et al.* (1978, p. 10)). The postwar development of this large regional free trade market is an important development in achieving scale economies.

The key point is that these large markets can take advantage of the potential scale economies on the cost side by access to large markets, either domestically or on a free trade basis, which provide an adequately large market on the demand side to sell the large volumes associated with low costs. If the economies of scale on the production side are large and the share of production costs in total company costs is high, these scale economies can be achieved without high concentration ratios if the market is large enough for a fair number of producing firms.

It should be recognized that market size is not the only factor that is relevant to achieve low unit costs and high productivity. The stock of capital and the age of the capital stock are also relevant. Although new technology moves fairly easily between countries, there are significant differences in the speed with which new technology is adopted in different countries. Education is a very important factor, especially when it is remembered that labor income is a very large share of net national income. Relevant dimensions of education for economic performance would include the general level of education (including literacy), vocational education, and experience and training on the job. These other considerations influence the differences in real GNP per capita and per person employed among countries with access to large markets.

For illustrations of all of these points in the context of comparisons of economic growth over time and differences in real income per person employed at a point in time between Europe and the United States, see Denison (1967) and Daly (1968).

17.3 THE OPTIONS FOR SCALE IN SMALL COUNTRIES

International trade is an obvious potential route by which a small country can obtain some of the gains from scale not possible in a small domestic market. A rough indication of the degree to which small countries are achieving interdependence with other countries in the world economy can be provided by the ratio of merchandise trade to GNP. Such ratios are typically higher for small countries than large countries. For example, the percentages of commodity exports to GDP for the United States, Japan and the EEC were 7, 12, and 11 (for extracommunity trade only) in 1976. For some of the smaller countries in Europe, total exports of goods were between 40 and 50 percent of GDP. On the other hand, comparable ratios for Australia and Spain were 14 and 9 percent. (Data from United Nations quoted in Industry, Trade and Commerce 1978).

There seem to be only three possible options open to a small country. Option 1 is to follow a policy of relative self-sufficiency, which can be attained by a significant degree of protection of domestic industry by high tariff and nontariff barriers to trade. The alternative options are to achieve a greater degree of specialization through low tariff and nontariff barriers to trade. One way to achieve this is through a regional free trade association with other countries, which we will regard as Option 2. Another way is through low tariffs, attained either as part of more comprehensive multilateral negotiations or unilateral action, which we will discuss as Option 3. We will review the effects of the options and give examples of small countries that have taken each of these routes.

Option 1 is the small country that uses high tariff and nontariff barriers to trade to achieve a certain degree of self-sufficiency, perhaps in industries that have high status either domestically or internationally, such as manufacturing. These tariff and nontariff barriers to trade tend to lead to higher prices for manufactured products in the country imposing them than the prices for comparable items in other countries, and thus consumers and producers pay higher prices. On the production side, producers are encouraged to establish plants that are less than the minimum efficient scale for that product, and to produce a wider range of products in plants of a given size than would emerge in a larger market with a greater degree of product specialization. This prevents the producers in small countries from attaining the potential economies of scale distinguished earlier in this paper and discussed in some other workshop papers. Some of the associated symptoms that can emerge are lower levels of productivity in relation to labor and capital inputs, high costs, and less openness to change. These tendencies handicap the manufacturing producers in developing export markets except in selected products.

These policies would receive no support from economists in the neoclassical tradition of Adam Smith and his modern descendants. Such policies would not lead to an efficient use of labor and capital resources.

It is also interesting that such policies for small countries would not be supported by Friedrich List, who provided the fullest alternative analysis and policy proposals of his period, for Germany and other less-developed and less-industrialized countries. He was a strong supporter of the elimination of the various barriers to trade between the various German states and their economic integration into a larger free market within a larger German *Zollverein*. He thus recognized that small states should group together into free trade areas to permit them to take advantage of the economies of scale. Furthermore, he also favored the adoption of free trade after a country had achieved high levels of real income. His long-term policy proposals were thus rather similar to the traditions of Adam Smith (see List 1885 pp. 89 and 115).

Are the costs of tariffs and nontariff barriers to trade large or small? Early estimates of the costs of tariffs suggested quite small costs frequently a fraction of one percent of GNP in the country concerned. These estimates covered only the costs to the consumer, as measured by the extent to which prices were higher than world prices. However, these estimates did not cover the production effects of less effective use of labor and capital on the supply side. When these other costs were included, the costs of tariffs were higher. In 1975 in Canada, for example, the costs of Canadian and U.S. tariffs on Canadian GNP were 8.2 percent of GNP – a very significant amount. Not since the great depression has there been a Canadian policy that offered such a substantial economic benefit (Wonnacott 1975 pp 177 and xxii; Daly and Globerman 1976 pp. 18 – 61). Similar estimates for the early 1960s were even higher, but some gains from tariff reductions under the Kennedy Round and the Canada–United States Automotive Agreement have since been attained.

There are examples of small countries that have followed a protectionist route. Australia (with 12 million people) is one, and they receive further protection by high transport costs since they are geographically isolated from other large manufacturing producers. (Tokyo is about 5,000 miles by air from Melbourne, Australia.) Canada (with 22 million people) was also in this category, but will have moved a significant distance from the tariff levels of the 1930s by the end of the 1980s once the Tokyo Round reductions have taken place.

Option 2 is the option in which a small country achieves specialization by entering a free trade area or common market for industrial products with other neighboring countries. Belgium, the Netherlands, and Luxembourg are early examples of small countries in the European Economic Community (with populations of 10, 14, and a third of a million, respectively). With complete free trade on industrial products, these small countries can have free trade access to markets of 290 million people, and a level of total real GNP approaching that of the United States. They still retain many aspects of political sovereignty (including foreign embassies, education, taxation, etc.) with local currencies (with some limits on exchange rate movements with other European countries).

TABLE 17.1 Indexes of real national income per capita for selected European Common Market countries, 1960 and 1973 (United States = 100, U.S. weights).

	1960	1973
Belgium	61	75.3
the Netherlands	61	68.4
France	66	76.1
FRG	73	77.4

Sources: Denison (1967, p. 22) and Kravis *et al.* (1978, p. 13).

It is interesting that by the early 1970s, the levels of real national income per capital in Belgium and the Netherlands (examples of smaller countries in the EEC with comparable data) were closer to those of the larger countries of France and the FRG than they had been in 1960. The levels of real income were also somewhat closer to that of the United States, and all the European Common Market countries (except Italy) had moved ahead of the United Kingdom. These changes can be seen in Table 17.1.

As part of the increases in real income and the narrowing in real income differences between the countries within the European Common Market, there has been a significant increase in the degree of specialization and a big increase in intra-European manufacturing. It is also interesting that these changes took place with essentially low levels of unemployment and low levels of bankruptcy among companies. There have, of course, been problems about regional differences and industrial adaptation in some countries, as pointed out in the discussion by Dr. Ansoff and Professor Gold.

Option 3 is the situation of countries that have achieved specialization and high real incomes by policies of low tariffs but with more autonomy on economic policy than may be feasible in a common market. European examples of countries in this situation are Sweden and Switzerland. Canada is in the process of shifting from a position closer to Option 1 to one closer to Option 3 by the end of the 1980s, but it is not one of the sharpest examples of an extreme situation. Sweden has specialized in such specialty fields as roller bearings and steel, but a recent Boston Consulting Group study points out a number of areas of vulnerability in traditional exports, partly associated with potential competition from developing countries. The study provides a framework of analysis to assist in identifying key areas for companies to concentrate on, and to isolate emerging areas of competition in developing countries – approaches that would be relevant for other countries in addition to Sweden (Boston Consulting Group 1978).

In the early 1970s, Sweden and Switzerland had levels of real GNP per capita close to those of France and the FRG; Sweden and Switzerland were slightly lower than the larger European countries.

These are examples of smaller European countries that have followed Options 2 and 3 and achieved high real incomes per capita comparable to larger countries in the same region. These countries are all involved in exporting some varieties of manufactured products, but the range of items both exported and produced is specialized and selective rather than covering a wide range of manufactured products.

For a discussion of intraindustry trade (emphasizing the theory, measurement, and testing of the concepts), see Grubel and Lloyd (1975). For an early study of country size and economic performance, see Robinson (1960). For an alternative interpretation of the effects of European economic integration, see Cantley (1979).

17.4 POTENTIAL PROBLEMS FOR SMALL COUNTRIES

An increased degree of involvement in the world economy by producers in small countries is bound to involve new problems as part of the movement into new and larger opportunities. A number of the more important of these will be mentioned.

There must be a reasonable number of products and firms that are already competitive or close to being internationally competitive, at current costs and exchange rates. These can provide an initial source of export earnings, and with further specialization they can develop further marketing outlets and broaden the range of products exported and the range of countries that can be exported to. These should build on existing and emerging areas of comparative advantage.

Increased economic interdependence can increase the risks of changes in demand, new competitors, or changes in policies and exchange rates on other countries. A greater amount of flexibility in corporate strategy and tactics would be necessary than when production and sales were taking place in a more stable, secure, and protected market.

One potential problem for a small country to consider in moving to a more open and specialized position in world markets is the abilities, training experience, and flexibility of its management. A history of technological and managerial backwardness could be an important handicap for a small country in moving into a more competitive international economy.

It was interesting in the discussion at the workshop how frequently the topic of management was mentioned both by speakers from the market economies and the planned economies, including academics, government people, and the business community. It is also relevant to the ease with which small countries can take advantage of scale economies by operating in larger markets through specialization in a smaller number of products and industries.

REFERENCES

Boston Consulting Group. 1978. A Framework for Swedish Industrial Policy. Boston: Minco.

Cantley, M.F. 1979. Scale, Protectionism and European Integration: The Structural Dynamics of Strategic Control in a Turbulent Field. WP-79-42. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Daly, D.J. 1968. Why Growth Rates Differ – A Summary and Appraisal. International Review of Income and Wealth. March: 75 – 93.

Daly, D.J., and S. Globerman. 1976. Tariff and Science Policies: Applications of a Model of Nationalism. Toronto: University of Toronto Press.

Denison, E.F. (assisted by J.-P. Pouillier 1967.) Why Growth Rates Differ   Postwar Experiences in Nine Western Countries. Washington, D.C.: The Brookings Institution.

Grubel, H.G., and D.J. Lloyd. 1975. Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Products. New York: Wiley.

Industry, Trade and Commerce. 1978. Canada's Trade Performance – 1960–1977, Vol. 1. General Developments. Ottawa: Supply and Services.

Kravis, I.B., A. Heston, and R. Summers. 1978. International Comparisons of Real Product and Purchasing Power. Baltimore: Johns Hopkins.

List. F. 1885. The National System of Political Economy. London: Longmans, Green and Co. Reprinted 1966. New York: Augustus M. Kelley.

Morawetz. D. 1977. Twenty-Five Years of Economic Development 1950 – 1975. Washington: The World Bank.

Robinson, E.A.G., ed. 1960. Economic Consequences of the Size of Nations. London: Macmillan, for International Economic Association.

Scherer, F.M., A. Beckenstein, E. Kauter, and R. Murphy. 1975. The Economies of Multi Plant Operation: An International Comparison. Cambridge: Harvard University Press.

Wonnacott, R.J. 1975. Canada's Trade Options. Ottawa: Information Canada.

CHAPTER 18 SCALE STRATEGIES FOR A SMALL COUNTRY – THE EXPERIENCE OF GDR INDUSTRY

H.-D. Haustein and G. Wittich
Hochschule für Ökonomie,
Berlin, GDR

18.1. THE PROBLEMS

In our view, scale has a broader meaning than size. Scale refers to the way productive forces are combined, while size implies the absolute degree of largeness or smallness. However, while scale and size are different terms and describe different economic concepts, they have a close relationship.

Obviously, the size of a country in terms of its area or population is an important factor influencing the potential for scale economies and the feasible scale strategies. But the influence is indirect and results, for example, from the amount of mineral and other natural resources. The potential for scale economies is influenced more directly by the size of the national economy, measured by such quantities as GNP, number of working people, fixed assets, level of consumption, the size of the internal market, etc. In a given social environment, both the size of the country and the size of its national economy limit the feasible scale strategies by restricting the natural and manpower resources and the scientific potential, constraining the accumulation power, and reducing the size of the internal market.

In particular, the size of the internal market limits the potential of small national economies to take advantage of economies of scale. One way in which small national economies can oversome this difficulty is to participate in a larger market.

As a result, the economy of small developed countries is characterized by a high ratio of foreign trade to either the national income or GNP. (In the case of the GDR, the ratio of foreign trade to national income was 65 percent in 1979.) However, this does not necessarily reflect the efficiency of the national economy or the influence of scale economies. It could be due to adverse terms of trade or an inappropriate industrial structure in relation to the pattern of demand in the external market. Thus, in order for a small national economy to participate effectively in international trade it is necessary for it to develop a strategy for the development of its industrial structure that will enable it to exploit scale economies. In the case of centrally planned economies this means that there have to be appropriate strategies both for the national economy as a whole and also for the different branches of industry. For the GDR the basis of such strategies is participation in CMEA, thus giving the advantages of a large planned market to smaller countries.

Economies of scale are not limited to production; they can also be in the fields of research and development, management, financing, sales, etc. (Cantley and Glagolev 1978). Economies of scale in the research and development field are now of special importance because the achievement of economies of scale in the production field is based on the transformation of scientific and technical knowledge into the production sphere. While the percentage of national financial and manpower resources going to scientific and technical purposes is very similar in highly industrialized countries, the absolute level of resources depends on the size of the national economy. This limitation of the absolute scientific and technical potential means that available resources should be concentrated on selected problems; for example, the main effort should be focused on creating future productive capability. In 1974 the goverment of the GDR decided on a first draft of a program for the long-term development of basic research in the fields of natural science, mathematics, and selected technical directions up to 1990. This first draft can be considered as a starting point for a strategy for scale in research and development.

Finally, the potential for scale economies is also determined at the level of the economic organization by such factors as the degree of specialization and concentration, the size and structure of plants and corporations, management capability, etc.

To summarize, the scale strategies of a small country are composed of a series of substrategies like:

- Strategies of production structure
- Strategies of research and development
- Strategies for participation in the international division of labor
- Strategies for economic organization

All these substrategies are interconnected in such a manner that it is impossible to say that one strategy will follow directly from another. In considering organization strategies it is necessary to take the historical background as a starting point because this background explains not only the given conditions and the conditions that have to be changed but also restricts future changes (both in size and in time).

18.2 THE HISTORICAL BACKGROUND

In 1950, our industry had 23,582 enterprises with approximately 42,000 production units (*Produktionsstaetten*). The share of socialist enterprises was 25.6 percent measured by the number of enterprises and 75.7 percent measured by the number of employees. On the one hand, we had a highly concentrated basic industry, for example, in chemicals, and on the other hand, many industry branches consisted of a large number of small and medium-sized enterprises.

From the very beginning of the planning system in the GDR, we were confronted with the task of determining the right scale and size of our production units and enterprises. This task was made more difficult by the need to overcome the disproportions induced by the division of the former German economy. So we had to establish new production units appropriate to the demands of our newly founded state and its economy. In the first 5-year plan (1950 – 1955), the rapid increase of production was achieved by an increase in the variety of different products. Subsequently, the increased demands of our foreign trade on the technical standard and effectiveness of our products required a more rational organization of our industry. So in most industry branches programs for specialization and concentration were drawn up and realized.

The effect of these programs can be shown in various ways. Table 18.1 shows the change in the size distribution of enterprises where size is measured by number of employees. In Figure 18.1 the data on size distribution is plotted on probability paper for the log normal distribution. The slope of the line is a measure of the degree of concentration of both enterprises and production units. It can be seen that the number of enterprises has fallen faster than the number of production units, indicating that what has occurred is more a process of centralization of management rather than concentration of production. The main reason for the centralization was the new conditions established by the complete nationalization of our industry. The former spontaneous process of concentration and specialization could be transformed into a centrally planned and guided one.

Characteristic of the long historical tradition of GDR industry is that most enterprises consist of small units in different territories. Our industrial corporation for production of socks and stockings (VEB Strumpfkombinat ESDA) consists of 200 production units in more than 70 towns and villages in 17 areas (*Kreise*). In 160 of the production units the number of employees is less than 100 and in 120 production units it is less than 50. More than 18,000 employees work in the corporation. When we try to assess the degree of concentration in this branch we have to take into account the fact that a high level of specialization was achieved in this corporation combined with the necessary flexibility. It is not possible to reduce abruptly the number of production units. More than that, smaller units are required in the future to meet fast-changing demand, to use local resources, and to produce needed special assortments of products.

However, we must state that we cannot be satisfied with the present level of organization in our industry. Table 18.3 shows that our metal-working industry is characterized by only a small share of specialized mass production and automated flow production. From 1965 the share of specialized mass production in mechanical production only rose from 8 to 11 percent. One reason for this is the fast growth of the range of articles produced. We can say that our metal-working industry produces more than two-thirds of the world range of major product groups in this branch.

In the industry of automation instruments, 280 basic products have the following distribution:

24 products with an output of more than 20×10^6 marks
38 products with an output from 5 to 20×10^6 marks
101 products with an output from 1 to 5×10^6 marks
117 products with an output of less than 1×10^6 marks.

The number of types is more than 50,000 and the number of articles is some 100,000. This illustrates our problem to reach economical series and to create better conditions for innovations. This situation leads to a low degree of concentration and a large amount of interlocking, which must be managed and controlled. On the other side in implementing new products we do not always reach the necessary scale for a short

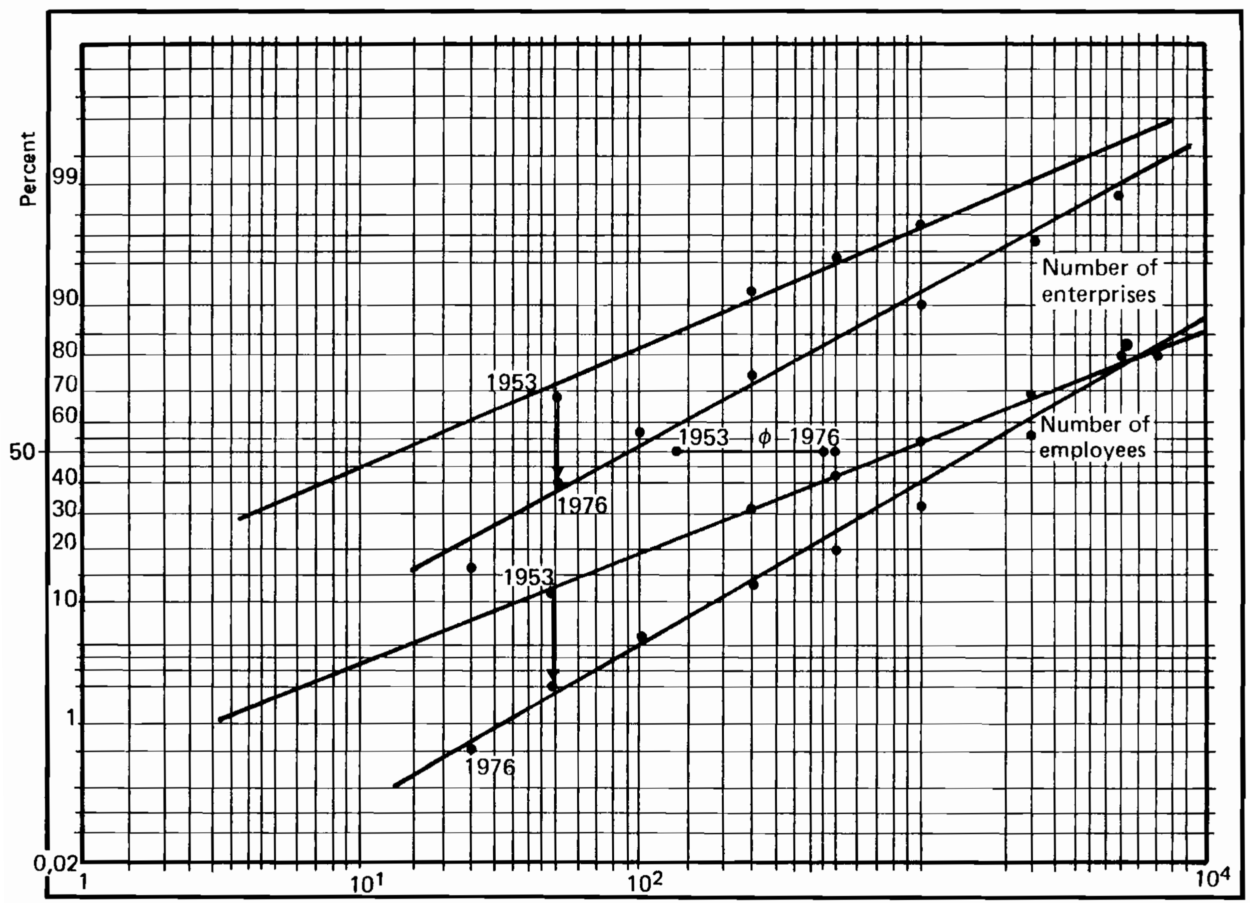


FIGURE 18.1 Frequency distribution of enterprises in GDR industry: 1953 – 1976.

Table 18.1 Size of Enterprises Measured by the Numbers of Employees (% of total number of enterprises).

	Number of employees up to 50	51 - 200	201 - 500	501 – 1,000	over 1,000
1953	69.2	20.4	5.4	2.4	2.6
1977	31.5	33.0	16.3	7.9	11.3

SOURCE: Statistisches Jahrbuch der DDR (1953, 1978, 1979).

Table 18.2 Indicators of size and scale in GDR industry

	1950	1960	1970	1977
Number of enterprises	23,582	16,038	11,700	6,480
Number of production units[a]	42,000	40,000	39,000	
Number of employees (x 1,000)	2,098	2,768	2,855	3,083
Industrial output (index; 1950 = 100)	100	294	536	812
Consumption of electrical energy (GWh)	13,626[a]	27,991	46,597	59,729
Fixed assets (10^6 marks)	67,334[a]	94,699	167,626	260,060
Employees per enterprise	89	173	244	476
Employees per production unit	50	69	244	
Fixed assets per enterprise (10^6 marks)	2.86	5.90	14.33	40.1
Consumption of electrical energy per enterprise (GWh)	0.58	1.75	3.98	9.22

SOURCE: Statistisches Jahrbuch der DDR (1953, 1977, 1978).
a Estimate.

payback period. Finally, we can state that under these circumstances production costs are too high and therefore it was necessary to rethink the whole organization problem.

18.3 A STRATEGY FOR ECONOMIC ORGANIZATION

In our discussion of scale strategies we indicated that a strategy for economic organization is related to strategies for production structure, external markets, technical progress, etc. Organization projects must consider these other aspects and create the necessary management facilities so that an overall scale strategy can be followed. For a centrally planned economy with nationally owned enterprises this means the establishment of guiding principles of concentration, centralization, and patterns of organizational association to achieve the desired objectives.

The decisive step in this path is the formation of corporations (*Kombinaten*) responsible for, in most cases, an industrial branch and directly subordinated to an industrial ministry. These corporations are the basic economic units. They consist of a number of legally independent enterprises and are managed by a director-general on the basis of central plan targets and their own forecasts. They are a modern type of industrial organization and the basis of a modern management system for our industry. The key point is that these large corporations can take advantage of the potential economies of scale.

The first 37 corporations were founded in 1968. On the basis of their experiences the next step, strengthening and partial reorganization of the existing corporations and the formation of new corporations, was implemented during the last 3 yr. At present there are 129 corporations producing nearly 90 percent of total industrial production. Each corporation usually has 20,000 to 40,000 employees (the largest has more than 80,000) and has a yearly output of several billion marks.

The formation of the corporations is not just a formal change in the organization structure of our industry nor just a formal union of enterprises into new organizational units. Every corporation has a special set of global objectives and tasks determined by the needs of the national economy as a whole (and thus by the current and future internal and external demand). These objectives and targets are components of the national economic plan. So that the objectives can be achieved with a high degree of efficiency, the corporation was given the necessary capabilities for research and development, project drafting, fabricating and assembling, specialized supplies, and marketing (for example, the corporations control about 90 percent of industrial research and development capability). The corporations are primarily an economic amalgamation of different enterprises that relate to the general tasks. It is clear that the development of such an economic organization is a process of growth, characterized by increasing division of labor and a higher level of specialization and concentration (Koziolek 1979). As compared with the former organization, the corporations are given, within the planned economy, more responsibility in the following fields:

1. Development and realization of new products and technologies
2. Satisfaction of the needs of the national economy
3. Foreign trade

TABLE 18.3 Distribution of mass production and specialized production organization in GDR metal-working industry since 1965 (%).

	Type of production organization								
		Mechanical production					Assemblage		
				Specialized on products					
Type of production	Total	Total	Nonspe-cialized	Total	Flow production mechanized	Flow production automated	Total	Flow production	Flow production automated
1965[a]									
Single production	19.0	10.0	9.0	2.0	0.0	0.0	9.0	0.4	0.0
Serial production	64.0	40.0	21.0	19.0	2.4	0.2	24.0	9.5	0.0
Mass production	17.0	12.0	3.0	8.0	2.1	0.3	5.0	4.1	0.0
Total	100.0	62.0	33.0	29.0	4.5	0.5	38.0	14.0	0.0
1971									
Single production	15.5	8.9	6.3	2.6	0.1	0.0	6.6	0.1	0.0
Serial production	66.8	42.4	19.7	22.7	2.6	0.4	24.4	0.2	0.3
Mass production	17.7	12.1	2.2	9.9	3.2	0.6	5.6	0.9	0.6
Total	100.0	63.4	28.2	35.2	5.9	1.1	36.6	1.2	0.9
1977									
Single production	14.5	8.3	5.9	2.4	0.1	0.0	6.2	0.1	0.0
Serial production	67.2	41.9	19.3	22.6	2.7	0.6	25.3	8.4	0.3
Mass production	18.3	13.0	2.1	10.9	3.4	0.9	5.3	4.5	0.2
Total	100.0	63.2	27.3	35.9	6.2	1.6	36.8	13.0	0.4

SOURCE: Statistisches Jahrbuch der DDR.
[a]Percentage of output; in the other years percentage of production time.

All things considered, the corporations possess greater potential for achieving scale economies. In order to achieve these economies it is necessary to exploit the advantages of the homogeneous and partially centralized management of all enterprises and institutions that constitute the corporation. The basis of the homogeneous management is the combination within the corporation of research and development, production, and marketing into an economic whole. It is well known that if highly specialized production is broken down into many enterprises and production units, management is more difficult, particularly in the field of innovation and technical change. By concentrating related economic activities into a single corporation it will be possible to manage the corporation with a high degree of flexibility and efficiency. However, this usually requires a new managerial structure with what are called principal enterprises (*Leitbetrieb*) charged with organizing joint activities in the fields of research and development, production, sales, etc., within a group of enterprises engaged in similar production. Also it requires new managerial methods, like management by objectives and the application of new management instruments like program planning.

The effectiveness of corporate management is strongly linked to the existence of stable objectives in the main fields of activities. So the corporation has to develop its own forecasts in the areas of research and development, production structure and production size, marketing, and the future development of specialization and concentration. In a centrally planned economy these forecasts must be coordinated with overall national economic strategies. This is done by an iterative coordination procedure between corporation, industrial planning ministries, and the state planning commission.

This example indicates that such important changes in the organizational structure of our industry require changes in all fields and on all levels of management and planning. Now we have to think about new patterns of management and planning at the central level and further development of the management and planning mechanism according to the greater economic potential and increased responsibility of the corporations. Starting from the main objectives and the facilities and responsibilities of the corporations, and using the method of program planning, we can find more appropriate standards for scale and for an optimal relationship between small, medium, and large enterprises.

REFERENCES

Cantley, M.F., and V.N. Glagolev, 1978. "Problems of Scale" The Case for IIASA Research. RM-78-47. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Koziolek, H. 1979. Reproduktion und Nationaleinkommen. Berlin: Verlag Die Wirtschaft. Paragraph 9.3.

Statistisches Jahrbuch der DDR. 1953, 1978, 1979 and other years. Berlin: Staatsverlag der DDR.

Part 5

SUMMARY AND CONCLUSIONS

CHAPTER 19 DIRECTIONS OF FUTURE RESEARCH

J.A. Buzacott and K. Tsuji
International Institute for Applied Systems Analysis,
Laxenburg, Austria

One of the main purpose of the workshop was to identify needs for further reseach on scale. In order to do this, participants were asked to respond to the following question:

> It would be helpful if you could give us a brief statement of what you feel would be appropriate directions for future reseach bearing in mind the special position and character of IIASA and the practical needs of policy makers and decision makers.

Although the question was asked within the contex of research appropriate to IIASA, the suggested research topics are all of general interest. A large proportion of participants filled in the questionnaire and some also made specific suggestions during the last session of the workshop. A further indication of research needs is the recurrence of certain themes in the discussion, each implying a general concern by participants with certain issues where further resolution seemed necessary.These responses and discussion form the basis of this chapter.

However, before discussing the specific research directions, there were some general concerns raised to the phrase *bearing in mind the special position and character of IIASA and the practical need of policy makers and decision makers*, as the conference participants also had a number of important points to make on this.

19.1 WHAT SORT OF RESEARCH ON SCALE IS APPOPRIATE TO IIASA

In his closing remarks to the conference, Tomlinson made the following points:

Tomlinson: Let me just remind you of certain major factors relating to our own particular situation here at IIASA and which affect what we do.

First of all, if we are going to undertake any work it will have to fulfil three conditions.

1. It must be of international interest and East - West. That is a basic criterion, fundamental to the whole role of IIASA -- a problem which is only of interest either to the market economies or to the socialist economies is not something we should tackle here.

2. Our work should and must be interdisciplinary – many or all of the disciplinary studies that one might undertake in this field can be done somewhere else besides IIASA. These two conditions, the interdisciplinary and the international nature of our work, give a special character to what we can do. We hope that it is a peculialy useful characteristic, but in any case, the conditions have to be borne in mind.

3. The third distinctive feature is what I've called coordinating. By the nature of the work we do and the relatively short periods that people stay with us, probably our most fruitful work in these applied areas is to identify common problems, and in some way coordinate the state of knowledge and the state of research with regard to those problems. It is unlikely that we shall be able to produce an entirely new idea and carry it through to its research conclusion within the length of time that an individual stays with us. We can, however, achieve a great deal by inviting good research workers to help identify the state of the art as well as continue their own lines of research and then to coordinate the knowledge, and sometimes perhaps even the research, of people who are interested in the same problems. I think this is an important feature.

Some other items also have to be borne in mind. We are expected to direct our efforts towards real problems, not research for its own sake. This means that what we do is directed primarily towards analysts and advisors, since we cannot undertake a direct consultancy role. Staff turnover provides us with a problem beause we don't keep people for very long periods of time; on the other hand, it does provide us with great opportunities. We can say that in order to tackle that problem we need a certain kind of person and there will be a vacancy for that kind of person because of the natural turnover.

Concerning research on scale, and the role of IIASA in such research, he had this to say.

Tomlinson: . . .one of the questions that came up early in our own work, and which I have been asking myself during the workshop, was: "Should we really be putting a focus on scale, or is it really just a rather small element in other problems?" My own reaction to this was that there is a need to put this focus on scale. There are a lot of people in the world working on innovation, on technology assessment, on management of organization. If we concentrate work on scale within a team here at IIASA we can ensure that there is an interaction of the various approaches and interests that can be brought to it. After all, we must not lose sight of the fact that the critical decision that has to be made by industries and governments is often "How big shall we actually build or allow this to be built?" We should concentrate our analysis on the real problem. By putting a focus on scale we can bring all the conflicting factors together. It is, in fact, the problem of identifying just what factors are important, combined with, in some cases, an inadequate understanding of their impacts and consequences which has led to mistakes on scale that have been and are still being made. So my conclusion from this discussion is that it is indeed worth our while to put a focus on the problems of scale.

The sheer existence of a team coordinating information and thinking and doing research is in itself a useful function.

In the subsequent discussion the following general issues were raised: (a) who will benefit from scale research and, in particular, why should industry be involved with it? and (b) how are problems of scale a problem for systems analysis?

19.1.1 Why Industry Should Get Involved in Research on Scale

Shutler: The justification I would like to see for industry to involve itself in these studies is that, whether one is talking about socialist systems or market economies, decisions about the scale of a new plant and decisions about mergers and takeovers are now regulated. The decisions have to be taken in a community context. Therefore, it is in the interest of industry that the people doing the regulating should understand the problems that industry faces and should understand the decision process through which industries go in choosing an optimum scale of operations at plant level or total industrial organizational level. Therefore, there is a real justification for carrying on this research provided only that one can involve individual companies across national boundaries, perhaps coordinated by steering groups of people in the individual countries linked via IIASA.

Savin: I think that we would agree from the industry side that there is not the need for study of industry problems for the sake of industry but it could be that from our point of view that the understanding of the problems of industry should be better known amongst society in general, goverments, and supernational organizations and the like. To that extent I think IIASA might have a role to fulfil – not to do the work but to draw attention to the implication of it.

van Dalen: . . . for whom is IIASA's research intended – is it intended for industry or is it intended for other bodies who have some control or some bearing on the way our society develops, for example, trade unions or governmental institutions who have to deal with economic parameters and who want to have at their disposal some general ideas and formulate how to respond to economic changes within private industry. I get the feeling that IIASA is mostly directing its efforts to these institutions rather than to industrial companies themselves.

Tomlinson: To whom is IIASA addressing itself? This clearly is a question we are always asking ourselves. I would say that most of the work of IIASA is addressing itself to what are in general rather large broad policy issues, trying to develop understanding which would be used at a governmental level, but we do not feel that this can or should be the sole purpose of IIASA. If we are genuinely an institute of applied systems analysis we have to take account of the fact that rather a large proportion of difficult decisions are taken within industrial organizations or in the negotiations that go on between productive organizations and government.

I would be very unhappy if we were not addressing ourselves seriously to problems that worry people within industry. We won't really be successful in this until the suggestions are coming as strongly from those organizations as they are from IIASA, although at this stage I think we have to put ideas forward. So I believe we do have two audiences, but we are only just beginning to talk to one of them. One of the features of this conference is the fact that it has had strong industry representation – that industry has talked and engaged in debate. It needs to happen more often.

19.1.2 Problems of Scale as a Problem for Systems Analysis

H. Wagner: IIASA is at a period now where the outside world is hungry to see that the results of systems analysis contibute value in thinking about issues. In this particular realm, systems analysis is going to contribute in either or both of two ways; one is the use of systems analysis by an industry or government to think through these scale/size issues, and the other possible contribution is that IIASA itself will use systems analysis to give some insights about the issues. Either one will be a viable way to proceed.

I think IIASA needs to have a very clear view early on as to what it intends to get out of any effort.

Then in his comments written on the last day of the workshop, Wagner wrote:

IIASA has a splendid opportunity to make a major contribution to an international understanding of industry size/scale issues. Its chance for success, however, appears modest if IIASA staff yields immediately to the temptation of taking any of a number of obvious next steps that are suggested by the work reported at this conference. I would counsel a more deliberative research strategy.

Specifically, IIASA staff should devote the next 12 months to clarifying in the context of systems analysis what are the key issues in theis area. The conference proved beyond doubt that even the working vocabulary of this area is muddled. Furthermore, which investigatory appoaches can stand the tests of science is itself a central issue. Thus, an assessment of what are meaningful issues, which ones are most important to research, and which ones can be studied effectively by IIASA cannot be sensibly made today. The multinational structure of IIASA, with its organizational resources to obtain entry into universities, research centers, public agencies, and industrial concerns, creates a singular opportunity to discover what are the essential research challenges and thereby to make an important impact on the future thinking in this area.

The target output of the initial 12 months' work would be a publication that proposes a comprehensive blueprint (or taxonomy) for research by systems analysts in this field; specifically, it woult state in detail the issues and hypotheses that need further attention and testing. For several of these topics, it could propose one or more legitimate avenues of study, describe the character (nature) of the possible results, and give an idea of the criteria that are appropriate to judge the scientific merit of the findings (validity and generalization).

The report, by way of background, should briefly discuss the relative merits of previous research, and point out any pitfalls to research methodology that were uncovered in this literature survey.

The concluding chapter would propose what research IIASA intends to pursue, the rationale for the choice, along with a plan and timetable.

The report, prior to its publication, should be circulated, in draft form, to knowledgeable individuals outside IIASA. These reviews can provide suggestions for reseach design improvements, alert IIASA staff to latent criticism, and possibly indicate where it would help to call upon scholars from disciplines not represented among MMT's own staff.

Unlike other substantive areas studied at IIASA, there is less urgency to address issues immediately. The truly pervasive and extensive aspects of these issues as well as the disturbingly limited impact made by systems analysis so far suggest that a careful "front-end" study would be highly valuable.

19.2 SPECIFIC DIRECTIONS FOR FURTHER RESEARCH

The main areas proposed for possible future study can be grouped under six main headings and we shall comment on each. They are:

1. Problems of taxonomy – definition and measurement
2. Organizational scale: size, structure, and environment
3. Scale and technology change (innovation)
4. Scale in industrial complex and associations
5. Implications of scale on national economic and industrial policies
6. Scale and society: issues for the public

Furthermore, the discussion at the workshop indicated a concern with two further topics:

7. Scale and learning
8. Scale and the uncertain future

19.2.1 Taxonomy – Definition and Measurement

Rochlin: I am most interested in developing descriptive "indications" of scale at all levels together with environmental interaction. Even if small is beautiful, what is small? At what points do quantitative shifts in scale entail qualitative shifts, for example, in the locus of problem area from one part of the system to another or from one level to another? These are systemic issues and therefore most appropriate to IIASA. However, this does not necessarily imply they are "modelable" in the mathematical sense.

Uhlmann: . . . Instead of proving again and again that there are diseconomies of size we should develop an adequate taxonomy of size. This taxonomy should take into account the level, the sector, the size establishing factors. We should overcome generalization like "small is beautiful." There are aspects of size which are beautiful and those which are not. To work on such a catalogue of indicators is a tyical task for systems analysis.

Comment. Various ways of categorizing the situations within which scale problems can be analyzed in comparable terms have been proposed in this book – levels, factors, etc. A taxonomy of problems of scale at each level and the relevant factors and criteria seems to be required before one can develop a more general methodology for determining scale.

19.2.2 Organizational Scale: Size, Structure, and Environment

Ansoff: I would suggest splitting the overall project into several related lines of enquiry.

1. . . . Problems of economies of scale in production
2. Design of the total logistic (productive) process of the enterprise. This would include tradeoffs between research and development, production, distribution, and marketing. The question to be studied is the effectiveness of the total process as a function of technology, size, location, interfunctional coupling, etc.
3. Design of the total enterprise, comprising both managerial and productive processes. Here research could be broken down between strategically stable and unstable environments. The latter is the key problem in the West today and is of great interest in the USSR.

Plug: . . . a main bottleneck in the size of organization is managerial in character, or rather managerial and organizational . . . what is required is a taxonomy of organizations and a theory of organizational and managerial functioning, differentiated enough to deal with different situations.

van Dalen: It is in the area of investigation of the industrial enterprise that shows promise for further research because the problems of organization and management structure lend themselves somewhat better to generalizations and abstraction.

Apter: Particular topics needing research, review, or reformulation should be given special attention and general hypotheses formulated:

1. Bureaucratization/coordination; cost, management, information
2. Social overhead/participation; human priorities, organizational implications
3. Political systems/controls; hierarchy, interferences, re-allocation, priority setting
4. Social structures/class; meritocracy, compensatory education, training
5. Adaptation/social learning; innovation, technology, capital vs. labor intensive

Stoyanov: Development of:

1. System of indicators describing the scale of an economic organization
2. Methodology for determining the scale of an organization
3. Methods and techniques for determining and controlling the scale of the economic organization.

Comment. There seem to be two related problems: (1) What is the optimal size of an organization? and (2) What is the optimal size and structure of the management of organization in a particular environment? Developing solutions to these problems requires an understanding of management as a human and social process and management as an information and control process. Thus the problem of organization scale requires an interdisciplinary team with experts in both organizational behavior and in information and control systems. Such a project is probably too broad in scope for IIASA.

However, since the structure of management will be influenced by the technology of management – for example, the use of computer networks and distributed data bases for communication, control, and information processing – an appropriate task for IIASA could be to study actual and planned networks in industry, both East and West, and develop models by which the performance of these networks can be evaluated and their influence on managerial structure and effectiveness assessed.

19.2.3 Scale and Technology Change (or Innovation)

Haustein and Wittich: In our opinion, the main problem for research about scale is long-range social planning of economic organizations and especially its interconnection with technology policy (or innovation policy).

Rees: The relationship between scale and the diffusion of innovations: are small firms more innovative than large firms in different countries?

Schenk: It seems very important to analyze the problem of scale explicitly with respect to foreseeable changes in technology (e.g., the impact of microelectronics or information technology).

Comment. The discussion on scale and innovation in Chapter 15 advanced a number of hypotheses about the relationship between scale and the type of innovation. Major product innovations occur in relatively small organizational unitis. For major process innovations, closely integrated medium-size organizations seem to be most effective, while large organizations with clearly set productivity improvement goals seem to be particularly good at minor process improvements and innovations.

However, it would seem that an understanding of the scale issues in innovation should come out of a more detailed understanding of the innovation process. It seems necessary to study the nature of the communication networks involved in recognizing and developing an innovation. Are such networks self-generated or do they need to be created and motivated from outside? What resources does the innovative group require and how are they acquired? A better understanding of the innovation process should lead to a clarification of its relationship to the scale of the organization in which it occurs.

However, there is one problem that may warrant study on its own – the way in which successive scale-up of plant proceeds. An understanding of what determines the maximum feasible plant size at a given time seems necessary in order to understand the decisions firms make on size of plant.

19.2.4 Scale in Industrial Complexes and Associations

Comment. A problem of paricular concern to the socialist countries following the reorganization of their industry in the early 1970s (but also of significance to large companies in capitalist counties) is the appropriate structure for large multiplant, multiproduct organizations (see Chapter 11 (Egiazarian and Glagolev)). These organizations have to be efficient both in producing existing products and in developing and marketing new products. It was pointed out by Utterback that it is difficult to combine both aspects in one organization. An organization structure and control system that promotes efficient production by specialization and concentration is usually inappropriate for recognizing and exploiting new product innovations, particlarly when the new product may threaten the market of existing products.

The solution to this problem appears to be quite complex. On the one hand it seems to be desirable to develop production systems that are more flexible and require less specialization to achieve maximum efficiency, perhaps by exploiting variety and diversity; on the other hand, it seems necessary to develop a better understanding of both the problems of management scale and the relation between innovation and scale before significant progress in improving corporation structure can be made.

However, it does not seem possible to develop solutions to this problem until some of the other scale problems are better understood -- in particular, management scale, innovation and scale, and effect of a turbulent environment and uncertain future on the planning process.

The only direction that research on this topic could take in the immediate future would be to do a comparative East West study of large organizations in a particular industry. The objective would be to compare their productive, innovative, and marketing efficiency and to relate differences to organization structure. IIASA is an appropriate place for this type of research.

19.2.5 Implications of Scale on national Economic and Industrial Policies

Daly: Market size and options for small countries. It raises interdisciplinary questions and has some connections with scale, international trade, and living standards and can affect corporate decisions. It is related to government policy decisions.

Stratton: The problems faced by small but developed economies in competing internationally, particularly if they have to import energy.

Horsnell: The appropriate size of industrial enterprise in the context of a low or no-growth or even negative growth economy with abundant supplies of labor and short supply of fossil fuel. This would currently be appropriate to Third World countries and will, in my view, become increasingly appropriate for developed countries.

Millendorfer: In the metal products branch of industry we have big firms and small firms. They produce different things and the question is, can we find abstract principles which govern the difference in firm size and products? Because, for example, of the bottleneck in energy, we will have in 10 or 20 years a very different mix of products and a very different demand structure. This has to lead to a different structure of size and scale of industry and we should try and develop an understanding of the process of transition.

Comments. The discussion on innovation in Chapter 15 also raised the question of what policies countries can adopt when they find that they have too much human and capital resources tied up in large-size plants in declining industries. While plant closures may be efficient, the social and political consequences are severe. It seems that an understanding of the way firms make scale and location decisions is a necessary basis for government policies aimed at modifying industrial structure.

Another aspect of industrial strategy of concern to a number of countries is the development of policy to improve the effectiveness of small and medium-size firms. Such policies may include support for new product or process development, production and marketing management consulting, and low-interest loans. In order to evaluate the benefits of such policies there should be an understanding of the benefits of small firms and the dynamics of their development and growth.

Since the workshop, IIASA has decided to look at questions of industrial strategy within a broader context than just that of scale. However, the specific problem of policies for support of productivity improvement and innovation in small firms should not be neglected. It seems that there are significant differences between socialist and capitalist countries (and within each of these groups) in the role and importance of the small firm sector to the economy. IIASA is in a unique position to permit interchange of information and ideas between East and West.

19.2.6 Scale and Society: Issues for the Public

Savin: Some studies on the effects of scale on society: the changes in the work ethic and the like, could be very interesting – as well perhaps as looking at the sometimes alleged situation that economies of scale sometimes go along with restriction in variety to the consumers.

Stratton: One of the things that concerns me is the relationship of what we have been talking about, increasing size and scale, to limits to growth. A few years ago all the debate was on limits to growth and here we are discussing growth as a good thing in itself, we think . . . We always tend to start off with economic growth, per se, but are we starting at the right point? Are we looking at the system as a whole? How does this relate to quality of life? If you were to start from quality of life you would perhaps be setting yourself very different targets for scale and size. So, are we basing too much on what we have done in the past and are we assuming that the whole economic and social structure is going to remain too much the same in the future?

Comments. The question of the impact of scale on the public and their attitudes to large-scale plants, organizations, and nations is becoming of increasing concern. It merits significant research; however, such research is probably not appropriate for IIASA because the findings may be politically sensitive. Indeed, IIASA's financial support is dominated by two large countries, which may result in symptoms within it of this distrust or apprehension concerning bigness.

19.2.7 Learning

This topic was discussed at great length (see Chapter 6). For example, it was suggested that the increase of the size of the largest plant with time or with cumulative production indicates that a learning process is occurring. However, there was disagreement as to whether this process can be described by a general law, that is, whether the process is inevitable. Some participants felt that this represents too great a generalization and that in order to understand the significance of learning it is necessary to determine exactly who learned what.

The literature and the workshop discussion support the idea that the appropriate way to think about learning is within a hierarchical context (Figures 6.2 and 6.3, Chapter 6). Learning occurs at a variety of levels within the firm. It is not automatic. Improvement is the result of decisions and actions taken by people at each level. However, it seems reasonable that, in making decisions at a particular level, learning can be assumed to occur at lower levels.

All the same, it is necessary to understand the learning process at each level. There are too many examples of firms that wrongly assumed that learning is automatic. There are limits to the learning curve and these limits can only be understood if the learning process is understood.

The learning effect can confuse evalutation of the benefits of increasing scale. An observed decrease in cost with successive increases in plant may not be necessarily due to real economies of scale. The results of experience in designing and operating one plant could be incorporated in the next plant. If economies of scale are overestimated, plants may be built too large and unforeseen problems in extrapolating the technology may occur.

Thus, it is essential to understand the learning process and its limits. It is also necessary to understand how one exploits the learning process to ensure that past mistakes are not repeated and all available knowledge about the technology and the market is used. This is particularly important when plant is being scaled up.

19.2.8 The Effect of Uncertainty about the Future

One of the major issues discussed was the way uncertainty about future demands, costs, prices, and technology will affect scale decisions. For example, compared with the time when decisions on the size of plant and equipment in the electrical generating industry were made, recent demand growth has been less, the performance of the generating units has been worse, and the cost of fuels has been higher. It was pointed out that there is a tendency in making scale decisions to assume that the factors influencing the future will be the same as those that operated in the past. Generally, the effect of this inability to make scale decisions that recognize the effects of uncertainty about the future are serious. However, it was pointed out that in the UK a lower than expected growth in the demand for electricity permitted inefficient out-of-date plant to be scrapped so that the cost penalty was negligible.

As pointed out in Chapter 7, while there are a number of mathematical models for determining the optimal size and timing of plant installations, almost all these models assume that the future is known. The few models that allow for uncertainty make very simple assumptions; either there is no correlation between demand growth in successive periods or constant demand growth per period, but with a value not known precisely at the time the scale decision is made.

Thus, there is a need for a better understanding of how to make scale decisions in the face of an uncertain future. This requires better methods for defining the extent of our knowledge about the future and the nature of our uncertainty, better models for allowing for the uncertainties, and approaches for making scale decisions that will not be affected if the future turns out differently from our expectations.

The objectives of research should be to:

1. Determine the effect of uncertainty about future demand, technology, costs, and prices on the optimal scale of a plant or a system
2. Suggest methods of improving the planning process so that both the amount and the effect of uncertainty can be reduced
3. Identify approaches found useful in one industry and which could be transferred to other industries and other countries
4. Provide the basis for more formal models of decision making about scale

It is also likely that the research would lead to an understanding of the way the structure of an industry is affected if the level of uncertainty increases, for example, owing to inflation or changes in the market or in raw material availability.

CHAPTER 20 CONCLUDING REMARKS

R. Tomlinson, J.A. Buzacott, and K. Tsuji
International Institute for Applied Systems Analysis,
Laxenburg, Austria

20.1 INTRODUCTION

The main purpose of this chapter is to discuss how far we are able to go on the basis of the workshop and of subsequent review and analysis in answering the two important questions that were set out at the beginning of the first chapter. They may be rephrased:

1. To what extent is scale a general problem, and who would be helped by a more detailed analysis?
2. How far can we go towards developing a general methodology for determining scale?

Before we discuss these points in detail, however, there are certain fundametal issues that need clarification and reiteration, since they can easily be ignored or forgotten with resultant confusion. These issues are:

1. Size vs. scale
2. The dynamics of scale
3. Broader criteria determining scale

20.1.1 Size vs. Scale

Too often the words *size* and *scale* are used synonomously; and indeed, *scaling-up* is often thought of as a process of simple magnification. Yet the process of making larger is often more complex than this, involving structural changes as well. (Indeed, without structural change the economies of scale can often not be realized.) We therefore follow Gold in making a distinction between size and scale and take *scale* to refer to size *and* structure. The point is worth clarification through an example.

The basic difference is outlined in Figure 20.1. Consider a production process producing a single product. One way of increasing the output is to increase its size, that is, to make it bigger by increasing some of its dimensions, using larger motors, using higher speeds,

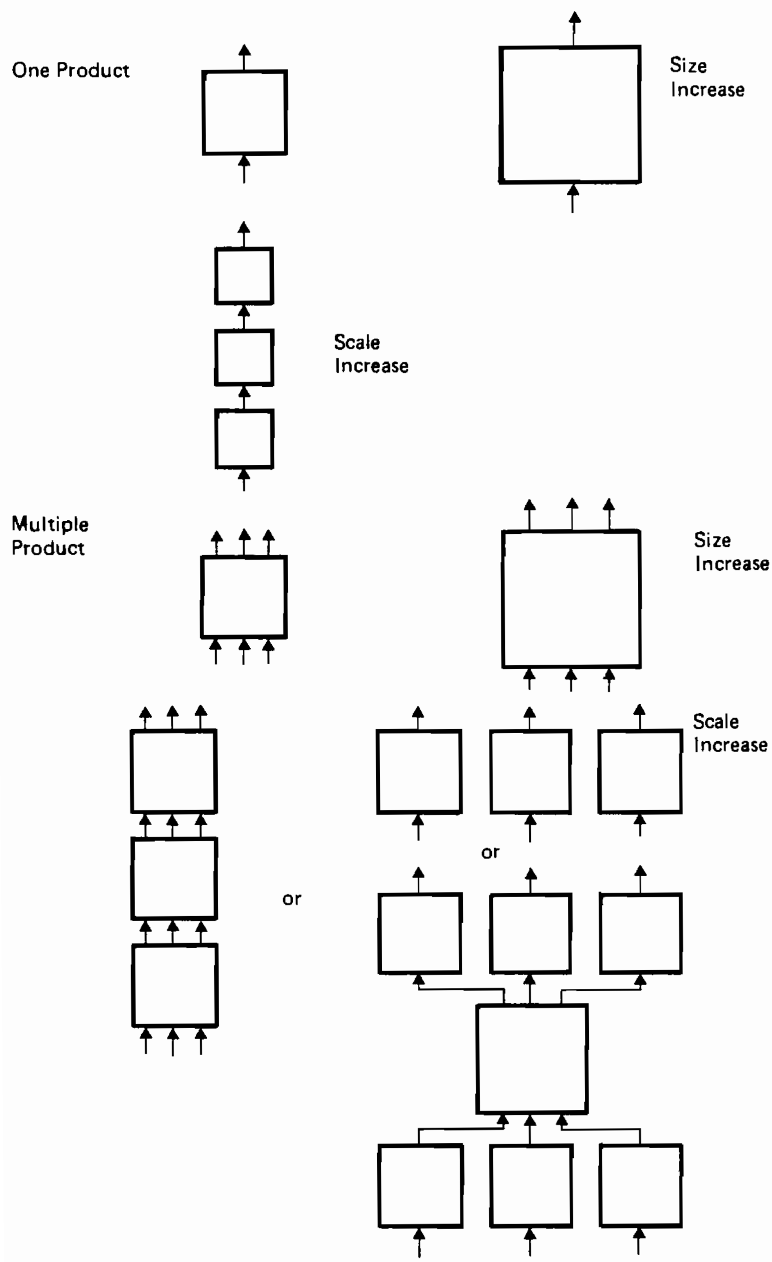


FIGURE 20.1 Size versus scale.

etc. However, rather than increasing size it is often more expedient to subdivide the functions of the process and develop specialized equipment for each function or stage.

Next, consider a process that produces a number of products. Again, one method of increasing the output is to increase the size of the process. However, in practice two steps would be taken; one is to specialize individual stages of the process, and the other is to specialize the process to individual products to eliminate change over time and interference. As a further step, more complex structures would develop in which some stages may be common to many products and others specialized. Indeed, it is conceivable that the process may become a set of specialized functions connected by some complex network. This is the basis of Gold's definition of scale: "the level of planned production capacity which has determined the extent to which specialization has been applied to the subdivision of the component tasks and facilities of a unified operation." That is, scale encompasses not only changes in size but also changes in the structure of the process or system.

20.1.2 The Dynamics of Scale

Many discussions of scale take place as if the decision were a static one; that one could, so to speak, feed in the current values of the various parameters and identify the optimal scale using standard optimization techniques. Such a procedure would be fallacious in a number of respects. In the first place, a particular scale decision is part of a time sequence of decisions within the organization concerned; it is affected and constrained by previous decisions and the resulting commitments to physical and organizational structures. Past problems and future uncertainties are as important to the analysis as is present knowledge. The dynamic behavior of the various parameters is also important, the accumulation of experience, market patterns, environment, changes in technology, etc. In fact, the more that one studies a real problem, the more these dynamics seem to gather in importance, and the more one becomes aware that an optimum based on static information may be local in time, and extremely dangerous. So it becomes increasingly important to see the question of scale in its full dynamics based on past experience and inserted into an uncertain, though seldom quite unpredictable, future.

20.1.3 Broader Criteria Determining Scale

One of the difficulties in understanding and making scale decisions is that the usual concern with achieving economies of scale through increasing size is not, on its own, a sufficient criterion for determining the appropriate scale. There are other criteria that have hitherto often been thought of as secondary, but which can become dominant. These other criteria include flexibility, systems effectiveness, complexity, and human motivation and performance. We shall give a short discussion of each.

20.1.3.1 Flexibility

When one operates on a small scale, particularly in a situation where a large number of similar units are concerned, it is relatively easy to react to a new situation, whether it relates to technological advances or change in customer demands. Individual units can be changed or modified at relatively low cost, and the changeover can be made gradually. When scale is increased, that may no longer be so. The whole equipment must be changed and the ramifications of that change are harder to identify. It is becoming more the rule that increased scale goes with increased specialization. Specialization also leads to lack of flexibiity, as is illustrated by the supertanker. When trade in one commodity has become poor, the supertanker has often been found to be too large for the facilitites for handling other commodities. Similarly, a coal mine may have its large, superefficient surface facilities geared to rail transport and cannot serve a customer equipped to receive the product by road transport. Yet with changing markets this may be undesirable. The wrong balance of specialization and flexibility can easily lead to economic disbenefits.

20.1.3.2 Systems Effectiveness

Although it is sometimes possible to look at a unit in complete isolation, more often than not the unit has to be seen as part of a system and its effectiveness judged by its contribution to the overall performance of the system. If the performance of that system is seriously weakened by the inclusion of units of disproportionate size, trouble can easily begin. Thus, in examining the desirable size of an individual generator, it is necessary to look at the impact on the whole of the electric power system. There have in fact always been rule-of-thumb methods for doing this, but they may not necessarily remain valid when both unit and system increase in size.

Another obvious application where the systems effect is of critical importance is in air transport – the subject of Chapter 3. Here the danger of running an aircraft at too low a percentage of capacity is a major factor in the actual operating costs.

Another major factor for consideration is the question of vulnerability and security, since systems performance is much more sensitive to the breakdown of large units than small ones. This problem becomes more severe the more closely the system is controlled and integrated, which is a natural concomitant of increased size. It is also worth remarking that it is in connection with systems behavior that extrapolation from past performance is most dangerous.

20.1.3.3 Complexity

In general, it is true to say that large systems have more elements within them, and are thus more complex systems requiring more care in coordination and control. This complexity can be expensive. Thus, in the electricity generation case, the large units now being installed are so complex that the installation time has escalated, and the reliability significantly reduced. The associated costs have been sufficient in some instances to outweigh the expected economies of scale. The problems associated with "trouble-shooting" also increase out of proportion to the increase in size.

The issue of complexity is not easily analyzed, and is too little researched. Some complexity can be reduced by the modular approach; some is inherent in the situation. It cannot be ignored.

20.1.3.4 Human Motivation and Performance

The human factor in scale decisions can be quite crucial and is slowly beginning to achieve the recognition it deserves. The human dislike of being caught inside the big bureaucratic machine is well known, though its consequences are not easily identified. It is often assumed that the answer lies in better control – though often it lies in more appropriate organization – e.g., one which reduces the perceived size of the working group. The same effect may not be apparent at all levels in the organization – senior managers may prefer large units because they have fewer subordinates to control, but the managers at unit level will face a more difficult task, leading to problems quite different from those to which they were accustomed.

20.2 TO WHAT EXTENT IS SCALE A GENERAL PROBLEM, AND FOR WHOM?

This topic is best discussed in relation to the requirements of potential decision makers with problems to solve, and who might be interested in obtaining better information or analytical techniques concerning problems of scale. We might approach this from a consideration of the five organizational levels already discussed, but it will be sufficient for our purposes to consider just three levels of decision makers, roughly corresponding to the topics discussed in Parts Two, Three, and Four of this book: the plant designer, the general manager, and the national policy maker. For convenience, we consider the decision maker and the systems adviser together as one team.

The first potential user, then, is the designer of a productive unit, or his analytical advisers. His job is to provide the very best design to meet specified conditions. Thus, whilst most of the secondary conditions set out above will be considered in his planning, many of them will be dealt with in a rather perfunctory sense by assuming certain parameter values, such as the acceptance of a certain growth rate in demand. Apart from the original elements in the design process, which already cannot be prescribed, most levels of management will almost invariably insist that a normative approach be adopted in the planning process, i.e., that a certain methodology be adopted, a certain set of calculations undertaken, and certain organizational norms accepted. It is critical to an understanding of what happens to recognize that there is a division of responsibility here, that certain pieces of analytical research are the responsibility of this level, and others should be undertaken at the higher level. It is also important to realize that the process adopted will determine the outcome. Decisions and simplifications must always be made at various stages in the planning process. Thus, in coal mining, one cannot leave all possibilities open until a final optimizing calculation is undertaken. The overall planning problem has to be structured and decisions made at an early stage based on imperfect information. It is often not even permissible to work on three different alternatives; in practice this would require three different planning teams leading to impossible organizational complications. Yet the plan(s) prepared must be resilient in the face of a range of possible futures. Thus, a methodology is needed, and at present every technology – and indeed different organizations within a technology – work out their own. If a general methodology could be developed, it would surely be useful.

In the second place, we need to consider the decision makers at what one might describe as the general management level. They have to evaluate proposals put to them (whose technical reliability they are unable to check in detail) and they must have faith in the procedures they have laid down. But the proposals have to be considered in the wider content of the uncertain external environment and general systems effects. They have the problem of ensuring that the proposals set out for them meet these wider requirements; and they must therefore ensure that they have the necessary methodology to include these other factors (including organizational factors) in their final judgment.

Thirdly, we need to be concerned with policymakers at the national level. They may not make individual decisions but through policy statements, regulation, and a variety of other controls and policy rulings, ensure that the decision of the general management will satisfy general criteria relating to national and perhaps international objectives. They too need to develop a methodology for this.

Apart from managers, decision makers, and their immediate advisers, we should also consider the needs of two groups of researchers: those who want to develop modeling techniques appropriate to particular situations, and those who want to develop improved understanding of what Simmonds calls "industrial behavior patterns." Those concerned with industrial behavior patterns require general approaches to understanding problems of scale and the way in which scale decisions are affected by the relationships between cost, performance, size, and structure. Eventually, the results of such research will have an impact on both policymakers and planners in contributing to a greater understanding of the significant factors. However, the research requirement, which should come before the application, must not be ignored if real progress is to be made.

20.3 A GENERAL METHODOLOGY: IS IT FEASIBLE?

The discussion at the workshop made it clear that problems of size and scale were widespread and not fully researched. It was not clear, however, just how far they could be considered general, in the sense that it would be possible to develop a generalized normative approach to a definable set of problems. Such an approach would require the development of a standard structure for scale problems with a statement of the technique available for analyzing the different structural elements. At first sight, it might have been assumed that such an approach would be most valuable at the level of the productive unit, for it is at that level where the mistakes are best documented. Nevertheless, it was felt that in some technologies the relevant factors at the level of the production unit were well known, and the available analytical technology adequately documented. Failures, if there were failures, came either from an unwillingness to recognize the symptoms and to undertake the appropriate analysis, or simply through the insertion of incorrect parameter values in a calculation that was structurally correct. This was often a result of the unthinking extrapolation of past experience, and emphasized the need to maintain continual vigilance in situations of rapid change, both technological and economic. But in other technologies the situation was much more open.

Whilst there was a degree of acceptance of the present position with regard to the scale of productive units in some technologies, there was no complacency at all when it came to questions relating to higher levels of organization and management. Questions of the scale of industrial complexes, the relationship of organizations to their economic environment, and the structure of multilevel complexes were felt, *inter alia*, to be ill-understood and needing further research. But the question of generality remains unresolved.

It is of course not obvious that a general methodology should be based on management levels, as the previous discussion would imply. There is a closer relationship and interpenetration of the problems at different levels than the structure of this book would suggest. In an industry such as the ethylene industry, the production technology and its scale characteristics are of dominant importance not only at the level of the plant but also at the level of national economic policies on trade, tariffs, investment subsidies, monopolies, and competition. It would seem that while level is a useful way of categorizing those decisions where scale is an important aspect, it may not be the appropriate way of categorizing the methodology or critical factors in understanding scale decisions and their impact.

A second categorization, from which we might develop a general set of models, is that of scale problems in which (a) the production technology is dominant, (b) the structure or relationships of specialized components are dominant, or (c) the psychological aspects of group or individual behavior are dominant. However, some scale problems require consideration of more than one of these three aspects. A subsidiary categorization of models and approaches focuses on the environment as either static, dynamic (but changing in a defined way), or involving elements of uncertainty and risk. We comment briefly on the three types of problems.

- *Production technology dominant*. The characteristics of the production technology essentially determined most of the size-related capital and operating costs. Formal mathematical models are quite well developed for both static and dynamic situations although, as pointed out in Chapter 7, the treatment of uncertainty is not yet satisfactory. Also, while the dynamic models are quite comprehensive, they are not yet well known.
- *Structure and relationships dominant*. The way in which functions are assigned to basic elements and the connections and interactions between these elements determines performance. Available models are mostly restricted to static situations and relatively simple networks. However, some of the modeling approaches developed to understand computing systems and communication networks may be extended to provide more general insights.
- *Psychological aspects of group behavior dominant*. Performance is determined by interaction between the members of the group, role assignment, leadership, goal setting, etc. There is a considerable descriptive and theoretical literature on group behavior but not much in the way of prescriptive guidelines. The sociotechnical systems approach tries to combine this with a consideration of the properties of the production technology and this work may have relevance to the study of problems of scale.

Alternatively, we might attempt a general methodology in terms of the broader criteria set out earlier in this chapter (Section 2). It should in theory be possible to write a series of authoritative statements on complexity and scale, flexibility and scale, scale within a systems environment, scale and human factors, etc., which could discuss the techniques currently available for exploring these questions in a given case. On the other hand, it is not yet clear that adequate research, particularly into the taxonomy of the subject, has been undertaken for such a venture.

20.4 CONCLUSION

This volume is the conclusion of 2 years of work, as well as a report on a workshop. What have we learned, and where do we go from here?

We have learned that there is a major problem that, in one form or another, recurrently, faces management at all levels of organization. We have learned that there is a large literature on the subject, but that this literature is for the most part within disciplines, and not well known to people from other disciplines. There is a need to develop a cross-disciplinary community and undertake more interdisciplinary work.

Although at the start of the work it was hoped that it would be possible to produce a general handbook setting out a methodology for tackling scale problems, this does not appear to be possible at our present stage of knowledge. There is need for more research of a taxonomical nature, on questions of definition and measurement. Only then are we going to be in a position to prepare a handbook, which remains a management need. Currently, the scientific study of the subject is going down the traditional reductionist path – identifying analytical problems and trying to solve them separately. To be of use to decision makers, it is just as important to study how to put the pieces together again. No one is doing that.

A number of major problem areas have been identified, which would each justify a conference and a book in themselves. The scale of institutions in relation to the surrounding economy has already been studied with some care in some countries, but not widely enough. Scale, management, and organizational size and structure is a key issue where the work of sociologists and organization theorists needs to be integrated constructively with that of management scientists and others. The question of industrial complexes and multiorganizations is a major one facing all parts of the developed world and has received all too little attention outside the socialist countries. Scale and technological change remains a critical element in thinking about industrial growth. Finally, an increasing number of public issues related to the question of scale are beginning to emerge in debate. In all these the work needs to be done on a full interdisciplinary basis, and it needs to be done now.

So far as IIASA is concerned, we have gone a long way in helping to structure the problem, to clarify certain critical issues, and to establish a new community. Although the issue of scale will continue to be of interest in IIASA research, e.g., in work on IIASA's Innovation Task, it will no longer constitute a separate task (a problem of budget scale). We hope that our stimulus will enable other researchers to carry on the work.

APPENDIXES

APPENDIX A

CONFERENCE PARTICIPANTS

Prof. H. Igor Ansoff
European Institute for Advanced Studies in Management
Place Stephanie 20, Bte 15–16
B–1050 Brussels, Belgium

Dr. David Apter
Dept. of Political Science
Yale University
New Haven, Connecticut 06520, USA

Dr. A.M. Belolipetsky
Institute for Systems Studies
Moscow, USSR

Mr. Gordon G. Betts
Manager, Technical Development
BP Chemicals Ltd.
Belgrave House, 76 Buckingham Palace Road
London SW1W 0SU, UK

Prof. John Buzacott
Dept. of Industrial Engineering
University of Toronto
Toronto, Canada

Prof. Donald J. Daly
Faculty of Administrative Studies
York University
4700 Keele Street, Downsview
Ontario M3J 2R6, Canada

Dr.-Ing. Johannes Dathe
Managing Director
Industrieanlagen-Betriebsgesellschaft mbH
Einsteinstrasse 20
8012 Ottobrunn, FRG

Dr. Frederic de Hoffmann
President, The Salk Institute
P.O. Box 1809
San Diego, California 92112, USA

Dr. J.F. den Hertog
N.V. Philips Gloeilampenfabrieken
Eindhoven, the Netherlands

Mr. H.H.J.M. Derkx
Estel NV
Postbus 401
6500 AK Nijmegen, the Netherlands

Dr. Manfred Dirrheimer, Jr.
International Institute of Management
Platz der Luftbrucke 1–3
1000 Berlin (West) 42

Mr. Gillis Een
R&D Group Staff, Strategic Process Planning
Alfa-Laval AB
Postfack S–147 00 Tumba, Sweden

Prof. Gevork A. Egiazarian
Faculty of Economics
Moscow University
Moscow 117234, USSR

Dr. John C. Fisher
Corporate R&D
General Electric Company
Bldg. K-1, Room 3C36
P.O. Box 8
Schenectady, N.Y. 12301, USA

Prof. Bela Gold
Director, Research Program in Industrial Economics
Case Western Reserve University
Cleveland, Ohio 44106, USA

Prof. Sten-Olof Gustavsson
Division of Industrial Management
Chalmers University of Technology
S–41296 Göteborg, Sweden

Prof. Dr. Heinz-Dieter Haustein
Hochschule für Ökonomie
Hermann Dunckerstr. 8
DDR–1157 Berlin, GDR

Dr. Gareth Horsnell
Depts. of Environment and Transport, Common Services
2 Marsham Street
London SW1P 3EB, UK

Mr. Alexander Kamerman
Estel NV
Postbus 401
6500 AK Nijmegen, the Netherlands

Prof. Todd La Porte
Institute of Governmental Studies
109 Moses Hall
University of California
Berkeley, California 94720, USA

Prof. Dr.hab. Alojzy Melich
President
Polish Academy of Science – Katowice branch
40039 Katowice, Poland

Dr. Johann Millendorfer
Study Group for International Analysis
Berggasse 16
1090 Vienna, Austria

Prof. Dr. Manfred Neumann
Universität Erlangen Nürnberg
Volkswirtschaftliches Institut
Lehrstuhl für Volkschaftslehre
Lange Gasse 20
D–8500 Nürnberg, FRG

Mr. Egbert Plug
NV Philips Gloeilampenfabrieken
ISA Research
Building HSM–4
Eindhoven, the Netherlands

Mr. Eric Price
Economics and Statistics Division
Dept. of Industry
Ashdown House
123 Victoria Street
London SW1, UK

Dr. Evka Razvigorova
Institute for Social Management
21, Pionerski pat st.
1635 Sofia, Bulgaria

Dr. Gustav Reczey
Ministry of Heavy Industries
Budapest, Hungary

Prof. John Rees
School of Social Science
The University of Texas at Dallas
Richardson, Texas 70580, USA

Prof. Gene I. Rochlin
Research Policy Analyst
Institute of Governmental Studies
University of California
Berkeley, California 94720, USA

Prof. Gerhard Rosegger
Case Western Reserve University
Cleveland, Ohio 44106, USA

Prof. Devendra Sahal
International Institute of Mangement
Griegstrasse 5–7
D–1000 Berlin (West)

Mr. C.T. Savin
Policy Research Unit
BP Ltd.
Britannic House
Moorlane, London EC2Y 9BU, UK

Dr. Winfried Schenk
Austrian Institute for Economic Research
Postfach 91
1103 Vienna, Austria

Mr. M.F. Shutler
Price Commission
Market Towers, Nine Elms Lane
Vauxhall, London SW8, UK

Dr. W.H.C. Simmonds
National Research Council of Canada
Montreal Road
Ottawa, Canada K1A OR6

Dr. J. Stachowicz
Dept. of Organization and Management
Polish Academy of Sciences
Bytom, pl. Kościuszki 9, Poland

Dr. Ivan Stanchev
College of Economics
Dept. of Management and Systems Theory
Exarch Jossif 14
Sofia 1000, Bulgaria

Prof. Josef Steindl
Austrian Institute of Economic Research
1103 Vienna, Postfach 91, Austria

Prof. Veselin Stoianov
Institute for Social Management
21, Pionerski pat st.
1635 Sofia, Bulgaria

Prof. A. Straszak
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw, Poland

Dr. Andrew Stratton
Imperial Chemical Industrial Ltd.
ICI House
Millbank, London SW1P 3JF, UK

Prof. Yutaka Suzuki
Faculty of Engineering
Osaka University
Yamada-Kami, Suita
Osaka, Japan

Dr. L. Uhlmann
IFO-Institute for Economic Research
5 Poschingerstrasse
Postfach 86-04-60
D–8000 München 86, FRG

Dr. James Utterback
Center for Policy Alternatives
Room E40-250
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA

Ing. J.D. van Dalen
Manufacturing Base Chemicals Division
Shell International Chemie Maatschappij B.v.
Postbus 162
Den Haag, the Netherlands

Ir. Rudy van Hees
TEO, Building EMB 1-18
N.V. Philips Gloeilampenfabrieken
5600 MD Eindhoven, the Netherlands

Dr. Harvey M. Wagner
School of Business Administration
The University of North Carolina at Chapel Hill
Carroll Hall 012 A
Chapel Hill, North Carolina 27514, USA

Mrs. Karin Wagner
Wissenschaftszentrum Berlin
Internationales Institut für Management
Platz der Luftbrücke 1-2
Berlin (West) 42

Dr. Gerhard Wittich
Hochschule für Ökonomie
Sektion LIS
Hermann-Duncker Strasse 8
DDR–1157 Berlin, GDR

IIASA PARTICIPANTS

Mr. Rolfe C. Tomlinson, Area Leader
Management and Technology Area

Mr. Mark F. Cantley
Management and Technology Area

Dr. Vladimir N. Glagolev
Management and Technology Area

Dr. Kiichiro Tsuji
Management and Technology Area

Prof. Bernard I. Spinrad
Energy Program

APPENDIX B

Principal Speakers	Affiliation	Title or Topic
TUESDAY, 26 JUNE 1979		
R.E. Levien	IIASA Director	Welcome to IIASA
R.C. Tomlinson	Leader, Management & Technology Area, IIASA	The work of the Management & Technology Area; objectives of the workshop
M.F. Cantley & V.N. Glagolev	Management & Technology Area, IIASA	Research topics in scale-related problems
J.C. Fisher	Corporate Research & Development, General Electric Company, USA	Optimum size of subcritical fossil fueled electric generating units
B.I. Spinrad	Oregon State University & IIASA	Scaling and learning in nuclear energy
G. Horsnell	Systems Analysis Research Unit, UK Department of the Environment	Productivity growth and related variables, an example from electricity production
G.G. Betts	Manager, Technical Development, BP Chemicals Ltd. UK	Implications of plant scale in the chemical industry with particular reference to ethylene plants
H. Derkx & A. Kamerman	Corporate Planning Department, Estel NV, the Netherlands	Studies of productivity and scale in the steel industry
R.N. van Hees	Central TEO, Philips Industries NV, the Netherlands	Main aspects determining the scale of an organization

Principal Speakers	Affiliation	Title or Topic
WEDNESDAY, 27 JUNE 1979		
J. Dathe	Industrieanlagen-Betriebs GmbH, FRG	Problems of scale in international air transportation
L. Uhlmann	IFO-Institute for Economic Research, FRG	Piecemeal growth of plants in industry?
D. Sahal	International Institute of Management, Berlin (West)	Scale, learning, and technological innovation
S.-O. Gustavsson	Division of Industrial Management, Chalmers University of Technology, Sweden	Motive forces for and consequences of different plant size
J. Stachowicz	Institute of Organizational and Management Problems, Polish Academy of Sciences, Bytom, Poland	Scale of collieries and their top-level management process requirements in the Polish coal mining industry
D.J. Daly	Faculty of Administrative Studies, York University, Canada	Scale economies and the options for a small country
I. Stanchev	College of Economics, Sofia, Bulgaria	Mathematical analysis of the influence of relationships on the determination of scale and organization structure of economic units of organization
J. Millendorfer	Study Group for International Analysis, Austria	Scale and other aspects of structural change in Austrian industry – quantitative studies
THURSDAY, 28 JUNE 1979		
G. Een	Research and Development Group, Strategic Process Planning, Alfa-Laval AB, Sweden	Food industry in the year 2000
G. Wittich & H.-D. Haustein	Hochschule für Ökonomie, GDR	Scale strategies for a small country – the experience of GDR industry
G.A. Egiazarian	Faculty of Economics, Moscow University, USSR	Problems of determining production scale in Soviet industry

Principal Speakers	Affiliation	Title or Topic
H.I. Ansoff	European Institute for Advanced Studies in Management, Brussels, Belgium	The management of management and the size of management
A.M. Belolipetsky	Institute for System Studies, USSR	The impact of an enterprise's size on organization structure
V. Stoianov & E. Razvigorova	Institute for Social Management, Sofia, Bulgaria	The "management" factor and the problem of size in economic organizations
M. Shutler	UK Price Commission, London, UK	Scale economies – the evidence from published reports of the British Price Commission
J. Utterback	Center for Policy Alternatives, Massachusetts Inst. of Technology, USA	New technology-based entrants, innovation, and production efficiency
FRIDAY, 29 JUNE 1979		
B. Gold	Case Western Reserve University, USA	Revising prevailing approaches to evaluate scale economies in industry
		General discussion of the issues raised by the workshop, future directions for research, and possibilities for research cooperation

SCALE IN PRODUCTION SYSTEMS

John A. Buzacott, Mark F. Cantley, Vladimir N. Glagolev, and Rolfe C. Tomlinson, Editors

When a new enterprise is planned or the scope of an organization is changed, a decision on scale has to be made. At a certain stage in the development of a technology, the economies of scale that have been well established are overtaken by factors that had hitherto been considered secondary. Mistakes of scale have been made, from hospitals to electricity-generating boards. The mistakes of being too large are the most eye-catching, but the mistakes of being too small are probably just as frequent, and just as important.

Twelve chapters of this book are based on reports of a workshop on scale and productive efficiency held at the International Institute for Applied Systems Analysis. The other eight chapters describe work done at the Institute; they also include discussion of the key issues relevant to scale that arose at the workshop, thus broadening the survey of this topic.

This book treats problems of scale at various levels: at the unit (the equipment) and plant levels; at the organization level; and at industrial and national levels.

The book includes practical case descriptions that will be of value to managers and decision makers, as well as material of value to research workers.

0 08 028725 5