

UNIVERSAL
LIBRARY

OU_162153

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

No 330
C 76

Accession No 37806

Author

Koyck L.M.

Title

Distributed lags and investment
analysis - 1954.

This book should be returned on or before the date last marked below.

**DISTRIBUTED LAGS AND
INVESTMENT ANALYSIS**

**CONTRIBUTIONS
TO
ECONOMIC ANALYSIS**

IV

EDITED BY

J. TINBERGEN

P. J. VERDOORN

H. J. WITTEVEEN



1954

**NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM**

DISTRIBUTED LAGS AND INVESTMENT ANALYSIS

BY

L. M. KOYCK



1954

**NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM**

COPYRIGHT 1954
N.V. NOORD-HOLLANDSCHE UITGEVERS MAATSCHAPPIJ
AMSTERDAM

Printed in the Netherlands

INTRODUCTION TO THE SERIES

This series consists of a number of hitherto unpublished studies, for the greater part, by Dutch authors. They are introduced by the editors in the belief that they represent fresh contributions to economic science.

The term *economic analysis* as used in the title of the series has been adopted because it covers both the activities of the theoretical economist and of the research worker.

Although the analytical methods used by the various contributors are not the same, they are nevertheless conditioned by the common origin of their studies, namely theoretical problems encountered in practical research. Since, for this reason, business cycle research and national accounting, research work on behalf of economic policy, and problems of planning are the main sources of the subjects dealt with, they necessarily determine the manner of approach adopted by the authors. Their methods tend to be "practical" in the sense of not being too far remote from application to actual economic conditions. In addition they are quantitative rather than qualitative.

It is the hope of the editors that the publication of these studies will help to stimulate the exchange of scientific information and to reinforce international co-operation in the field of economics.

THE EDITORS

CONTENTS

page

CHAPTER I. INTRODUCTION

I.1.	Estimating the time-shape of economic reactions, a neglected point in econometrics	1
I.2.	Scope of this study	3

CHAPTER II. THE TIME-SHAPE OF ECONOMIC REACTIONS AND THEIR STATISTICAL ESTIMATION

II.1.	Distributed lags and their causes	5
II.2.	Description of the time-shape of an economic reaction	9
II.3.	Asymmetry and irreversibility of economic reactions	12
II.4.	Difficulties in the statistical estimation of long-term elasticities	17
II.5.	An approximation of the time-shape of a reaction	19
II.6.	A method of estimating the long-term elasticity and the speed of adjustment	22
II.7.	A comparison with Tinbergen's method of measuring long-term elasticities	25
II.8.	Irving Fisher's short-cut method for calculating distributed lags	30
II.9.	Inconsistent estimates of the time-shape, development of a method to find consistent estimates . .	32

CHAPTER III. DISTRIBUTED LAGS IN THE ADJUSTMENT OF CAPACITY TO OUTPUT

III.1.	Capacity-output adjustment; importance of the problem	40
III.2.	Importance of correct economic analysis in econometric research	44
III.3.	Restriction of the problem: investments induced by demand	44
III.4.	The acceleration principle irrelevant to problems on the speed of capacity adjustment	45

	page
III.5. A model on entrepreneurial decisions concerning capacity and its rate of utilization	48
III.6. Short-run adjustments to changes in output . . .	50
III.7. Long-run adjustments to changes in output . . .	55
III.8. Rising supply curve of capital funds and expansion of capacity	58
III.9. Secular growth and building ahead of demand . .	61
III.10. Uncertain expectations and fluctuating demand	64
III.11. Distributed lags in capacity adjustment; a final refutable hypothesis	67

CHAPTER IV. EMPIRICAL INVESTIGATIONS ON THE TIME-SHAPE OF THE REACTION OF CAPACITY TO CHANGES IN OUTPUT

IV.1. Railway freight traffic	74
IV.2. The electric light and power industry	91
IV.3. Portland cement industry	98
IV.4. Steel industry	103
IV.5. Petroleum refining	105
IV.6. Summary of results	106

APPENDIX	111
--------------------	-----

PREFACE

The main ideas of this study go back to practical economic research at The Netherlands Economic Institute at Rotterdam during a couple of years.

I owe a great debt to the directors of The Netherlands Economic Institute, who afforded me the opportunity to finish this study. Especially Prof. H. W. LAMBERS and Prof. J. TINBERGEN have given me constant encouragement to go on with the work.

Part of the preliminary results of the empirical studies were discussed with Prof. R. M. GOODWIN at the Varese meeting of the Econometric Society in the autumn of 1950. Subsequently, I presented some of the results in a paper read before the Cambridge meeting of the Econometric Society in August 1952. On that occasion Mr H. B. CHENERY drew my attention to his article "Overcapacity and the Acceleration Principle" in *Econometrica*, January 1952, dealing with some aspects of our study. Both Prof. GOODWIN and Mr CHENERY gave some valuable hints for further research.

As to the statistical difficulties encountered in the course of this study, I received much help and constructive criticism from my friend and colleague Mr J. A. HARTOG, lecturer in statistics at the Netherlands School of Economics in Rotterdam. The core of the analysis from a statistical point of view, section 9 of the second chapter, resulted from his highly appreciated criticism.

I have had interesting discussions with my research-associates; the assistance of Mr E. H. MULDER, who helped me with the calculations and checked the manuscript, was especially valuable.

I also thank Mr M. VAN BEUZEKOM, who corrected my English.

L. M. KOYCK

Rotterdam, June 1954

CHAPTER I

INTRODUCTION

I.1. ESTIMATING THE TIME-SHAPE OF ECONOMIC REACTIONS, A NEGLECTED POINT IN ECONOMETRICS

At the beginning of this century, MARSHALL said in a famous address to an undergraduate audience in Cambridge: "Speaking generally the nineteenth century has in great measure achieved *qualitative* analysis in economics; but it has not gone farther. It has felt the necessity of *quantitative* analysis, and has made some rough preliminary surveys of the way in which it is to be achieved: but the achievement itself stands over for you" ¹).

The first half of this century has indeed shown great activity in the direction pointed out by MARSHALL. The rise of econometrics as a new scientific method in which economic theory and statistical research are applied in combination is there to demonstrate this. The late twenties and early thirties are marked by a number of bold and determined advances into the new territory. The pioneering work of SCHULTZ, FRISCH, TINBERGEN and ROOS belongs to the early "Sturm und Drang" period, as LEONTIEF called it ²), in the development of econometrics.

There was certainly much to criticize in early attempts of testing economic theories and giving quantitative expression to them by bringing them into relation with observations. It is a natural reaction to the first flush of enthusiasm and

¹) A. MARSHALL, The old generation of economists and the new, *Quarterly Journal of Economics*, 1897.

²) In *A Survey of Contemporary Economics*, edited by H. S. ELLIS, Philadelphia 1948, pp. 388-411.

over-optimism of the "Sturm und Drang" period, that in a second, more reflective stage in the development of the new field of empirical quantitative analysis in economics more stress is laid on the many weak points.

Considerable progress has been achieved in this second stage toward the understanding of proper and improper application of statistical inference to economic analysis and the work on improving the tools of econometrics is in full swing. But econometrics is a young branch of science and very much remains to be done.

One of the hitherto somewhat neglected points, and at the same time an important practical topic of the day, is the estimation of the difference between short-run and long-run elasticities. Up to now the attempts to estimate elasticities in economics have been mainly restricted to the estimation of short-run elasticities. "In the field of long-term reactions", as TINBERGEN says, "very little has as yet been undertaken. Also here, we have a vast unexplored territory for scientific work" ¹⁾).

An economic reaction, say the reaction of the quantity of a certain commodity demanded brought about by a change in its price, is a process in time. Part of the reaction may be worked out after the lapse of one period, another part of the reaction taking place during the second period and so on, until after a certain number of periods the total reaction is worked out and demand is fully adjusted to the new price. In other words, economic reactions have a certain time-shape.

Pure theory says that a fall in price, under certain conditions, causes a rise in the amount demanded. Econometrics should give a numerical estimate of this rise, when price falls by say 10 %. It may be important in practical problems to know how much of this rise may be expected to be realized after a short-period, say a year, and how much the rise will

¹⁾ J. TINBERGEN, *Econometrics*, London 1951, p. 156.

be in the long-run, i.e. when sufficient time has elapsed for the full adjustment of the quantity demanded. Generally it may be important to have some empirical knowledge of the time-shape of an economic reaction. It is the purpose of this study to deal with this problem.

I.2. SCOPE OF THIS STUDY

In Ch. II a theoretical analysis of the time-shape of economic reactions is given. A method for a quantitative description of the time-shape is developed. An approximation of the time-shape is introduced, which, we believe, may make it possible to surmount the specific difficulties of estimating an empirical time-shape, and consequently to estimate the long-term elasticity and speed of adjustment of economic reactions.

The last section of Ch. II gives some mathematical statistics, it discusses a complication of our estimation procedure and develops a method to overcome this difficulty. From a statistical point of view, it is the core of Ch. II.

The ideas of Ch. II will be applied in a number of empirical studies, for the application the reaction of capacity to changes in output of an industry is chosen. But before doing so, the main argument is interrupted by a theoretical analysis of the reaction which is to be estimated. This analysis is given in Ch. III, where traditional tools of economic theory are combined in the investigation of the adjustment of capacity to output, leading to a refutable hypothesis in mathematical form.

In Ch. IV the main argument is resumed. The ideas of Ch. II and III are combined in the empirical studies of the reactions of capacity to output in a number of American industries, viz. railway freight traffic, the electric light and power industry, the portland cement industry, the steel industry and the petroleum refining industry.

Estimating structural relations in economics is a difficult

job, the estimation of the time-shape of economic relations is even more difficult and complicated. In order to make scientific progress in a complicated situation, it is useful to proceed step by step, introducing the various complications successively rather than simultaneously. This study, it is hoped, is one minor step in a great number of successive steps still ahead in the field of estimating structural economic relations from time series data. This implies that the empirical investigations of Ch. IV are only made to illustrate the treatment of one kind of difficulty, the estimation of the time-shape of economic reactions. Many other difficulties in the estimation procedures of Ch. IV are purposely left out of consideration.

The series of "contributions to economic analysis" is introduced by the editors in the belief that they represent fresh contributions to economic science. Therefore it may be useful to indicate what we believe to be new in this booklet. Apart from the empirical studies of Ch. IV, the approximation of the time-shape of economic reactions and the method of estimation which is then derived are new. Also new, as far as we know, is the mathematical-statistical investigation of section 9 of Ch. II. As said before, Ch. III is an application of traditional economic theory.

CHAPTER II

THE TIME-SHAPE OF ECONOMIC REACTIONS AND THEIR STATISTICAL ESTIMATION

II.1. DISTRIBUTED LAGS AND THEIR CAUSES

Some general remarks about the time-shape of economic reactions will now be made. When speaking of economic reactions we mean the reactions of economic subjects e.g. producers or consumers, as described by means of behaviour equations.

In order to make the analysis as general as possible to begin with, let us introduce an endogenous economic variable y that reacts to changes in an exogenous variable x . We can think of y as representing the quantity of a certain commodity produced in a given period and of x as the price of that commodity; or y may be the demand for a consumers' good and x the disposable income of consumers; or one can think of any other behaviour relation in which y is an endogenous and x an exogenous variable.

As a starting point we suppose that x has been constant during a period long enough for the adjustment of y to this level of x to be completed. In order to avoid non-essential complications it is supposed that y only depends on x . This being taken for granted and y being fully adjusted to the constant value of x , a constant value of y will correspond with the constant level of x ¹⁾. This value of y represents the equilibrium in the behaviour of the households concerned corresponding

¹⁾ It is implicitly assumed here that the process of adjustment is converging. This assumption is generally made in economic theory: the equilibrium of the firm and that of the consumer are supposed to be stable and the hypotheses of falling marginal productivity and falling marginal utility are usually introduced to that end.

to the given value of x ; it is the equilibrium level of supply in a behaviour equation of suppliers or the equilibrium level of demand in a demand equation.

Let us now suppose that at time $t = 0$ the exogenous variable in the behaviour equation, i.e. x , increases and remains constant at this higher level afterwards. How can the reaction of y as a process in time be described?

As the increase of x disturbs the existing equilibrium of y , changes in economic behaviour will follow. In other words y will move towards a new equilibrium as determined by the new level of x . The reaction of y will not be concentrated in any one point of time, it will be spread over a period of time. The general case is that there will be a distributed lag, long or short, in the reaction of y to x .

There are several reasons for the reactions of producers or consumers being spread over a certain period of time. This is a rough classification of these reasons:

a. *objective reasons*

1. technological
2. institutional

b. *subjective reasons.*

The technological reasons are related with the fact that production of investment and consumers' goods requires time. A decision of an entrepreneur to expand production will result in a larger flow of products only after a certain lapse of time ¹⁾).

The technological lifetime of durable goods, investment goods and consumers' goods, is another instance of objective resistance to a complete and immediate adaptation to changed

¹⁾ This technological lag plays an important role in various business cycle theories. In the theory of Aftalion e.g. the time required for the production of investment goods is the fundamental reason for periodic overproduction. In the modern theories of FRISCH, TINBERGEN, KALECKI a.o. the period of production is *one* of the dynamic factors.

circumstances. A well-known textbook illustration is given by G. J. STIGLER. When dealing with the effects of a decrease in the price of electricity he says: "durable goods (e.g. gas stoves) may have to be worn out before a shift is made to the product whose price has fallen" ¹⁾).

The restrictions in the behaviour of entrepreneurs with respect to durable investment goods are of great importance in dynamic theory. Suppose we have a firm in a situation of full-equilibrium, the size of production corresponding with optimal utilisation of its plant. Now sales alter, leading to an initial change in stocks. After a period of changing its stocks the firm will, as a rule, adjust its production to the new level of sales. But generally the producer is neither able nor willing to alter the size of his plant simultaneously with the change in production. As a consequence of the durability and immobility of investment goods the firm will have excess or short capacity. In case of excess capacity a reduction of this capacity is feasible only after a certain lapse of time, the length of which depends on the age distribution and the lifetime of the various investment goods. In case of short capacity the entrepreneur may postpone the decision to expand his plant for some time, until he considers increased sales as sufficiently permanent. Furthermore some time will elapse between the decision to expand and actual expansion, its duration depending, among other things, on the terms of delivery of investment goods.

Institutional factors resulting in a lagged reaction are the rules observed — through law or custom — when contracts are made. To give some examples: payment of salaries at the end of the month, of dividends once a year or a quarter; the period of notice agreed upon in a labour-contract or a lease; the terms of redemption of loans, etc. Thus, an increase in demand for agricultural products, *ceteris paribus*, will induce land-owners to ask higher rents, but this reaction is spread

¹⁾ G. J. STIGLER, *The theory of price*, New York 1949, p. 94.

over a period of time, the length of which depends on the dates of expiration of the leases.

Subjective reasons for lagged reactions are imperfect knowledge of the market and psychological inertia of economic subjects. A fall in price may not be known to every potential buyer immediately, in that case the total effect of the change in price will only be realised when every potential buyer is fully aware of the price cut. Imperfect knowledge of the market may not be very important with respect to goods bought regularly and frequently, say weekly. It may be more important in the case of goods bought infrequently like furs, and it is a more important factor in the behaviour of consumers than in that of producers or dealers.

Psychological inertia prevents instantaneous readjustment of the behaviour of economic subjects to a changed situation, in that case habit leads to lagged reactions.

Once a consumer has found that a certain brand of coffee satisfies his taste at given prices and qualities, he will probably not make the mental efforts necessary to compare again the various brands of coffee as to quality and price every time he buys coffee. Habit is a great saver of mental effort and may persist although the reason for it has disappeared. It may require a considerable period of time for the consumer to rearrange his expenditure plan after a fall in price of better qualities or a deterioration in quality of the coffee he has been using.

The lag in the behaviour of an individual economic subject — whatever the reason for the lag — may in some cases consist of one definite period of time. The decision of a single entrepreneur to increase production may be realised in a higher output at one certain point of time, after which output remains constant at the higher level for some time. A rise in the price of gasoline may induce the owner of a gasoline-truck to replace it by a dieseloil-truck. Then again the time lag between the rise in gasoline price and the moment of

replacement consists of one discrete period of time. A higher rate of salaries for office-clerks may lead to a decision of an entrepreneur to dismiss a number of clerks and to buy a calculating machine after one definite period of time, say one month, depending on the stipulations of the labour-contract. Generalising we may say that the lag in the reaction of an individual economic subject may have one determinate length.

When dealing with the reactions of a number of individuals, e.g. all consumers or all entrepreneurs in a certain industry, the total reaction will be spread over a certain period of time. Some individuals will react nearly immediately or after a short lag, some will be very slow for one reason or another and will react after a long period, others may be grouped between these extremes. Generally the lag in the reactions of a number of subjects will be distributed over a period of time, it will be a "distributed lag".

II.2. DESCRIPTION OF THE TIME-SHAPE OF AN ECONOMIC REACTION

In order to illustrate the description of the time-shape of reactions, we return to the starting point of the previous section. We supposed there that x has been constant sufficiently long for y to be fully adapted to the constant level of x . At time $t=0$ an increase of x was introduced, causing a change in y in the direction of a new equilibrium of y .

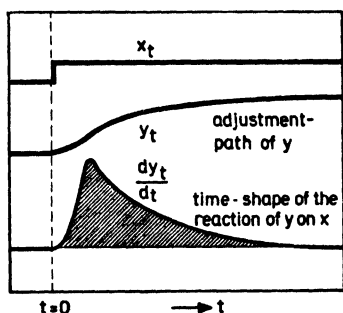
In graph 1 a graphical description of a possible adjustment process is given by the adjustment path of y .

We define the time-shape of the reaction of y to x as the time series of the change in y per unit of time caused by a change in x ; in graph 1 this time-shape is given by the curve dy_t/dt ¹⁾.

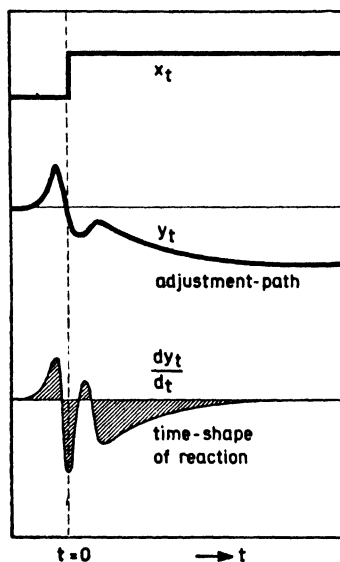
The time-shape of graph 1 is simple, though perhaps in many cases rather realistic. More complicated shapes are possible however. To give an example, let y be the demand (say per month) for and x the price of a certain not im-

¹⁾ A more exact definition will be given on p. 12.

mediately perishable good. If the rise in price at time $t=0$ is known to buyers beforehand or is anticipated by them, demand y may increase before prices start to rise, only to decline afterwards and this perhaps more sharply than would have been the case had there been no anticipating increase in demand. In graph 2 this kind of time-shape is illustrated.



Graph 1



Graph 2

It should be pointed out here, however, that the time-shape of a reaction, as defined above, depends on the length of the unit-period. When e.g. y is not demand per month but per year or quarter, the time-shape of the reaction becomes more regular.

Following the graphical description of the time-shape, an analytical and more general manner of describing time-shapes will be given now.

In II.1. it was shown that in general the lag between y and x will be a distributed one. Analytically this can be written as follows:

(II.2.1) $y_t = f(x_t, x_{t-1}, x_{t-2}, \dots)$, or as a linear approximation:

$$(II.2.2) \quad y_t = \alpha_0 x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots = \sum_{i=0}^{\infty} \alpha_i x_{t-i}.$$

The coefficients α in (II.2.2) are reaction coefficients. Supposing that the reaction of y after a disturbance of an equilibrium situation will lead to another equilibrium, x remaining constant after the disturbance, the series $\sum \alpha_i$ must be converging. This implies that $\lim_{n \rightarrow \infty} \alpha_n = 0$.

The series of x -values in (II.2.2) is a discontinuous one, the lags introduced are discrete lags. If one wants to work with continuous lags the equation (II.2.2) must be replaced by the integral-equation $y_t = \int_0^{\infty} \alpha_i x_{t-i} di$.

Henceforth we shall use a reaction equation of the form (II.2.2) as this analysis is simpler and more easily suited to empirical verification. Equation (II.2.2) can be considered as an approximation of an integral-equation with continuous lags.

It can easily be proved that the graphical description of a reaction as given in graph 1 corresponds with the analytical description as given by equation (II.2.2). We again choose the units in which and the level from which x and y are measured in such a way that $y_t = 0$ and $x_t = 0$ for $t \leq 0$ and $x_t = 1$ for $t > 0$.

The values of y for $t > 0$ follow directly from equation (II.2.2), they are given below:

t	0	1	2	3	.	.	.	∞
y_t	0	α_0	$\alpha_0 + \alpha_1$	$\alpha_0 + \alpha_1 + \alpha_2$	.	.	.	$\sum_0^{\infty} \alpha_i$

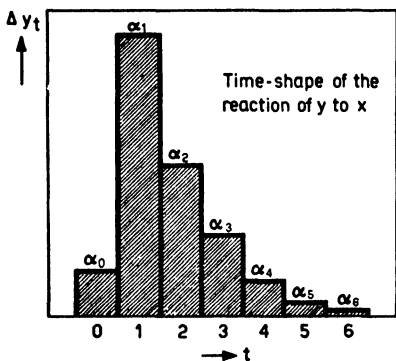
The change in y per unit of time follows from (II.2.2):

$$(II.2.3) \quad \Delta y_t = y_{t+1} - y_t = \sum_{i=0}^{\infty} \alpha_i \Delta x_{t-i}.$$

The values of Δy_t for $t \geq 0$ resulting from (II.2.3) are given below:

t	0	1	2	3	.	.	.	.	∞
Δy_t	α_0	α_1	α_2	α_3	.	.	.	.	$\alpha_\infty = 0$

This time series Δy_t is represented in graph 4, it corresponds with the curve of dy_t/dt in graph 1. Similarly as in the case of



Graph 3

continuous lags (the curve of dy_t/dt in graph 1), the curve represented in graph 3 can be used to define the time-shape of the reaction of y to x in the discontinuous case. We can now define the time-shape of the reaction of y to x in analytical terms as follows: the time-shape of the reaction of y to x is the series of reaction coefficients α_i of equation (II.2.2).

The linear approximation of the general form (II.2.1) has the great advantage of being easily manageable in empirical work. Another approximation of (II.2.1), also manageable in empirical calculations, is an exponential function:

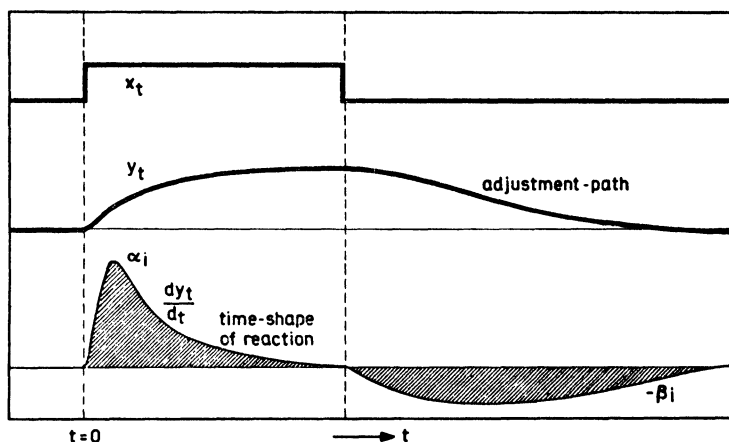
$$(II.2.4) \quad y_t = x_t^{\alpha_0} x_{t-1}^{\alpha_1} x_{t-2}^{\alpha_2} \dots = \prod_{i=0}^{\infty} x_{t-i}^{\alpha_i}.$$

In this case, a one per cent increase in x results in a change in y of α_0 per cent during the first period, an additional α_1 per cent during the second period, and so on.

II.3. ASYMMETRY AND IRREVERSIBILITY OF ECONOMIC REACTIONS

In general it is possible that the time-shape of the reaction of y to a rise of x is different from the time-shape of the reaction of y to a fall of x . An illustration of this case is given in graph 4, we can say that in this case the reaction is asym-

metric. The fact that reactions may be asymmetric has been stressed in theoretical essays in recent years ¹⁾. Anticipating the statistical investigations of later chapters, it can be said – to give an example – that the reaction of the stock of railway freight cars in the U.S.A. to an increased demand for transport is much faster than the reaction to a fall in demand for transport.



Graph 4. Asymmetric reactions

When the reaction is asymmetric, equation (II.2.3) can be split up into two equations, one for an increase of y and one for a fall of y :

$$(II.3.1) \quad \Delta y_t = \sum_{i=0}^{\infty} \alpha_i \Delta x_{t-i}$$

$$(II.3.2) \quad \Delta y_t = \sum_{i=0}^{\infty} \beta_i \Delta x_{t-i}.$$

In the same way equation (II.2.1) leads to:

$$(II.3.3) \quad y_t = \sum_{i=0}^{\infty} \alpha_i x_{t-i}$$

$$(II.3.4) \quad y_t = \sum_{i=0}^{\infty} \beta_i x_{t-i}.$$

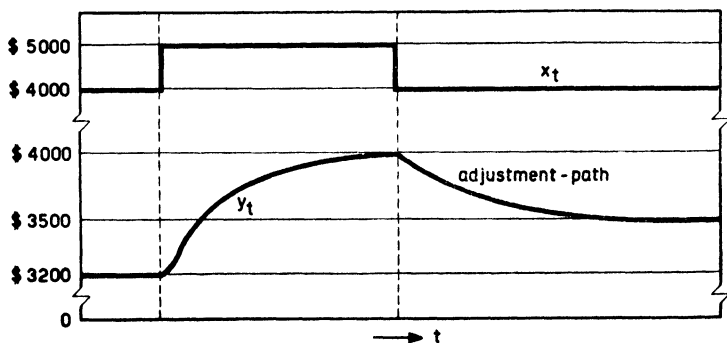
¹⁾ J. R. HICKS, *A contribution to the theory of the trade cycle*, London 1950. R. M. GOODWIN, *A non-linear theory of the cycle*, *Review of Economics and Statistics*, 1950.

In graph 4 it was implicitly assumed that the later fall of x to the level it had reached before $t=0$ caused a fall of the equilibrium value $\bar{y}$ to the original level of $\bar{y}$ before $t=0$. In other words the full extent of the reaction is the same in the case of a rise and a fall of y , i.e. $\sum \alpha_i = \sum \beta_i$; it is only the speed of the reaction which is different in both cases.

There may be cases where not only the speed but also the extent of the reactions differ. For example, a rise in someone's income from say \$ 4.000, to \$ 5.000 will lead to an increase in consumption expenditure from say \$ 3.200 to \$ 4.000; if afterwards income should fall again to \$ 4.000 it might well be possible that consumption expenditure would not fall from \$ 4.000 to the original level of \$ 3.200 again but might stay at a higher level, of say \$ 3.500, the reason for this being e.g. the fact that one does not want to give up all the pleasures enjoyed at the higher income level, not even in the long-run.

In this case the long-run marginal propensity to consume in case of a rise in income is not the same as the long-run marginal propensity to consume in case of a fall in income. When x is income and y consumption expenditure in (II.3.1) and (II.3.2), this can also be pointed out by saying $\sum \alpha_i \neq \sum \beta_i$.

This being the case we shall say that the reaction is irreversible. A graphical illustration of the example of irreversibility given above, is found in graph 5.



Graph 5. Irreversibility

MARSHALL alludes to the irreversibility of reactions, when he makes the following remarks about one of his statements, where reversibility was assumed implicitly. "It must however be admitted that this theory is out of touch with real conditions of life, in so far as it assumes that if the normal production of a commodity increases and afterwards again diminishes to its old amount, the demand price and the supply price will return to their old positions for that amount. Whether a commodity conforms to the law of diminishing or increasing return, the increase in consumption arising from a fall in price is gradual; and, further, habits which have once grown up around the use of a commodity while its price is low, are not quickly abandoned when its price rises again. If therefore after the supply has gradually increased, some of the sources from which it is derived should be closed, or any other cause should occur to make the commodity scarce, many consumers will be reluctant to depart from their wonted ways" ¹⁾.

Summarising one can distinguish the following cases:

- 1) for every i $\alpha_i = \beta_i$ in (II.3.1) and (II.3.2). The reaction is symmetric and reversible;
- 2) for one or more values of i α_i in (II.3.1) differs from β_i in (II.3.2), nevertheless $\sum \alpha_i = \sum \beta_i$. In this case the reaction is asymmetric and reversible;
- 3) for one or more values of i α_i in (II.3.1) differs from β_i in (II.3.2), and $\sum \alpha_i$ differs from $\sum \beta_i$. The reaction is asymmetric and irreversible. It will be clear that irreversibility implies asymmetry.

Hitherto an equilibrium situation was used as starting point for the analysis of a lagged reaction and it was assumed that a new equilibrium was attained before another disturbance was introduced (see graphs 4 and 5 e.g.). This was done for expository purposes, to make it easy to arrive at a systematic

¹⁾ A. MARSHALL, *Principles of economics*, London 8th ed. 1938, p. 807.

description of the time-shape of reactions as given by equations (II.2.2) to (II.3.4). In general these equations indicate how y is determined not only by the present but also by a series of past values of x . One might say that y is determined by a history of values of x , or rather by a weighted average of historical values x_{t-i} where $i=0, 1, \dots \infty$, the reaction coefficients α_i being proportional to the weights of this average.

The elasticity of a reaction can now be specified as to the period of time to which the reaction relates. From equation

$$(II.2.2) \quad y_t = \alpha_0 x_t + \alpha_1 x_{t-1} + \dots$$

follows that the elasticity of the reaction of y to x during the first period, say the first "year" after the change in x is:

$$(II.3.5) \quad \varepsilon_1 = \alpha_0 \cdot \frac{x}{y}.$$

The elasticity of the reaction during the first and second year is $\varepsilon_2 = (\alpha_0 + \alpha_1)x/y$ etc. The long-term elasticity is the elasticity of the total reaction on the change of x , i.e.

$$(II.3.6) \quad \varepsilon_t = \left(\sum_{i=0}^{\infty} \alpha_i \right) \frac{x}{y}.$$

The short-term reaction will usually be defined as the reaction during the first period, then the short-term elasticity ε_s will be identical with ε_1 , so

$$(II.3.7) \quad \varepsilon_s = \alpha_0 \frac{x}{y}.$$

When the reaction is asymmetric or irreversible the short-term and long-term elasticities are easily calculated from equations (II.3.1) and (II.3.2).

It will be clear from what is said above that it may be quite insufficient to speak of *the* elasticity of an economic reaction; in many cases it will be necessary to specify the period of time to which the elasticity applies.

II.4. DIFFICULTIES IN THE STATISTICAL ESTIMATION OF LONG-TERM ELASTICITIES

The usual method of estimating the coefficients α_i in (II.2.2) is the method of multiple correlation, where y_t has to be explained by a series of explanatory variables x_{t-i} . In theory i may go from zero to infinity, in practice it will be permissible to stop the series of explanatory variables at x_{t-n} , where x_{t-n} is so far back in history that one can reasonably suppose that the total influence of x -values more than n periods passed, can be neglected.

This method will yield unsatisfactory results in all cases where the reaction is spread over a long period so that n is larger than say 2 or 3 periods. In estimating the coefficients α for values of x more than 2 or 3 periods lagged, the relative margins of error in these α 's will generally be large. One of the main reasons for the unreliability of these estimates will be the presence of multicollinearity; the time series $x_t, x_{t-1}, x_{t-2}, \dots$ etc. being highly intercorrelated as a rule.

The method for estimating distributed lags developed by F. L. ALT may serve as an example to illustrate these difficulties¹⁾. In outline his method is as follows. At first y_t is explained by x_t , then a second calculation is made where x_{t-1} is added, and so gradually $x_{t-2}, x_{t-3}, \dots$ etc. are added, "as long as the resulting coefficients α make sense". "After some time the coefficients will become erratic" (ALT also points out that this is usually the consequence of the high intercorrelation of the lagged explanatory variables). "This is a signal that too many independent variables are being introduced, and the process should be stopped. Computation of the standard errors of the regression coefficients may suggest stopping the process at an earlier stage, if the standard errors are so large that the coefficients can no longer be considered

¹⁾ F. L. ALT, Distributed lags. *Econometrica*, 1941, p. 113 ff. We used our notation in the exposition of Alt's ideas and calculations.

as giving reliable information about the shape of the distribution function”¹⁾).

This method was applied by ALT, among others, to the relation between the domestic consumption of residual fuel oil and general business conditions (as represented by an over-all index of new orders compiled by the Institute of Applied Econometrics) in the United States.

Denoting quarterly averages (calculated from monthly figures) of residual fuel oil consumption by y and quarterly averages of new orders by x and using a quarter as the unit period, the results of ALT’s calculations for the period January 1930 to June 1939 were:

(a) $y_t = 8,37 + 0,171 x_t$

(b) $y_t = 8,27 + 0,111 x_t + 0,064 x_{t-1}$

(c) $y_t = 8,27 + 0,109 x_t + 0,071 x_{t-1} - 0,005 x_{t-2}$

(d) $y_t = 8,32 + 0,108 x_t + 0,063 x_{t-1} + 0,022 x_{t-2} - 0,020 x_{t-3}$.

“Apparently the first two equations are reasonable. So is the third, because the coefficient of x_{t-2} is practically equal to zero, and the equation is practically identical with (b). In the fourth equation, however, the coefficients become erratic. Apparently the intercorrelation between the independent variables has become too high relative to the number of variables used, and the coefficients are unreliable. The second equation was therefore used”.

The difficulties of the usual method are clearly illustrated by this case. The second equation was used, not because the influence of new orders of two and more quarters ago was zero, but because the influence could not be safely estimated. This is a serious handicap when it might reasonably be supposed that the reaction is spread over a long period (this was probably not so in the case analysed by ALT).

Knowing the value of every individual coefficient α_i of the series of coefficients may not be very important. It will be im-

¹⁾ *Op. cit.* p. 116.

portant to know at least two things: in the first place the sum of all the coefficients, because it is this sum $\sum \alpha_i$ that determines the long-term elasticity $\varepsilon_t = (\sum \alpha_i) x/y$; in the second place, the period that must elapse for the total reaction to work out; this gives us an indication of the speed of adjustment to the equilibrium.

II.5. AN APPROXIMATION OF THE TIME-SHAPE OF A REACTION

We shall now introduce some assumptions concerning the series α_i and prove that this will often enable us to estimate:

- a. the long-term reaction as given by $\sum \alpha_i$;
- b. the period that has to elapse for the long-term reaction to work out, or in other terms, the speed of adjustment.

It might be expected that the technical, institutional and psychological hindrances to an immediate adaptation to changed data are gradually overcome. This idea was incorporated in the adjustment paths of y as represented in the graphs of the previous sections, the influence of a change in x on y gradually tapering off as time proceeds until the full effect on y is realised.

It was put forward in (II.2.2) that the value of y_t was determined by a complete set of values which x had assumed in the past, or in other words, by a weighted average of x_t, x_{t-1}, x_{t-2} etc., the weights $\omega_0, \omega_1, \omega_2$, etc. being proportional to the reaction coefficients $\alpha_0, \alpha_1, \alpha_2$, etc.

The previous statement about the tapering off of the effect of a change in x means that in the series of weights $\omega_0, \omega_1, \omega_2, \dots$ the weights will decrease from a certain index onwards, gradually approaching zero. After a certain point the influence of the lagged values of x on y will become the smaller, the longer the lag.

In order to reduce the number of explanatory variables, x , and hence the difficulties of intercorrelation between them and the amount of work in calculations, we now introduce an assumption about the weights ω_i and, which comes to the

same thing, about the coefficients α_i . As an approximation we introduce the hypothesis that from a certain index i onwards the weights ω_i and reaction coefficients α_i will decrease in a constant proportion.

Or, wording our hypothesis more carefully: we suppose that the series of coefficients α_i $i = 0, 1, 2, \dots$ from a certain index $i = k$ onwards can be approximated by a converging geometric series, thus:

$$(II.5.1) \quad \alpha_{k+m} = \lambda \alpha_{k+m-1},$$

where $m \geq 0$ and $0 \leq \lambda < 1$.

It is interesting to see the consequences of hypothesis (II.5.1) with respect to the adjustment path of y .

From the hypothesis (II.5.1) follows that:

$$(II.5.2) \quad \begin{cases} y_t = \alpha_0 x_t + \alpha_1 x_{t-1} + \dots + \alpha_{k-1} x_{t-k+1} + \alpha_k x_{t-k} + \alpha_k \lambda x_{t-k-1} + \\ \quad + \alpha_k \lambda^2 x_{t-k-2} + \dots \quad 0 \leq \lambda < 1. \end{cases}$$

The values of y for every $t \geq 0$ are easily calculated from (II.5.2) when x_t rises from a constant level $x_t = 0$ for $t \leq 0$ to a level $x_t = 1$ for $t > 0$.

For y_{k+n} where $n > 0$, and writing β for $\sum_{i=0}^{k-1} \alpha_i$, we find that:

$$(II.5.3) \quad y_{k+n} = \beta + \alpha_k \lambda + \alpha_k \lambda^2 + \dots + \alpha_k \lambda^{n-1} = \beta + \frac{\alpha_k}{1-\lambda} - \frac{\alpha_k}{1-\lambda} \cdot \lambda^n.$$

The new equilibrium value $\bar{y}$ is found to be:

$$(II.5.4) \quad \bar{y} = \lim_{n \rightarrow \infty} y_{k+n} = \beta + \frac{\alpha_k}{1-\lambda}.$$

The distance of y_{k+n} from the equilibrium $\bar{y}$ is found from (II.5.3) and (II.5.4); when this distance is denoted by d_{k+n} :

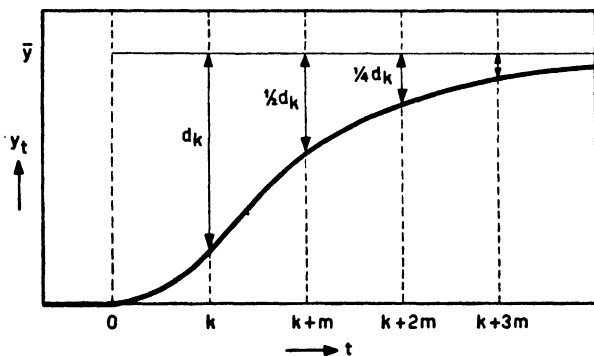
$$(II.5.5) \quad d_{k+n} = \bar{y} - y_{k+n} = \frac{\alpha_k}{1-\lambda} \cdot \lambda^n.$$

At last from (II.5.3) the change in y_{k+n} per unit of time is easily calculated:

$$(II.5.6) \quad \Delta y_{k+n} = \alpha_k \lambda^n = (1-\lambda) d_{k+n}.$$

The following conclusions can be drawn from these equations:

- a. y is an exponential function, in n from $t=k$ onwards, i.e. an exponential function of time, when time is measured from $t=k$ as a basing point. The adjustment path of y for $t > k$ will be an exponential curve. See (II.5.3) and graph 6.



Graph 6. Adjustment path of y following from hypothesis (II.5.1)

- b. The distance of y_{k+n} from the equilibrium value of y is getting smaller proportionately as time goes on, e.g. constantly halved in the same period m ; see (II.5.5) and graph 6.
- c. From $t=k$ the speed of adjustment of y , i.e. Δy_{k+n} , is proportional to the distance of y_{k+n} from the equilibrium level of y . See (II.5.6).

Summarising it can be said that the assumption made is equivalent to an approximation of the adjustment path from the point $t=k$ onwards by an exponential curve.

When (II.2.4) $y_t = x_t^{\alpha_0} x_{t-1}^{\alpha_1} \dots$ is used as behaviour equation the assumptions introduced above lead to $y_t = x_t^{\alpha_0} x_{t-1}^{\lambda \alpha_0} x_{t-2}^{\lambda^2 \alpha_0} \dots$ (for the sake of convenience it is assumed here that the approximation is valid from α_0 on).

A one per cent increase in x will change the equilibrium of y by $\alpha_0/(1-\lambda)$ per cent. After n periods y has reached a value y_n

which is $\alpha_0(1-\lambda^n)/1-\lambda$ per cent above the old level and hence still at a distance of $\alpha_0\lambda^n/1-\lambda$ per cent from $\bar{y}$. The percentage change in y_n per unit period is $\alpha_0\lambda^n$ or $(1-\lambda)$ times the percentage distance of y_n from the equilibrium $\bar{y}$. In this case the approximation of proportionately decreasing coefficients leads to the assumption that the percentage change in y_n is proportionate to the percentage distance of y_n from the equilibrium level of y .

The great advantages of the approximation made above will become more evident in the next sections.

II.6. A METHOD OF ESTIMATING THE LONG-TERM ELASTICITY AND THE SPEED OF ADJUSTMENT

Let us begin with the simple case where the whole adjustment path of y is approximated by an exponential curve, i.e. $k=0$. In this case

$$(II.2.2) \quad y_t = \sum_{i=0}^{\infty} \alpha_i x_{t-i}$$

can be written as:

$$(II.6.1) \quad y_t = \sum_{j=0}^{\infty} \alpha_0 \lambda^j x_{t-j} \quad 0 \leq \lambda < 1.$$

For Δy_t a simple relation can now be deduced as follows:

$$(II.6.2) \quad y_{t+1} = \alpha_0 x_{t+1} + \alpha_0 \lambda x_t + \alpha_0 \lambda^2 x_{t-1} + \dots$$

$$(II.6.3) \quad \lambda y_t = \alpha_0 \lambda x_t + \alpha_0 \lambda^2 x_{t-1} + \dots$$

$$(II.6.4) \quad y_{t+1} - \lambda y_t = \alpha_0 x_{t+1}, \text{ or}$$

$$(II.6.5) \quad y_{t+1} - y_t = \alpha_0 x_{t+1} - (1-\lambda)y_t, \text{ or}$$

$$(II.6.6) \quad \Delta y_t = \alpha_0 x_{t+1} - \gamma y_t, \text{ when } \gamma = 1 - \lambda.$$

A comparison of (II.6.6) and (II.2.2) makes it clear that the hypothesis made, reduces the estimation of reaction coefficients to statistically manageable proportions, instead of estimating a large number, in principle an infinite number, of coefficients in (II.2.2), only two coefficients in (II.6.6) have to be estimated.

As said before we wish to know the coefficient for the long-run reaction α_t and an indication of the speed of adjustment.

When equilibrium is attained $\Delta y_t=0$, hence it follows from (II.6.6) that:

$$(II.6.7) \quad \bar{y}_t = \frac{\alpha_0}{\gamma} x_t.$$

The coefficient for the long-run reaction α_t resulting from (II.6.7) is:

$$(II.6.8) \quad \alpha_t = \frac{\alpha_0}{\gamma}.$$

It is easily proved that the same value for α_t results from (II.6.1) by adding all the α 's:

$$(II.6.9) \quad \alpha_t = \sum_{j=0}^{\infty} \alpha_0 \lambda^j = \frac{\alpha_0}{1-\lambda} = \frac{\alpha_0}{\gamma}.$$

As to the speed of adjustment, it will now be shown that the coefficient γ of (II.6.6) is a good indicator of the speed of adjustment. Substitution of (II.6.7) in (II.6.6) gives:

$$(II.6.10) \quad \Delta y_t = \gamma(\bar{y}_{t+1} - y_t), \quad 0 < \gamma \leq 1,$$

here Δy_t is the actual change from y_t to y_{t+1} , $\bar{y}_{t+1} - y_t$ is the change that would be necessary to attain equilibrium at $t+1$, so (II.6.10) says that the actual change of y during one period is a proportion γ of the change that would be necessary to attain equilibrium in that period. The higher γ (within the boundary-condition $0 < \gamma \leq 1$), the higher the speed of adjustment and conversely.

For the highest value $\gamma=1$ it follows from (II.6.10) that

$$(II.6.11) \quad \Delta y_t = \bar{y}_{t+1} - y_t,$$

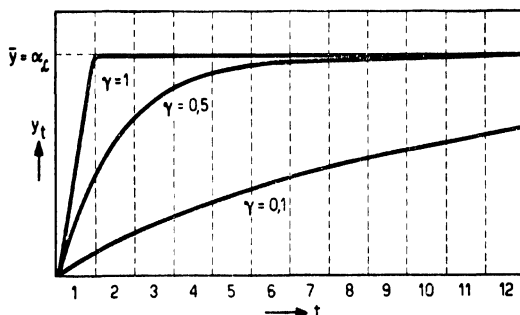
or

$$(II.6.12) \quad y_{t+1} = \bar{y}_{t+1};$$

the speed of adjustment is so great that y always reaches its equilibrium in the unit period of time. In this extreme case

(II.6.1) reduces to $y_{t+1} = \alpha_0 x_{t+1}$, all lagged values of x in the reaction equation have disappeared. The other extreme when $\gamma \propto 0$ is a situation when the speed of adjustment is nearly zero and equilibrium is reached after a period that may surpass any preconceived finite period.

A graphical illustration of the meaning of γ is given in graph 7. The adjustment path of y when $y_t=0$ and $x_t=0$ for $t \leq 0$ and $x_t=1$ for $t > 0$ is drawn for three values of γ : $\gamma=1$, $\gamma=0,5$ and $\gamma=0,1$.



Graph 7. Adjustment paths with different speeds of adjustment

In the general case where $k \neq 0$ the equation for Δy_t is deduced along the same lines:

$$(II.6.13) \quad y_{t+1} = \alpha_0 x_{t+1} + \alpha_1 x_t + \dots + \alpha_k x_{t-k+1} + \alpha_k \lambda x_{t-k} + \alpha_k \lambda^2 x_{t-k-1} + \dots$$

$$(II.6.14) \quad \lambda y_t = \frac{\alpha_0 \lambda x_t + \alpha_{k-1} \lambda x_{t-k+1} + \alpha_k \lambda x_{t-k} + \alpha_k \lambda^2 x_{t-k-1} + \dots}{\lambda}$$

$$(II.6.15) \quad y_{t+1} - \lambda y_t = \alpha_0 x_{t+1} + (\alpha_1 - \alpha_0 \lambda) x_t + \dots + (\alpha_k - \alpha_{k-1} \lambda) x_{t-k+1}, \text{ or}$$

$$(II.6.16) \quad \Delta y_t = \delta_0 x_{t+1} + \delta_1 x_t + \dots + \delta_k x_{t-k+1} - \gamma y_t.$$

In (II.6.16), $k+2$ reaction coefficients have to be estimated; when the exponential approximation of the adjustment path is introduced e.g. at the fourth coefficient of (II.2.2), then $k=2$ and four coefficients have to be estimated. Hence, when $k=0, 1$ or 2 , the *specific* difficulties of estimating the very large, in principle infinite, number of coefficients of (II.2.2) have been overcome. In order to make this method quite clear and to compare it with other methods used in estimating

distributed lags a concrete example will be worked out in the following section.

II.7. A COMPARISON WITH TINBERGEN'S METHOD OF MEASURING LONG-TERM ELASTICITIES

In view of the problems facing Western European countries with regard to balancing their foreign accounts and when estimating the degree of success obtained by devaluations or wage reductions as methods to equilibrate their balance of payments, TINBERGEN was interested in long-term foreign trade elasticities ¹⁾. "The elasticity of the volume of imports or exports of any one country with respect to its price level may be either conceived of as a short-run elasticity or as a long-run elasticity". A lot of work had been done so far in measuring these elasticities but no attempt had been made, however, to include lagged values of the price series and hence the measurements must essentially be seen as attempts to measure the short-run elasticities.

TINBERGEN thought it probable that the long-run elasticities were much higher than the short-run elasticities for a number of reasons mentioned by him, and made an attempt to measure a long-run import elasticity. It is the method used by TINBERGEN in this attempt which interests us now, and which will be compared with the method described in the previous sections.

TINBERGEN estimated the long-term substitution-elasticity between British and American pig iron in the United States of America during the period 1879–1914.

As variables were introduced:

Q_t — quantity of British pig iron imported into the U.S.A. as a percentage of total American consumption.

P_t — import-price (including duties) as a percentage of home price in the U.S.A.

¹⁾ *Metroeconomica* I, Dec. 1949, reprinted in *Econometrics*, p. 231 ff.

An attempt was made to explain the fluctuations in $\log Q_t$ by means of the fluctuations in $\log P_t$ with a distributed lag. For the distributed lag the following variables were introduced: $\log P_t$, $\log P_{t-1}$, $\log P_{t-2}$, $\log \bar{P}_{t-3/6}$ and $\log \bar{P}_{t-7/10}$, where $\bar{P}_{t-3/6}$ indicated the average price for the years $t-3$ to $t-6$ inclusive and $\bar{P}_{t-7/10}$ the average price for the years $t-7$ to $t-10$ inclusive.

By the usual method of multiple correlation the regression-coefficients of the explanatory variables were estimated. The result was, writing q_t for $\log Q_t$, p_t for $\log P_t$, etc.:

$$(II.7.1) \quad q_t = -3,1p_t - 0,2p_{t-1} - 0,1p_{t-2} - 0,4\bar{p}_{t-3/6} - 0,7\bar{p}_{t-7/10}.$$

The distributed lag in (II.7.1) represents the fact that a change in p has an effect on q that is spread over a number of years. The total effect of a change in p , e.g. a fall of the price of British pig iron, on q is not immediately realised, due to temporary hindrances; psychological inertia has to be overcome, current contracts must expire before new contracts can be made, one is prepared to make the expenses involved in establishing commercial relations with British exporters only when the fall in price has continued for a considerable time, etc.

It is to be expected that these hindrances to an immediate adaptation to changed data will gradually be overcome as time proceeds, the influence of the fall in p tapering off until the full effect on q has been realised.

As can be seen from (II.7.1) TINBERGEN's coefficients are not gradually decreasing in absolute value when the lag increases; on the contrary, from a lag of two years onwards the general tendency of the coefficients is to rise in absolute value with increasing lags (cf. graph 8 on p. 29). This does not seem very plausible. And why stop the calculation of elasticities at p_{t-10} when the coefficients are still rising with increasing lags?

One could imagine some cases where the influence of p_{t-10}

being higher than that of say p_{t-2} might be realistic. If British pig iron were used for the production of specific goods—American pig iron being a very poor substitute for British pig iron—and if these specific goods were to have an average life of about ten years, then the price-ratio of about ten years ago might have a greater influence on imports—via replacement pressure—than the price-ratio of say two years ago. For such a non-specific raw material as pig iron, however, this construction would be highly artificial.

Let us now see what our method comes to. The assumption is essential that, from a certain point, the coefficients are decreasing in a constant proportion with increasing lags. For instance

$$(II.7.2) \quad q_t = \alpha_0 p_t + \alpha_1 p_{t-1} + \lambda \alpha_1 p_{t-2} + \lambda^2 \alpha_1 p_{t-3} \dots \quad 0 \leq \lambda < 1$$

from which the following equation can be derived, as shown in the previous section,

$$(II.7.3) \quad \Delta q_{t-1} = \varepsilon_0 p_t + \varepsilon_1 p_{t-1} + \varepsilon_2 q_{t-1}$$

where the ε 's are expressions in α_0 , α_1 and λ .

When the approximation of the series of coefficients is introduced from the second coefficient onwards the resulting equation for q is even more simple, at the cost of higher abstraction of course. It is easily found that in this case:

$$(II.7.4) \quad \Delta q_{t-1} = \xi_0 p_t + \xi_1 q_{t-1}.$$

When the short-term reaction is defined as the reaction during the first year, the short-term elasticity that follows from (II.7.3) or (II.7.4) is resp. ε_0 and ξ_0 . The long-term elasticities can be calculated from (II.7.3) or (II.7.4), they are resp.

$$\frac{-(\varepsilon_0 + \varepsilon_1)}{\varepsilon_2} \text{ and } \frac{-\xi_0}{\xi_1}.$$

Using the same statistical material as TINBERGEN did, we estimated the coefficients of (II.7.3) and (II.7.4).

First the coefficients of equation (II.7.4) were computed:

$$(II.7.5) \quad \Delta q_{t-1} = -1,91 p_t - 0,54 q_{t-1} + 4,21.$$

The correlation-coefficient for (II.7.5) proved to be $R = 0,82$. In order to test the assumption that the residuals are non-autocorrelated, an assumption on which the method of estimation is based, the VON NEUMANN ratio was calculated. This ratio proved to be $v = 1,76$. For a random series the expected value of v is given by $E(v) = 2,06$ and the standard deviation is given by $S(v) = 0,33$. Apparently the data do not contradict the hypothesis that the successive values of the residuals are independent at the 5 % significance level¹⁾.

Using equation (II.7.3) instead of (II.7.4), we found:

$$(II.7.6) \quad \Delta q_{t-1} = -2,15 p_t + 0,93 p_{t-1} - 0,36 q_{t-1} + 2,69,$$

the correlation-coefficient proved to be $R = 0,85$, the VON NEUMANN ratio is $v = 2,47$ whilst again $E(v) = 2,06$ and $S(v) = 0,33$. In this case too the assumption that the residuals are non-autocorrelated cannot be rejected at the 5 % level of significance.

From (II.7.5) as well as from (II.7.6) the series of coefficients in the corresponding equation for q_t can easily be derived. From (II.7.5) it follows that:

$$(II.7.7) \quad \left\{ \begin{aligned} q_t = & 7,74 - 1,91 p_t - 0,87 p_{t-1} - 0,40 p_{t-2} - 0,18 p_{t-3} - \\ & - 0,08 p_{t-4} \dots \end{aligned} \right.$$

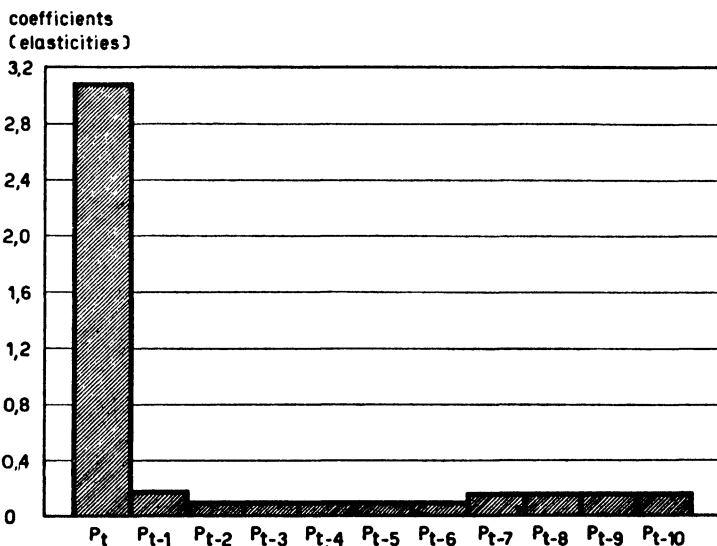
In the same way it is found from (II.7.6) that:

$$(II.7.8) \quad \left\{ \begin{aligned} q_t = & 7,39 - 2,15 p_t - 0,43 p_{t-1} - 0,28 p_{t-2} - 0,18 p_{t-3} - \\ & - 0,11 p_{t-4} \dots \end{aligned} \right.$$

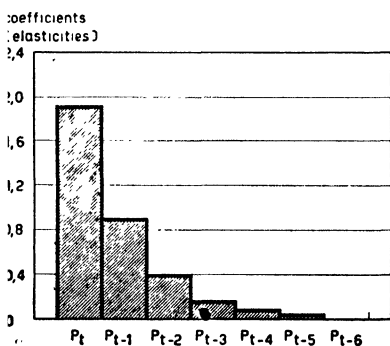
It is interesting to compare the time-shapes of the reactions, as represented by the series of coefficients in (II.7.7), (II.7.8) and TINBERGEN's equation. In order to do so a graphical representation of the coefficients (elasticities) of p_t , p_{t-1} , p_{t-2} , etc., is given in the graphs 8, 9 and 10.

¹⁾ See e.g. G. TINTNER, *Econometrics*, New York 1952, pp. 252 ff.

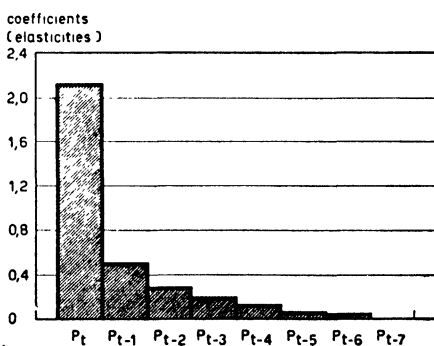
The difference between TINBERGEN's approach and ours is clearly illustrated. In our approach the influence of the price-ratio p is gradually approaching zero as p is more remote, while it follows from TINBERGEN's study that e.g. the elasticity of the price-ratio of six years ago is as high as that of two



Graph 8. Time-shape of the reaction according to Tinbergen's equation



Graph 9. Time-shape of the reaction according to our equation (II.7.7.)



Graph 10. Time-shape of the reaction according to our equation (II.7.8.)

years ago and that the elasticity of the price-ratio of even ten years ago is nearly twice as large as that of two years ago. As said before, it seems to us that our approach is more realistic in this respect; moreover it gives a great saving in computational work as a reward.

II.8. IRVING FISHER'S SHORT-CUT METHOD FOR CALCULATING DISTRIBUTED LAGS

After the development of our ideas about distributed lags and speed of adjustment, Prof. VERDOORN drew our attention to a short-cut method for calculating distributed lags published by IRVING FISHER ¹⁾.

FISHER points out that "any cause (such, for instance, as price change) produces a supposed effect (on, say, the volume of trade) only after some lag in time; and that, moreover, this effect is not all felt at once, at a single point of time, but is distributed over a number of points of time". The best distribution, FISHER goes on, presumably has the general form of some type of probability curve. The working assumption is made that, immediately following the cause, the effect is very small, but that it rises fast, reaching its mode after a very short interval and then tapering off slowly. "If time itself is plotted, the probability curve is evidently very skew. That is, the type is logarithmically normal".

However, the amount of computational labour required by the "logarithmically normal distribution method" is very large, and therefore FISHER advances a short-cut method which, according to FISHER, saves some nine-tenth of the labour.

By this short-cut method the influence of any given variable x on y is assumed to be the greatest at the very next time unit, and then to taper off by equal decrements for each successive time unit. "That is, the distribution curve becomes

¹⁾ I. FISHER, Note on a short-cut method for calculating distributed lags, *Bulletin de l'Institut International de Statistique*, La Haye 1937.

a straight line, beginning one time unit after the cause". In successive trials the best linear distribution curve (in our terminology time-shape curve) is estimated, the "best" curve being considered by FISHER as that which gives maximum correlation.

To be explicit, FISHER tries consecutively:

$$(II.8.1) \quad y_t = \alpha (3x_t + 2x_{t-1} + x_{t-2})$$

$$(II.8.2) \quad y_t = \alpha (4x_t + 3x_{t-1} + 2x_{t-2} + x_{t-3})$$

$$(II.8.3) \quad y_t = \alpha (5x_t + 4x_{t-1} + 3x_{t-2} + 2x_{t-3} + x_{t-4}) \text{ etc.}$$

Usually four or five trials, FISHER says, will suffice for practical purposes. Then the trial giving the highest correlation is chosen.

It is of some interest to compare FISHER's method with the one suggested in this study. This was done by BERGER in a paper published in *Metroeconomica*, August 1953¹⁾.

BERGER tried both methods in an explanation of the British part of Indian sheet and plate glass imports. The ratio of British imports to total imports in India was denoted by Q , the ratio of British glass prices to competing glass prices by P and their logarithms by q and p respectively.

Using our method one calculation was made, giving:

$$(II.8.4) \quad \Delta q_{t-1} = 2,92 - 1,03 p_t - 0,46 q_{t-1}$$

from which a long-run elasticity of $\varepsilon_L = -1,03/0,46 = -2,24$ was derived.

For FISHER's methods four trials were made in explaining q_t , viz. by:

$$(II.8.5) \quad 3p_t + 2p_{t-1} + p_{t-2} \quad \varepsilon_L = -2,16 \quad R = 0,858$$

$$(II.8.6) \quad 4p_t + 3p_{t-1} + 2p_{t-2} + p_{t-3} \quad \varepsilon_L = -2,24 \quad R = 0,881$$

$$(II.8.7) \quad 5p_t + 4p_{t-1} + 3p_{t-2} + 2p_{t-3} + p_{t-4} \quad \varepsilon_L = -2,38 \quad R = 0,836$$

$$(II.8.8) \quad 6p_t + 5p_{t-1} + 4p_{t-2} + 3p_{t-3} + \\ + 2p_{t-4} + p_{t-5} \quad \varepsilon_L = -2,48 \quad R = 0,751.$$

¹⁾ J. BERGER, On Koyck's and Fisher's methods for calculating distributed lags, *Metroeconomica*, August 1953.

The complete second equation, giving the highest correlation, is:

$$(II.8.9) \quad q_t = 6,37 - 2,24 \cdot \frac{4p_t + 3p_{t-1} + 2p_{t-2} + p_{t-3}}{10} \quad \varepsilon_L = -2,24.$$

A comparison of the final results, the reaction coefficients of the present and lagged price-ratio's p , gives:

	p_t	p_{t-1}	p_{t-2}	p_{t-3}	p_{t-4}
FISHER's approach	-0,90	-0,67	-0,45	-0,22	0
Our approach	-1,03	-0,56	-0,30	-0,16	-0,09.

The methods did not yield appreciably different results, and BERGER showed that generally "it is plausible that the sum of the coefficients—the long-term elasticity—is approximately invariant against a replacement of FISHER's by KOYCK's method".

From a theoretical point of view we do not see reasons why the assumption of a linear decrease of lagged influences should be preferred to our assumption of proportionately decreasing influences. On the other hand, if FISHER's short-cut method saves about 90 % of labour, our method saves about another 50 or 60 % compared with FISHER's short-cut method.

II.9. INCONSISTENT ESTIMATES OF THE TIME-SHAPE, DEVELOPMENT OF A METHOD TO FIND CONSISTENT ESTIMATES

Though serious difficulties in the problem of distributed lags will often be overcome by our method, another difficulty of a statistical nature is involved. An exposition of this difficulty and a method to surmount it will be given in this section.

First the difficulty. Let us rewrite the fundamental system of equations for the simple case and introduce a disturbance u in the equation of which we wish to estimate the coefficients:

$$(II.9.1) \quad y_t = \alpha x_t + \alpha \lambda x_{t-1} + \dots + u_t$$

$$(II.9.2) \quad \lambda y_{t-1} = \alpha \lambda x_{t-1} + \dots + \lambda u_{t-1}$$

$$(II.9.3) \quad y_t = \alpha x_t + \lambda y_{t-1} + u_t - \lambda u_{t-1}.$$

Let a and l be estimates of α and λ which are found by minimizing $\sum (y_t - ax_t - ly_{t-1})^2$ for variations in a and l . Then a and l are determined by the normal equations:

$$(II.9.4) \quad \begin{cases} \sum y_t x_t = a \sum x_t^2 + l \sum y_{t-1} x_t \\ \sum y_t y_{t-1} = a \sum x_t y_{t-1} + l \sum y_{t-1}^2 \end{cases}$$

$$(II.9.5) \quad \begin{cases} \sum y_t x_t = a \sum x_t^2 + l \sum y_{t-1} x_t \\ \sum y_t y_{t-1} = a \sum x_t y_{t-1} + l \sum y_{t-1}^2 \end{cases}$$

The computation of a and l from these normal equations is quite easy, but it will be proved, however, that a and l are not consistent estimates of the true α and λ ¹⁾.

From (II.9.4) and (II.9.5) it follows that

$$(II.9.6) \quad \begin{cases} \text{plim } a \cdot E x_t^2 + \text{plim } l E x_t y_{t-1} = E y_t x_t \\ \text{plim } a \cdot E x_t y_{t-1} + \text{plim } l E y_{t-1}^2 = E y_t y_{t-1} \end{cases}$$

$$(II.9.7) \quad \begin{cases} \text{plim } a \cdot E x_t^2 + \text{plim } l E x_t y_{t-1} = E y_t x_t \\ \text{plim } a \cdot E x_t y_{t-1} + \text{plim } l E y_{t-1}^2 = E y_t y_{t-1} \end{cases}$$

For $\text{plim } l$ e.g. the following solution results:

$$(II.9.8) \quad \text{plim } l = \frac{\begin{vmatrix} E x_t^2 & E y_t x_t \\ E x_t y_{t-1} & E y_t y_{t-1} \end{vmatrix}}{\begin{vmatrix} E x_t^2 & E x_t y_{t-1} \\ E x_t y_{t-1} & E y_{t-1}^2 \end{vmatrix}}.$$

For l to be a consistent estimate of λ , $\text{plim } l$ must be equal to λ . Let us now investigate the relation between $\text{plim } l$ and λ .

From (II.9.3) it follows:

$$(II.9.9) \quad \begin{cases} \alpha E x_t^2 + \lambda E y_{t-1} x_t + E u_t x_t - \lambda E u_{t-1} x_t = E y_t x_t \\ \alpha E x_t y_{t-1} + \lambda E y_{t-1}^2 + E u_t y_{t-1} - \lambda E u_{t-1} y_{t-1} = E y_t y_{t-1} \end{cases}$$

$$(II.9.10) \quad \begin{cases} \alpha E x_t^2 + \lambda E y_{t-1} x_t + E u_t x_t - \lambda E u_{t-1} x_t = E y_t x_t \\ \alpha E x_t y_{t-1} + \lambda E y_{t-1}^2 + E u_t y_{t-1} - \lambda E u_{t-1} y_{t-1} = E y_t y_{t-1} \end{cases}$$

The disturbances u_t as well as u_{t-1} are, by hypothesis, uncorrelated with $x_t, x_{t-1}, x_{t-2}, \dots$, hence $E u_t x_t = 0$ and $E u_{t-1} x_t = 0$. This can be substituted in (II.9.9).

From (II.9.1) it follows further that

$$(II.9.11) \quad E u_t y_{t-1} = \alpha E u_t x_{t-1} + \alpha \lambda E u_t x_{t-2} + \dots + E u_t u_{t-1} = E u_t u_{t-1}$$

$$(II.9.12) \quad E u_{t-1} y_{t-1} = \alpha E u_{t-1} x_{t-1} + \alpha \lambda E u_{t-1} x_{t-2} + \dots + E u_{t-1}^2 = E u_{t-1}^2.$$

¹⁾ An estimate $\hat{\theta}$ of a parameter θ is said to be consistent if $P(\hat{\theta} \rightarrow \theta) \rightarrow 1$ as $n \rightarrow \infty$ (or $\text{plim } \hat{\theta} = \theta$) where n is the size of the sample; "plim" means "the probability limit of".

This can be substituted in (II.9.10). This gives:

$$\begin{aligned} \text{(II.9.13)} \quad & \left\{ \begin{array}{l} \alpha E x_t^2 + \lambda E y_{t-1} x_t \\ \alpha E y_{t-1} x_t + \lambda E y_{t-1}^2 + E u_t u_{t-1} - \lambda E u_{t-1}^2 \end{array} \right. = E y_t x_t \\ \text{(II.9.14)} \quad & \left\{ \begin{array}{l} \alpha E x_t^2 + \lambda E y_{t-1} x_t \\ \alpha E y_{t-1} x_t + \lambda E y_{t-1}^2 + E u_t u_{t-1} - \lambda E u_{t-1}^2 \end{array} \right. = E y_t y_{t-1}. \end{aligned}$$

From these equations it follows:

$$\text{(II.9.15)} \quad \lambda = \frac{\begin{vmatrix} E x_t^2 & E y_t x_t \\ E x_t y_{t-1} & E y_t y_{t-1} - E u_t u_{t-1} + \lambda E u_{t-1}^2 \end{vmatrix}}{\begin{vmatrix} E x_t^2 & E x_t y_{t-1} \\ E x_t y_{t-1} & E y_{t-1}^2 \end{vmatrix}}.$$

Comparing plim l as given by (II.9.8) with this expression for λ , it will be clear that generally plim $l \neq \lambda$. Only in the special case when $E u_t u_{t-1} = \lambda E u_{t-1}^2$ it follows that plim $l = \lambda$.

Now it is reasonable to suppose in an economic theory as given by (II.9.1) that the disturbances u will be autocorrelated. Let us suppose that u_t is generated by an auto-regressive scheme

$$\text{(II.9.16)} \quad u_t = \xi u_{t-1} + \varepsilon_t$$

where u has a zero mean and a variance which is constant over time: $E u_t = 0$ and $E u_t^2 = E u_{t-1}^2$. Further, ε_t is uncorrelated with u_{t-1} and there is no autocorrelation between the ε 's, or $E \varepsilon_t u_{t-1} = 0$ and $E \varepsilon_t \varepsilon_{t-1} = 0$. Finally, it is assumed that $0 \leq \xi < 1$. That $\xi < 1$ follows already from the assumptions made earlier, this can be proved as follows. From (II.9.16) we have $E u_t^2 = \xi^2 E u_{t-1}^2 + E \varepsilon_t^2 + 2\xi E u_{t-1} \varepsilon_t$, substituting $E u_t^2 = E u_{t-1}^2$ and $E u_{t-1} \varepsilon_t = 0$ this becomes: $(1 - \xi^2) E u_t^2 = E \varepsilon_t^2$ from which follows $1 - \xi^2 > 0$ or $\xi < 1$.

Moreover $0 \leq \xi$ implies the reasonable assumption that economic time series do not have a component which has a period of two years.

The scheme (II.9.16) is not quite an arbitrary hypothesis. There is empirical evidence that (II.9.16) is not contradictory to quite a large body of economic data.

From (II.9.16) it follows:

$$(II.9.17) \quad Eu_t u_{t-1} = \xi Eu_{t-1}^2.$$

An expression for $\text{plim } (\alpha - a)$ and $\text{plim } (\lambda - l)$ can easily be derived by subtracting (II.9.13) and (II.9.14) from (II.9.6) and (II.9.7) respectively. This gives:

$$(II.9.18) \quad \left\{ \begin{array}{l} \text{plim}(\alpha - a) \cdot Ex_t^2 + \text{plim}(\lambda - l) \cdot Ex_t y_{t-1} = 0 \\ \text{plim}(\alpha - a) \cdot Ex_t y_{t-1} + \text{plim}(\lambda - l) \cdot Ey_{t-1}^2 = (\lambda - \xi) Eu_{t-1}^2. \end{array} \right.$$

$$(II.9.19) \quad \left\{ \begin{array}{l} \text{plim}(\alpha - a) \cdot Ex_t^2 + \text{plim}(\lambda - l) \cdot Ex_t y_{t-1} = 0 \\ \text{plim}(\alpha - a) \cdot Ex_t y_{t-1} + \text{plim}(\lambda - l) \cdot Ey_{t-1}^2 = (\lambda - \xi) Eu_{t-1}^2. \end{array} \right.$$

The difficulty, alluded to at the beginning of this section, can now be formulated as follows. We wish to estimate the true coefficients α and λ . By applying the least-squares method of estimation to the observed values y and x in the usual way the "direct" estimates a and l can be calculated from (II.9.4) and (II.9.5). But unless $\xi = \lambda$ these are not consistent estimates of α and λ .

A method to arrive at consistent estimates of α and λ , or at least at an interval within which these estimates must lie, will now be developed.

Let z_t be defined as:

$$(II.9.20) \quad z_t = y_t - ax_t - ly_{t-1},$$

i.e. z are residuals which can be calculated from the observed values of x and y , when a and l are computed from the normal equations (II.9.4) and (II.9.5).

We then have:

$$(II.9.3) \quad u_t - \lambda u_{t-1} = y_t - \alpha x_t - \lambda y_{t-1}$$

$$(II.9.20) \quad z_t = y_t - ax_t - ly_{t-1}$$

$$(II.9.21) \quad u_t - \lambda u_{t-1} = z_t - (\alpha - a)x_t - (\lambda - l)y_{t-1}$$

$$(II.9.22) \quad u_t^2 + \lambda^2 u_{t-1}^2 - 2\lambda u_t u_{t-1} = z_t^2 - 2(\alpha - a)x_t z_t - 2(\lambda - l)y_{t-1} z_t + (\alpha - a)^2 x_t^2 + 2(\alpha - a)(\lambda - l)x_t y_{t-1} + (\lambda - l)^2 y_{t-1}^2$$

or

$$(II.9.23) \quad \left\{ \begin{array}{l} u_t^2 + \lambda^2 u_{t-1}^2 - 2\lambda u_t u_{t-1} = z_t^2 - 2(\alpha - a)x_t z_t - 2(\lambda - l)y_{t-1} z_t + \\ + (\alpha - a)\{(\alpha - a)x_t^2 + (\lambda - l)x_t y_{t-1}\} + \\ + (\lambda - l)\{(\alpha - a)x_t y_{t-1} + (\lambda - l)y_{t-1}^2\}. \end{array} \right.$$

Let L denote the left hand side of this equation and R the right hand side.

First $E(L)$ is considered; as $Eu_t^2 = Eu_{t-1}^2$ and $Eu_t u_{t-1} = \xi Eu_{t-1}^2$ this is equal to $(1 + \lambda^2 - 2\xi\lambda)Eu_{t-1}^2$. Now $E(R)$; it may be assumed that $Ex_t z_t = 0$ and $Ey_{t-1} z_t = 0$ (because the correlation between explanatory variables and residuals in (II.9.20) is zero).

For the expressions between $\{ \}$ we get:

$$\begin{aligned} & \text{plim } (\alpha - a)Ex_t^2 + \text{plim } (\lambda - l)Ex_t y_{t-1} \text{ and} \\ & \text{plim } (\alpha - a)Ex_t y_{t-1} + \text{plim } (\lambda - l)Ey_{t-1}^2 \end{aligned}$$

respectively. Using (II.9.18) and (II.9.19) this is equal to zero and $(\lambda - \xi)Eu_{t-1}^2$ respectively. So we have for $E(R)$ the following expression:

$$Ez_t^2 + \text{plim } (\lambda - l) \cdot (\lambda - \xi)Eu_{t-1}^2.$$

Combining the results for the left hand side and the right hand side, we get:

$$(II.9.24) \quad (1 + \lambda^2 - 2\xi\lambda)Eu_{t-1}^2 = Ez_t^2 + \text{plim } (\lambda - l) \cdot (\lambda - \xi)Eu_{t-1}^2$$

or:

$$(II.9.25) \quad Eu_{t-1}^2 = \frac{Ez_t^2}{(1 + \lambda^2 - 2\xi\lambda) - (\lambda - \xi)\text{plim}(\lambda - l)}$$

or

$$(II.9.26) \quad Eu_{t-1}^2 = \frac{Ez_t^2}{1 - \xi\lambda + (\lambda - \xi)\text{plim}l}.$$

This expression for Eu_{t-1}^2 is now substituted in (II.9.14). For (II.9.13) and (II.9.14) we then have:

$$(II.9.27) \quad \left\{ \begin{aligned} \alpha Ex_t^2 + \lambda Ey_{t-1} x_t &= Ey_t x_t \end{aligned} \right.$$

$$(II.9.28) \quad \left\{ \begin{aligned} \alpha Ex_t y_{t-1} + \lambda Ey_{t-1}^2 &= Ey_t y_{t-1} + \frac{(\lambda - \xi)Ez_t^2}{1 - \xi\lambda + (\lambda - \xi)\text{plim}l}. \end{aligned} \right.$$

Now we introduce the estimates $\bar{\alpha}$ and $\bar{\lambda}$ as defined by the following equations:

$$(II.9.29) \quad \left\{ \begin{aligned} \bar{\alpha} \sum x_t^2 + \bar{\lambda} \sum y_{t-1} x_t &= \sum y_t x_t. \end{aligned} \right.$$

$$(II.9.30) \quad \left\{ \begin{aligned} \bar{\alpha} \sum x_t y_{t-1} + \bar{\lambda} \sum y_{t-1}^2 &= \sum y_t y_{t-1} + \frac{(\bar{\lambda} - \xi) \sum z_t^2}{1 - \xi\bar{\lambda} + \bar{\lambda}(\bar{\lambda} - \xi)}. \end{aligned} \right.$$

If ξ were known these equations would make it possible to calculate $\bar{\alpha}$ and $\bar{\lambda}$ which—as is easily seen when (II.9.29) and (II.9.30) are compared with (II.9.27) and (II.9.28) above—are consistent estimates of α and λ . For ξ , though unknown, only a restricted interval is admitted by hypothesis, viz. $0 \leq \xi < 1$.

With the interval allowed to ξ an interval for $\bar{\alpha}$ and one for $\bar{\lambda}$ will correspond. The smaller these intervals for $\bar{\alpha}$ and $\bar{\lambda}$ the less serious will be the fact that we do not know ξ exactly.

As can be seen in the chapter on empirical applications this will often be the case. It can be shown that this will generally be so when the multiple correlation coefficient is high. In order to prove this let us subtract (II.9.4) and (II.9.5) from (II.9.29) and (II.9.30). This gives:

$$(II.9.31) \quad (\bar{\alpha} - a) \sum x_i^2 + (\bar{\lambda} - l) \sum y_{t-1} x_t = 0$$

$$(II.9.32) \quad (\bar{\alpha} - a) \sum x_t y_{t-1} + (\bar{\lambda} - l) \sum y_{t-1}^2 = \frac{(\bar{\lambda} - \xi) \sum z_t^2}{1 - \xi \bar{\lambda} + l(\bar{\lambda} - \xi)}$$

from which $\bar{\lambda} - l$ is found to be:

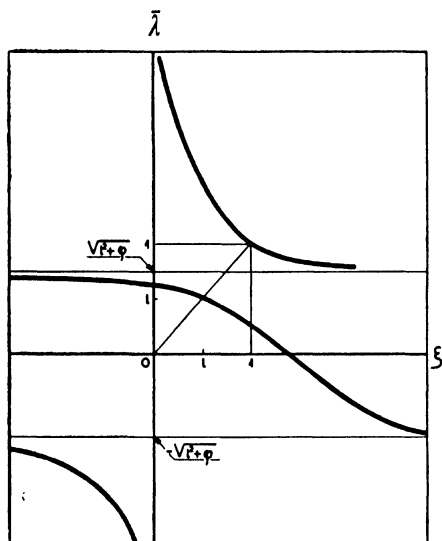
$$(II.9.33) \quad \left\{ \begin{aligned} \bar{\lambda} - l &= \frac{\bar{\lambda} - \xi}{1 - \xi \bar{\lambda} + l(\bar{\lambda} - \xi)} \cdot \frac{\sum z_t^2}{\sum y_{t-1}^2 - \frac{(\sum x_t y_{t-1})^2}{\sum x_t^2}} = \\ &= \frac{\bar{\lambda} - \xi}{1 - \xi \bar{\lambda} + l(\bar{\lambda} - \xi)} \cdot \frac{\sum y_t^2}{\sum y_{t-1}^2} \cdot \frac{\frac{\sum z_t^2}{\sum y_t^2}}{1 - \frac{(\sum x_t y_{t-1})^2}{\sum x_t^2 \cdot \sum y_{t-1}^2}}. \end{aligned} \right.$$

When $(\sum y_t^2)/(\sum y_{t-1}^2)$ is approximately equal to 1 this can also be written as:

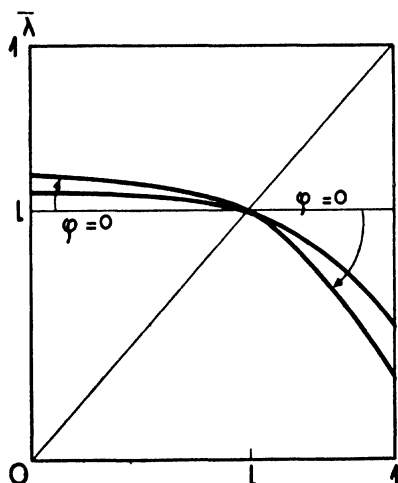
$$(II.9.34) \quad \bar{\lambda} - l \sim \frac{\bar{\lambda} - \xi}{1 - \xi \bar{\lambda} + l(\bar{\lambda} - \xi)} \cdot \frac{1 - R_{1,23}^2}{1 - r_{23}^2},$$

where $R_{1,23}$ is the multiple correlation coefficient of $y_t^* = ax_t + ly_{t-1}$ and r_{23} the correlation between x_t and y_{t-1} . When $(1 - R_{1,23}^2)/(1 - r_{23}^2)$ is near zero it is evident that $\bar{\lambda}$ is approximately independent of ξ and also approximately

equal to the direct estimate 1. An algebraic analysis of the relation between $\bar{\lambda}$ and ξ as given by (II.9.33) was made. The results are presented in graph 11. For the relevant intervals $0 \leq \xi < 1$ and $0 \leq \bar{\lambda} < 1$ the relation between $\bar{\lambda}$ and ξ for alternative values of $(\sum y_t^2)/(\sum y_{t-1}^2) \cdot (1 - R_{1,23}^2)/(1 - r_{23}^2)$, which is denoted by φ , is given in graph 12. When $\varphi = 0$, $\bar{\lambda} = l$ for all values of ξ ; when φ increases the $\bar{\lambda}$ - ξ -curve changes as indicated by the arrows in graph 12.



Graph 11



Graph 12

Let us now summarise the scheme of computation.

1. Compute the elements of the following matrix.

$$\begin{bmatrix} \sum x_t^2 & \sum x_t y_{t-1} & \sum x_t y_t \\ \cdot & \sum y_{t-1}^2 & \sum y_t y_{t-1} \\ \cdot & \cdot & \sum y_t^2 \end{bmatrix}$$

Then the "direct" coefficients a and l are calculated, further $\sum z_t^2$ is also calculated.

2. For alternative values of ξ between zero and one the corresponding values of $\bar{\lambda}$ and $\bar{\alpha}$ are calculated from (II.9.31) and (II.9.32).

The same line of reasoning can be followed when a more general hypothesis on the time-shape of the reaction is made in the original equation (II.9.1), e.g. when the geometric series for the reaction coefficients is introduced from the second coefficient onwards:

$$y_t = \alpha x_t + \beta x_{t-1} + \lambda \beta x_{t-2} + \lambda^2 \beta x_{t-3} \dots + u_t.$$

In Ch. IV a number of applications for this case will be given. It can be seen there, that the method of estimation developed in this section makes it possible to estimate $\bar{\lambda}$, $\bar{\alpha}$ and $\bar{\beta}$ within narrow limits in most cases.

CHAPTER III

DISTRIBUTED LAGS IN THE ADJUSTMENT OF CAPACITY TO OUTPUT

III.1. CAPACITY-OUTPUT ADJUSTMENT; IMPORTANCE OF THE PROBLEM

In this chapter the ideas on distributed lags previously expounded will be applied to the adjustment of productive capacity to output. This is an important relation in modern dynamic theory, where cyclical and secular movements are explained mainly by the double interaction between investment and output: first the effect of investment on output through the multiplier and second the effect of output on investment through the acceleration principle in some models (e.g. FRISCH, HARROD, SAMUELSON, HICKS) or through associated changes in profits in other models (e.g. KALECKI, TINBERGEN, KALDOR)¹).

Changes in capacity form a considerable part of total investment, therefore the object of our study, the dependence of capacity on output, is closely related with the second relation just mentioned. It is a well-known fact that in this relation the speed of adjustment of capacity to output has an important bearing on the stability of the economic system described in those theories²). Accordingly, applying the ideas

¹) R. F. HARROD, *Economic Journal*, 1951, p. 261 "it suddenly struck me that the key to the trade-cycle problem lay in a proper understanding of the interaction of the multiplier and the accelerator". N. KALDOR, *Economic Journal*, 1951, p. 833, "Modern cycle theories all explain the cyclical movement as a consequence of the double inter-relation between investment and income".

²) J. R. HICKS, *A contribution to the theory of the trade cycle*,

on distributed lags and speed of adjustment, as developed in the previous chapter, to the reaction of capacity to output in certain industries seems worth while.

In economic literature various meanings are attached to the concept of capacity; some remarks about the sense attached to capacity in this chapter may therefore be useful.

The object of our study is the adjustment of the number of a certain kind of durable producer's goods to changes in the flow of output. For our purpose capacity should be an index registering changes in this stock of capital goods over time, i.e. it should be an index of the amount of a given kind of fixed equipment in physical terms.

1. When the production department of a plant consists of a number of identical machines total capacity for our purpose could be indicated by the number of available machines. But machines are practically never identical, differences in age will always be present. This difficulty of differences in age is not specific to the measurement of machines, it is equally present when the amount of labour performed in a given period is measured and the labour force is composed of men of different ages.

When an old machine produces the same quantity of output per hour as a new one, allowance being made for normal periodic interruptions such as cleaning, normal repairs, etc., they are of equal capacity for our purpose. Capacity is sometimes used in economic literature as referring to the stock of year-services a machine will be able to produce before it is scrapped; a new machine which is expected to have an economic life of say ten years, then having double the capacity of an older machine which is expected to have a life of only five years. It should be clear that we do not use capacity in this sense.

Oxford 1950, p. 37. "It will be my contention . . . that the main cause of fluctuations is to be found in the effect of changes in output (income) on investment".

Only when the kind of machinery in question is such that the output per machine-hour of old and new machines shows minor differences or when the age-distribution of the stock of machines is practically constant over time, can capacity be measured by counting the number of available machines.

2. Apart from differences in age machines usually also differ in size, so that output per machine-hour is not the same. In this case, capacity as an index of the amount of equipment can be measured by the flow of output which will result in a given period under closely specified, fixed conditions. Essentially this will mean the construction of a weighted average, a machine which gives an output per hour twice as big as that of another machine being weighted twice as large as the other one.

Machines are not self-operating. They require the help of labour and power, and further supplies of raw materials. Even when all machines of a plant are in operation and one type of product is produced (with a uniform quality of raw materials), output can be varied by altering the number of hours per day and days per week during which the plant is operated, and further to a certain extent by changes in the speed at which each machine is operated by labour. This may be called a change in the intensity of use of the available stock of machinery.

Changes in the speed at which machines are running, together with changes in the use of power, usually require changes in the intensity of labour. The intensity of use of the stock of equipment is therefore related, among other things, to labour as a factor of production cooperating with the equipment, and more especially to working hours and the intensity of labour.

When a fixed number of working hours and intensity of labour are selected as a base it might approximately be said that a fixed intensity of operation, say i_t , is chosen. Capacity, as an index of the amount of equipment, can then be measured

by the output of the plant achieved in a given period if all available machines were actually operated at this fixed intensity of operation, allowance being made for normal periodic interruptions in the use of machines.

It is not very important for our purpose which fixed level of i is chosen as a base, essential is that it remains fixed so that comparisons in the stock of the equipment over time are possible. Capacity in the sense of a *stock* of equipment is measured by a *flow* of output which results from it under fixed, specified conditions.

3. In economic theory another concept of capacity is often used. For a plant with a given number of available machines there is an optimum intensity of use $\bar{i}$ at which the costs per unit of output are at their minimum. Capacity is then measured by the output which would result if the plant were actually operated at this optimum intensity of operation $\bar{i}$. In order to distinguish between both concepts one might speak of technical capacity when a fixed i is used and of economic capacity when $\bar{i}$ is used in defining capacity.

This economic capacity is not constant for a given physical equipment, it may change e.g. with changes in the wage-system. In comparing changes in physical equipment over time this concept is inadequate. Fortunately existing statistics of capacity generally use a definition which makes capacity constant for a given physical equipment, so that capacity refers to technical capacity as indicated above.

When output is not homogeneous in character, technical capacity should be expressed in a fixed composition of output. In industries where significant changes in composition and quality of output are common, there are indeed great difficulties in applying a concept and analysis of capacity. In the empirical studies which follow, the output as well as the machinery of the industries investigated are highly standardised, so that no serious difficulties of capacity measurement are encountered.

III.2. IMPORTANCE OF CORRECT ECONOMIC ANALYSIS IN ECONOMETRIC RESEARCH

Our ultimate goal is to get some idea of the empirical adjustment path of capacity in some industries. But before starting statistical measurements it is necessary to have reached an economic analysis of what is going to be measured. The need and importance of correct economic analysis in econometric studies can hardly be overestimated.

In this chapter therefore a more systematic analysis of capacity and investment behaviour, with special stress on the time-shape thereof, will be given. We shall have to deal not only with the question how *much* capacity will be needed for equilibrium in a given situation, a question amply discussed in economic theory, but also with the much more neglected question how *fast* additional capacity will be built (or capacity will be scrapped). An attempt will be made to link this theory of investment-decisions with the general theory of costs and short-period and long-period equilibrium in production.

III.3. RESTRICTION OF THE PROBLEM: INVESTMENTS INDUCED BY DEMAND

The capacity-output relation has sometimes been introduced in economics as a purely technically fixed relation. It should be clear from the beginning however that the capacity-output ratio is not a technically fixed ratio, it is a variable in the investment behaviour of the entrepreneur. In most industries there is a certain degree of flexibility in plant and equipment, so that a change in output is possible without immediately changing capacity. The adjustment of capacity to output is a problem of choice in the behaviour of the entrepreneur.

All investments are not connected with the capacity-output relation, we shall restrict ourselves to investment in fixed capital. And even then every investment does not bring about a change in capacity; apart from investment in capacity

(capital widening) there may be investment in automatism and durability of fixed capital without increasing capacity.

Investment in capacity may be the consequence of a great variety of causes. A rough and not always sharp distinction is that between autonomous investment (a response to technological innovations: the building of new or the expansion of old plants in order to produce new products or to produce a product by means of a previously unknown method of production) on the one hand and investment induced by changes in demand, wages, rates of interest etc. on the other hand. Our analysis will not deal with autonomous investments in capacity.

Our problem has to do with the response of capacity to changes in output. These changes in output may be a consequence of a) changes in the demand for the firm's product: a shift of the firm's sales curve, and b) changes in the quantity demanded (moving along a given sales curve) due to changes in the price the entrepreneur sets on his products as a response to higher or lower prices of raw materials, wages etc. For the industries studied in the next chapter the first cause was much more important than the second in causing changes in sales and output. Investment in capacity was mainly investment induced by changes in demand.

III.4. THE ACCELERATION PRINCIPLE IRRELEVANT TO PROBLEMS ON THE SPEED OF CAPACITY ADJUSTMENT

Is not the theory of the acceleration principle ready for use in our problem? It certainly deals with the relation between capacity and output and already has a long history. The first development of this theory, it seems, can be attributed to T. N. CARVER¹⁾ and a more detailed analysis to AFTALION²⁾. The name acceleration was introduced by J. M. CLARK in a

¹⁾ T. N. CARVER, A suggestion for a theory of industrial depressions, *Quarterly Journal of Economics*, 1903.

²⁾ AFTALION, *Les crises périodiques de surproduction*, Paris 1913.

study "Business acceleration and the law of demand" ¹⁾). Since then it has gained a place in many theories of the business cycle e.g. in those of HANSEN, HABERLER, J. M. CLARK, FRISCH, and others.

More recently the acceleration principle has been combined with the multiplier principle of J. M. KEYNES, in order to give an explanation of business cycles and, to go a step further, to study the possibilities of a steady long-term economic development ²⁾).

The principle in its rigorous form starts from the assumption, that there is proportionality between the stock of investment goods and the flow of output over a certain period, or in other words between net investment and the change in output during a given period. This can be shown most appropriately by the nearly classical example given by HABERLER.

Suppose that the output of shoes in a certain year is five million pairs, requiring a thousand machines of a certain type. If in the two succeeding years respectively six and six and a half million pairs are required, then twelve hundred machines will be needed in the first of these years and thirteen hundred in the next. New investment of two hundred machines will then be required in the first and of one hundred machines in the second year, that is, in proportion to the increases of one million and a half million, respectively, in the number of pairs of shoes produced.

The acceleration principle assumes that there is an immediate and perfect adjustment of the stock of fixed capital goods to current output, in other words that total capacity is always optimal for the current output.

It was pointed out some fifteen years ago by TINBERGEN that there are boundary conditions for the validity of the acceleration principle. "Very strong decreases in consumers'

¹⁾ In the *Journal of Political Economy*, 1917.

²⁾ R. F. HARROD, *Towards a dynamic economics*, London 1948.

goods production (i.e. output) must not occur. If the principle were right, this would lead to a corresponding disinvestment and this can only take place to the extent of replacement”¹⁾.

Machines, once made, cannot be unmade at will. When output decreases, excess capacity may arise, an increase in demand later on and in output subsequently will not make an expansion of capacity necessary as long as there is excess capacity available.

At the other end there is also a limit for the working of the acceleration principle: positive net investment may be limited by the capacity of the industries producing the investment goods required for the expansion of capacity.

But even within these limits there is no reason to suppose an immediate adjustment of capacity to current output. This assumption is the major shortcoming of the acceleration principle in its simplest form in the explanation of short-term fluctuations in investments; in reality there is “sometimes too much and sometimes too little” capacity.

Although it may be admitted that under favourable conditions the acceleration principle can answer the question regarding how *much* additional capacity will be *required* by an increase in output, it does not answer the question how *fast* this capacity will be *built*. Hence the acceleration principle cannot do much to meet our requirements being an analysis not only of the volume of investment but also of the timing of investment in capacity.

“The theory of investment”, as KLEIN said²⁾ “must be treated as a part of the general theory of economic behaviour and must be very systematically developed rather than pulled out of the magician’s hat as the “acceleration principle” or some other catchword”.

¹⁾ J. TINBERGEN, Statistical evidence on the acceleration principle, *Economica*, 1938; reprinted in *Econometrics*, London 1951, pp. 215–219.

²⁾ L. R. KLEIN, Notes on the theory of investment, *Kyklos*, 1948, p. 97.

III.5. A MODEL ON ENTREPRENEURIAL DECISIONS CONCERNING CAPACITY AND ITS RATE OF UTILIZATION

The entrepreneurial decisions on the capacity of his plant are closely related with other decisions he has to make, or to use the language of economic theory, capacity will be simultaneously determined with other decision variables.

A simplified model will be introduced to analyse these decisions. The simplifications are chosen in such a way, that the most essential aspects of our problem come to the front, leaving aside, we hope, all that is not strictly essential.

The following notation will be used in the analysis:

- k_{-1} — the number of machines (capacity) inherited from the past period, i.e. available at the beginning of the period to be considered.
- g — the number of old machines (i.e. inherited from the past) that is to be used again in the current period.
- s — the number of old machines to be scrapped.
- r — the number of old machines to be idle during the current period; $r + s + g = k_{-1}$.
- c — the number of new machines to be purchased: new capacity.
- k — the number of machines available in the current period, after scrapping old and buying new machinery
 $k = k_{-1} + c - s = r + c + g$.
- $\Delta k = k - k_{-1}$ = change in capacity.
- i — hours per week active machines are operated (or more generally intensity of operation of active machines).
- u — output in the current period.

Following Marshallian lines two types of equilibria will be studied:

- a. We first assume that capacity is fixed: $c = s = 0$ so that also $\Delta k = 0$. What will be the equilibrium values of the decision variables g , r and i for given levels of output u ? Let this be called short-period equilibrium.

- b. Total capacity is variable, and fully adjusted to the level of output, there is no idle capacity: $r=0$. What will be the decision variables, especially c and g which determine the total capacity k , for given levels of output u ? Let this be called long-period equilibrium.

In order to answer these questions a number of assumptions have to be made. Some of these assumptions undoubtedly are drastic simplifications; they will be modified later on.

a) We suppose, to begin with, that only one product is made.

b) Machines are identical as far as their technical function in production is concerned, they may differ i.a. in age, size, automatism, repair and maintenance required. The stock of machines is measured in "capacity-units", i.e. according to their output per machine-hour if they are operated at a fixed intensity of operation i , and allowance is made for normal interruptions. This technical rate of output is supposed to remain constant during the effective life of the machines.

c) It is supposed that there are a large number of machines in the plant, so that the plant may approximately be considered as being perfectly divisible as far as machines are concerned.

d) The output of a machine in use during a given period can be varied by changes in the intensity of use i . Increasing this intensity of use beyond a certain point will give rise to a decreasing output per machine-hour and hence to increasing costs per unit of output as a consequence of falling productivity of labour. The increase in unit costs will be strengthened when, at a given intensity of operation, higher wage rates will have to be paid (overtime, night shifts, etc.).

e) The available machines can be ordered according to their variable costs of operation, repair and maintenance per active machine-hour; let these costs be called m . For a given machine m will, other things equal, increase with the aging of this machine; we suppose that the variable costs of repair

and maintenance for an active machine-hour in a future period is independent of the activity (or non-activity) of the machine in the current period. This assumption excludes some intricate problems of time-substitution of repair and maintenance.

f) New machines can be purchased at given prices, while sufficient money capital is available at a given rate of interest. There may be economic progress in the efficiency of machines, e.g. it may be that the new machines are now more durable or will need less repair and maintenance over their expected lifetime than was the case with new machines formerly. But there is no change in the technology of production and there are no net economies of large-scale production.

g) A final assumption has to be made of the entrepreneur's behaviour. It is supposed that he will try to produce a given output u with least total costs K .

Now it is possible to answer the questions put in the beginning of this section about the equilibrium values of the relevant decision variables.

III.6. SHORT-RUN ADJUSTMENTS TO CHANGES IN OUTPUT

In the Marshallian short-run analysis output can be varied

1. by changes in the intensity i with which a given number g of machines is operated, and
2. by changes in the number g of old machines in operation at a given intensity of operation i ; there is a ceiling for this change of course: $g \leq k_{-1}$ ¹⁾.

Increasing i gives an increase in output and in costs, let us call this increase in costs per unit of additional output short-run marginal intensive costs: K'_i . Likewise, increasing g gives increases in output and costs, let us call this increase in costs per unit of increased output short-run marginal extensive costs: K'_e .

¹⁾ Increasing g is equivalent to decreasing the number of idle machines r , when $g = k_{-1}$ all machines are in use: $r = 0$.

In mathematical short-hand: when $u = f(g, i)$ and $K = \varphi(g, i)$,

$$K'_i = \frac{\partial K / \partial i}{\partial u / \partial i} \quad \text{and} \quad K'_e = \frac{\partial K / \partial g}{\partial u / \partial g}.$$

For total costs K of a given output u to reach a minimum, the marginal intensive costs K'_i must be equal to the marginal extensive costs K'_e ; when this condition is fulfilled no substitution between i and g can lower costs. The equilibrium condition $K'_i = K'_e$ is indeed the well-known substitution principle applied to the model in question.

Mathematically: minimising K for a given value $u = u_0$, subject to the production function $u = f(g, i)$ as a side relation, gives:

$$\frac{\partial K}{\partial i} + \lambda \frac{\partial [f(g, i) - u_0]}{\partial i} = 0 \quad \text{or} \quad \frac{\partial K}{\partial i} + \lambda \frac{\partial u}{\partial i} = 0$$

$$\frac{\partial K}{\partial g} + \lambda \frac{\partial [f(g, i) - u_0]}{\partial g} = 0 \quad \text{or} \quad \frac{\partial K}{\partial g} + \lambda \frac{\partial u}{\partial g} = 0,$$

where λ is a multiplier of LAGRANGE. Eliminating λ from these equations gives:

$$\frac{\partial K / \partial i}{\partial u / \partial i} = \frac{\partial K / \partial g}{\partial u / \partial g} \quad \text{or} \quad K'_i = K'_e.$$

When $g = k_{-1}$ and $\partial u / \partial i = 0$ the maximum output of the given equipment is obtained. No increase in output can be obtained without increasing the number of machines.

Marginal intensive costs K'_i consist of costs of operating labour, repair and maintenance (raw material costs may be considered as constant per unit of output and hence can be omitted in the substitution analysis).

For low values of i K'_i may be constant; with increasing intensity of operation however, a point will be reached where the marginal productivity of labour begins to fall (due to stepping up the speed of operating labour, overtime work, night shifts etc.), so that K'_i rises. This rise may be accentuated by increasing costs of repair and maintenance per machine-

hour when i is high and machines are "overworked". Further at a certain level overtime-rates may have to be paid for labour, so that K'_i may jump up to a higher level (see graph 13).

Marginal extensive costs K'_g are the unit costs of the marginal machine to be used, they consist of the costs for the operating labour and repair and maintenance. When g is at a low level, only the newest and most efficient machines will be in use. By increasing g the unit costs K'_g will generally rise, the additional machines put in operation having higher unit costs of operation, repair and maintenance.

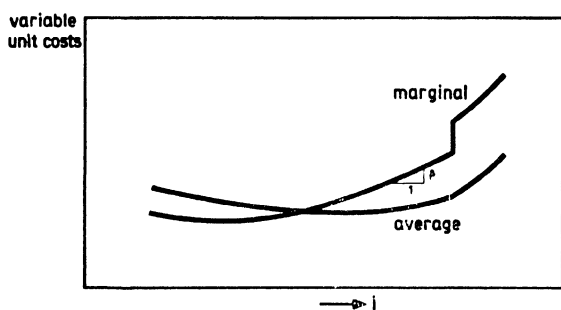
Let us now see what happens when output is varied, and i and g are so adjusted that the equilibrium condition $K'_i = K'_g$ remains fulfilled. This makes the short-run equilibrium values of i and g both functions of u . From $u = f(g, i)$, $g \leq k_{-1}$ and $K'_i = K'_g$ it follows that g and i are determined as functions of u : $g = g(u)$, $g \leq k_{-1}$ and $i = i(u)$.

There is a certain volume of output u for which g is just equal to k_{-1} , no machines being idle. Increases in output can then only be achieved by raising the intensity of operation. For lower outputs a balanced decrease of i and g exists which will maintain the least costs equilibrium. How is this adjustment to lower outputs going to be divided between i and g , i.e. between working less intensively (i.a. fewer hours) and using fewer machines?

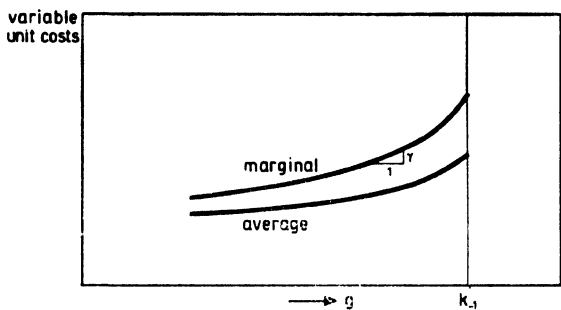
This will depend on the influence of i on marginal intensive costs, mainly determined by the productivity of labour, on the one hand and the influence of g on marginal extensive costs, mainly determined by the costs differentials among machines, on the other hand.

A general relation between variable costs and intensity of work is given in graph 13; the influence of the number of machines in use g on variable costs is represented in graph 14.

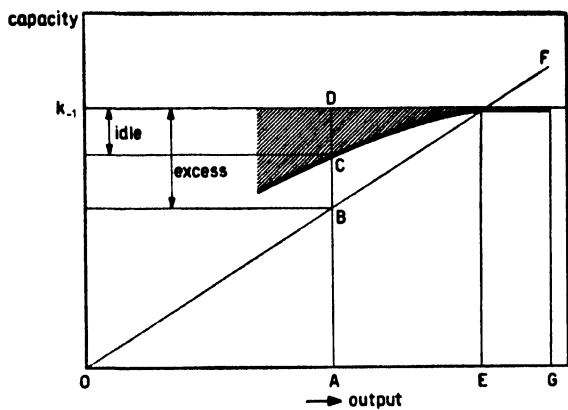
Now, the smaller the influence of i on the marginal intensive costs (i.e. the smaller the slope β in graph 13) and the greater the cost differentials among machines (γ in graph 14),



Graph 13



Graph 14



Graph 15

the more the weight of adjustment to lower outputs will fall on the intensity of operation i (hence the fewer machines will be left idle and the more short-time work will be introduced), in order to keep costs as low as possible, i.e. to maintain $K'_i = K'_c$. The same holds of course for adjustments to higher outputs as long as $g \leq k_{-1}$.

The relation between capacity operated, which is identical with g , and output is given in graph 15. Anticipating the next section, a long-run optimum intensity of operation i is supposed to exist, corresponding with a long-run optimum capacity-output ratio, represented by the slope of OF in graph 15.

An optimum output OE corresponds with the available capacity k_{-1} . For a lower output OA , active capacity (i.e. the number of machines in use) would be AB if i should remain at its long-run optimum level; we pointed out above, however, that generally the intensity of operation will be lowered when idle capacity arises, so that active capacity is higher than AB say: AC . Idle capacity, corresponding with the number of machines not in operation, is then equal to CD ; BC is that part of capacity which is in operation as a consequence of the lowered intensity of operation, e.g. due to short-time work. This is excess capacity, however, from a long-run equilibrium point of view, the total excess capacity being equal to BD .

Outputs higher than OE can only be obtained with the available capacity by increasing i , there is a maximum output OG when i is so high that the marginal productivity of labour is zero.

The difference between idle capacity and excess capacity should especially be kept in mind in empirical studies, when statistics refer to idle capacity. This is the case for example in the statistics on the cotton textile industry in the United States, where the number of active and idle spindles is recorded.

III.7. LONG-RUN ADJUSTMENTS TO CHANGES IN OUTPUT

In the previous section the possibility of purchasing new machines was excluded by the assumption that $c=0$. This assumption will be released now, hence an additional decision variable, c , has to be determined.

Output can now be varied not only by changes in i and g , but also by changes in c ($c \geq 0$). When the purchase of new machines is considered, the output expected for the current period has not only to be taken into account but, as a consequence of the durability of machines, also the output which is expected over a much longer period of time. This means that the entrepreneur's expectations about the future demand for his product (and also about future wage-rates, raw material prices etc.) play a much more essential role than in the previous section.

Let us first suppose that the elasticity of his expectations equals unity: when demand changes from say d_1 to d_2 the entrepreneur expects that it will remain constant at the new level (d_2 ¹). In this case the entrepreneur will generally be planning for a constant output stream, so that we can put off the study of a fluctuating output for the time being.

When purchasing new machines is considered, as a rule, the problem of choice between various degrees of automatism and durability of machines has to be solved. There will be a certain type of machine which gives lowest annual costs per unit of annual output of the new machine. Let these costs, which consist of depreciation, interest, costs of the operating labour, repair and maintenance, be denoted by K_n .

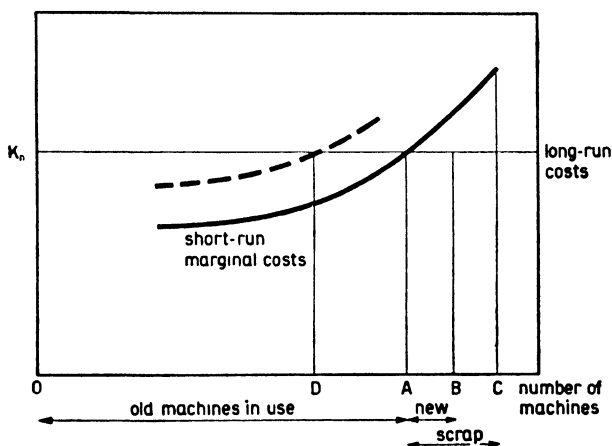
The buying of a new machine gives an increase in annual costs over a certain period and *ceteris paribus* an increase in the annual output over the same period. Let the increase in these costs per unit of increased output be called long-run marginal costs. When there are no net (dis)-economies of

¹) See J. R. Hicks, *Value and Capital*, Oxford 1939, p. 205.

large scale, these long-run marginal costs will be constant and equal to the average long-run costs K_n .

For least costs equilibrium these long-run marginal costs must be equal to short-run marginal costs (both intensive and extensive, but as these are equal in equilibrium we need not always add these adjectives).

In graph 16 the short-run marginal costs (assuming balanced changes in i and g so that $K'_i = K'_e$ always holds) is shown as a function of g (which of course, like i , is a function of u).



Graph 16

For low output levels long-run marginal costs (which, unlike short-run costs, include depreciation and interest) will be higher than short-run marginal costs. There is a certain level of output u_A however, where short-run marginal costs are just equal to long-run costs K_n .

This is the case in graph 16 when $g = OA$ old machines are used. For outputs above this level u_A , the equilibrium condition $K_n = K'_i = K'_e$ requires g to remain constant at OA , and i constant at an equilibrium level say i_0 , additional output will be produced most efficiently with new machines. For outputs above u_A changes in u will be proportional to changes in the

number of new machines c . An output u_B e.g. ($u_B > u_A$) will be produced with OB machines, of which OA are old and AB to be purchased.

When the number k_{-1} of machines inherited from the previous period is OC , it follows that AC is the number of these old machines which will be scrapped (when there is uncertainty about future demand it might be preferred to keep some old machines as a reserve) and BC the decrease $k_{-1} - k$ in the total stock of machines. For higher outputs than u_B the point B will be situated more to the right of course; the number AB of new machines purchased may be larger than the scrapping AC of old machines, an expansion in the stock of machines being the result.

For an output u_D below u_A no new machines will be purchased for the current period, because short-run marginal costs are lower than long-run costs. There is no possibility of long-period equilibrium in the current period. As time goes on however, the short-run costs will increase as a consequence of the aging of machines which, it is supposed, causes higher costs of repair and maintenance per unit of output¹⁾ (the point A in graph 16 moving to the left). Hence, after a certain lapse of time long-period equilibrium will be attained, represented by the dotted cost curve in graph 16, where the stock of machines is reduced by scrapping to OD which is again in full equilibrium with the output u_D . The capacity-output ratio has reached its long-term equilibrium, as this is determined by the optimum intensity of operation i_0 (for which $K'_i = K_n$).

The analysis, so far, is highly static because a perfectly elastic supply of capital funds was assumed, planning for a constant output stream, and no uncertainty about entrepreneurial expectations. It will be necessary to remove these assumptions in the next sections.

¹⁾ I.e. as time goes on the short-run marginal costs curve shifts upwards.

III.8. RISING SUPPLY CURVE OF CAPITAL FUNDS AND EXPANSION OF CAPACITY

Buying new machines means the investment of money capital. It was supposed hitherto that the supply of this capital to the firm was perfectly elastic, which implies that any firm can, at the given rate of interest, draw as much outside capital funds, in equity form or by borrowing, as the entrepreneur thinks advisable.

This assumption undoubtedly does great violence to facts. The capital market is an imperfect market, the fact that the amount of capital a firm obtains in a given period is limited otherwise than by the entrepreneur's considerations of profitability of investments alone, is normal for the greater part of the firms in reality, though perhaps less common among large than among small firms ¹⁾. The existence of these limits will have an important bearing on the adjustment of capacity to a higher long-period level.

In order to see what kind of consequences the imperfect supply of capital funds may have for our problem, it is advisable to start with a classification of possible sources of funds for a firm that wishes to buy new machines. The main sources are:

1. the internal funds of the firm;
2. outside funds to be obtained by borrowing;
3. outside funds to be obtained by equity participations.

The first source, internal capital funds, consists of the non-operating assets of the firm or the entrepreneur, in so far as these are available in a liquid or widely marketable form, e.g. time deposits and government bonds. These assets will accrue out of the depreciation reserves and the ploughing back of profits but may also be supplied in some cases out of the entrepreneur's private assets. However this may be, the

¹⁾ See e.g. A. G. HART, *Anticipations, uncertainty and dynamic planning*, New York 1951, p. 39 ff.

supply of internal funds has certain limits, becoming very inelastic beyond a certain point.

The second source, borrowing, also has its limitations. Professor HAYEK pointed out ¹⁾ that the assumption of "perfect competition" on the loan market is inappropriate, simply because successive (additional) loans to the same borrower will never present the "same commodity" in the sense in which the term is used in the theory of competition. This is due to the fact that in given circumstances the security a borrower has to offer is not as good for a large amount as for a small one. "In consequence, every prospective borrower will have to face an upward sloping supply curve of credit". In other words, a higher gearing of the capital structure is disliked by the creditors, since it increases the risk that the enterprise will not be able to meet their claims to interest and repayment. Moreover, when the entrepreneur or the equity holders are uncertain about the future profits, they also will dislike high gearing. As Dr KALECKI has shown ²⁾, there is an increasing subjective risk accompanying the increase of indebtedness relative to the firm's own capital, for the greater this ratio the greater is the reduction of the entrepreneur's or equity holders' income when the average rate of profit falls short of the rate of interest.

The third source, outside equity participations, often is not unlimited either. Ordinarily a limited amount of funds can be had from the relations of the entrepreneurs on terms which amount to inactive partnership. To go further with equity financing, stock must be publicly floated. There may be serious objections to this in the case of small firms, the family may be unwilling to weaken its hold on the management of the business, or the institutions of the capital market may imply the impossibility of floating stock. And when

¹⁾ F. A. VON HAYEK, The Ricardo Effect, *Economica*, 1942.

²⁾ M. KALECKI, *Essays in the theory of economic fluctuations*, London 1939.

stock can be sold considerations may occur limiting the amount of equity finance or at any rate making the terms of finance less favourable, the larger the amount required. "A given enterprise can usually get a larger amount of outside finance only by seeking subscriptions from capitalists who are less interested, or less well informed, or less trusting – and who consequently need to be tempted by better terms" ¹⁾).

Summarising it can be said that for many firms the supply curve of capital funds beyond a certain point is rising. Or rather, as was pointed out by Professor HART ²⁾ "the firm is confronted with a series of opportunities for raising funds, each with its own interest rate and its own rules, and each – ordinarily – having fairly definite quantitative limits. That is, the entrepreneur obtains his capital on a "stepped" market".

In times of high demand and output these limits may restrict the expansion of capacity. There are firms for which availability of capital funds is an important limitation of investment. These limits, however, are sliding limits as a rule. Corresponding with high demand and output there is the realisation of, sometimes high, net profits, part of which are ploughed back in course of time. This means that the supply curve of funds shifts to the right as long as retained net profits are accumulated, because this lowers the indebtedness relative to the firm's own capital. An increase in its own capital will therefore increase the funds it can raise at more or less the same cost of borrowing. Furthermore, the risk of indebtedness for the entrepreneurs or equity holders, will be reduced. For, as was shown by Dr KALECKI, this risk is an increasing function of the ratio of indebtedness to the firm's own capital. Also, the difficulties which some firms may have in the floating of stock will lessen as profits keep on flowing, the caution of sceptical capitalists being gradually

¹⁾ B. TEW, *The finance of investment, Oxford Economic Papers*, 1952.

²⁾ *Op.cit.* p. 44.

overcome. Thus, as long as the high level of sales and profits continues, the availability of investable funds rises with the passage of time.

Let us now see the consequences of all this for the capacity-output problem. The rising supply curve of funds makes long-run marginal unit costs K_n beyond a certain point, an increasing function of the number of new machines purchased in a given period. The rise in unit costs may be tempered to a certain degree, by purchasing cheaper e.g. less automatic and durable machines so that with a limited amount of funds more machines can be obtained. (This is of course HAYEK's well-known RICARDO effect).

When output rises to a level sufficiently high for the rising part of the long-run marginal cost curve to become effective, this "scarcity" of funds will reduce the purchase of new and also the scrapping of old machines. The least costs equilibrium $K_n = K'_i = K'_e$ entails an intensity of operation i which is higher and hence a capacity-output ratio which is lower than would have been the case if the supply of funds were perfectly elastic.

This is a temporary equilibrium however. When output remains at the high level and good profits continue, the factors which limited the purchase of new machines will become less stringent or completely ineffective after some time because of the ploughing back of profits and of the resulting more favourable terms at which outside finance can be obtained.

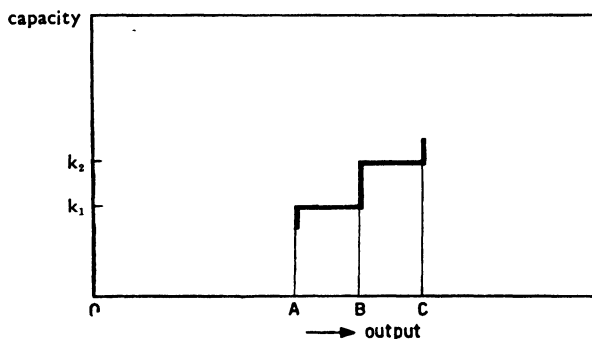
Additional new machines can be purchased then, partly in order to reduce the number of old machines g and the intensity of operation i , partly to increase total capacity until the long-period equilibrium between capacity and output is attained.

III.9. SECULAR GROWTH AND BUILDING AHEAD OF DEMAND

When we asked how a certain output u could be produced with minimum costs, it was hitherto assumed that the entrepreneur expected sales and output to remain stable at

this level, so that he was planning for a constant output stream. This assumption gave the analysis a highly static character and it must now be abandoned for a more realistic approach.

Let us first assume that the entrepreneur expects a regular increase in the demand for his product, so that he is planning for an output growing at a constant rate from the present level u_0 onwards.



Graph 17

When he is subjectively certain, i.e. when he has no doubts about this growth rate, this does not essentially change our conclusions. To maintain equilibrium the expansion of capacity must keep pace with the growth in output.

There are obvious reasons in many industries however, which make it impossible for the entrepreneur to expand his capacity step by step as his sales grow.

These reasons are the lumpiness and durability of machine units.

When equipment is not perfectly divisible, as was supposed up till now, the equality between marginal short-run and long-run costs, as developed in the previous sections, cannot always hold exactly. Let us see what this means in the case of the constant output stream first. The capacity-output relation becomes a stepped curve in this case, see graph 17. For outputs between OA and OB the variable costs of a more

intensive use of the capacity k_1 are lower than the unit costs corresponding with the increase of capacity with one unit to k_2 . When output should be at OB the expansion of capacity becomes just advantageous. From a technical point of view there is excess capacity then, but this is not unplanned excess capacity, it is equilibrium capacity from an economic point of view.

The phenomenon just discussed becomes much more important when there are "economies of large scale" and the entrepreneur plans for a gradually growing market. The economies of scale aimed at above are those of the large machine. It is a well-known fact that small machines often are much more expensive relative to their capacity than large ones, because an increase in the capacity of a machine requires a less than proportional increase in material and labour in the construction of it. In that case it will be cheaper to make a small number of large additions to capacity than a large number of small additions.

Let us suppose that an entrepreneur expects to expand the capacity of his plant say by 20 % in a certain future period as a consequence of expected growth in sales. When he has e.g. the choice between an increase of 20 % now and an increase of 10 % now and another 10 % afterwards, it may well be that the second alternative entails a higher present value of expected costs.

The deliberate holding of excess capacity from a technical point of view will be common in those cases. This again is no undesired excess capacity from an economic point of view, it is the well-known policy of building ahead of demand.

Another well-known fact can be explained by the analysis just given. When capacity is considerably expanded at a certain moment, there will be a drying up of investment in capacity, not because sales do not increase but just because the entrepreneur has expanded his plant as much as he had planned to do for the time being.

III.10. UNCERTAIN EXPECTATIONS AND FLUCTUATING DEMAND

The case of cyclical fluctuations in demand will now have to be considered. It is essential then to discard the assumption of subjective certainty of entrepreneurial expectations. The entrepreneur in the world of reality is uncertain about his expectations, he knows that they may afterwards prove to be wrong. There is of course also uncertainty when demand has shown a steady development in the past, but the more it has fluctuated in the past the more uncertainty will be felt about the future.

The uncertainties of expectations in investment planning apply both to future returns of sales and future expenses. The increments in expenses occasioned by having more machines are, as a rule, far less surrounded by uncertainty than the expectations of sales. Entrepreneurs know that demand curves for their products will not be constant over time but that they will shift according to a variety of causes such as general business conditions, buyers' tastes, prices of substitute products and so on; in short, uncertainty about buyers' and competitors' behaviour surrounds the firm's future sales. In so far as the additional costs entailed by additional machines are uncertain, it is largely correlated with the uncertainty regarding sales volume because costs of raw materials, labour and repairs are to a large extent dependent on actual output. When output is lower as a consequence of disappointing sales, these costs will also be lower.

Let us now deal with some of the probable consequences when, apart from a secular movement, there are cyclical fluctuations in sales. As a starting point we suppose that in the course of the cyclical upswing after a depression, a point is reached where available capacity is optimal (with due considerations to secular growth) for the current output. When sales continue to rise, what will be the reaction of capacity when entrepreneurs are uncertain about the future course of sales?

No hard and fast rules can be given here, but for several reasons it seems a plausible hypothesis to say that not an immediate but a lagged adjustment of capacity will follow this rise in output as a consequence of the risk that the rise in sales will be temporary. The longer the high levels of sales hold on, the more is it likely to be considered as a "normal" development, and the stronger the incentive to adjust capacity will be.

It is granted of course that for the higher output an expansion of capacity is cheaper than a more intensive use of the available capacity if the higher level of sales proved to be permanent. The line of thought of the previous sections would then be applicable. But there are other possibilities to be considered, sales may continue to rise further or may fall to lower levels again, and there is lack of symmetry between the consequences of favourable and unfavourable possibilities.

This is best shown when, as an alternative to an immediate expansion of capacity, we consider the case where this expansion is put off for some time, say a year, in order to see what the market will do, increasing output in the meantime by a more intensive and extensive use of available capacity. (In view of the possibilities of temporary peak outputs, it often occurs that not all relevant old and obsolete machinery is scrapped in the past, but that part of it is kept as a reserve; this is another kind of excess capacity which is desired from an economic point of view).

When sales do not fall back or are even rising further, profits per unit of output are lower over the year in question by the difference between average short-run and long-run costs. But "all is not lost that is delayed", after the year in question there is still the opportunity of expanding capacity. When sales fall back to a lower level, on the other hand, profits are higher or losses smaller when the expansion of capacity has been put off, and this may be so over a considerable period of time as a consequence of the fact that, if machines had been bought the fixed costs contingent upon them

would have continued. "Haste makes waste", says the proverb; as the entrepreneur gets further information about his market as time proceeds, he is free to alter his provisional plan but if he has adjusted his capacity immediately he may have undesired excess capacity during a considerable period.

But there is more; HART and FELLNER¹⁾ developed the principle of "linkage of risks", which also leads to asymmetry between favourable and unfavourable "surprise".

Due to institutional characteristics, as FELLNER pointed out, it is very likely that the provisions for uncertain developments of sales will be influenced more by the potentiality of unfavourable developments than by that of favourable developments. If the actual outcome of a capacity expansion is unfavourable, due to a drop in sales, by a margin sufficient to cause losses, then these losses are likely to be significantly increased in consequence of a sudden deterioration in the firm's relations with its creditors. Originally moderate losses frequently do not stay moderate but tend to become cumulative because the standing of the firm deteriorates. Immediate losses exceeding a certain percentage of the firm's net worth may even throw it into forced liquidation or bankruptcy.

This characteristic of losses has no counterpart in the event of immediate gains when sales prove to develop favourably. It is quite possible to argue that consistently favourable business results improve the credit standing of a firm and that this tends to increase the profitability of its ventures further; but the losses arising in the event of forced liquidations have no equivalent counterpart on the favourable side. The factor by which originally moderate losses are multiplied if credit facilities are withdrawn, is entirely out of line with

¹⁾ W. FELLNER, *Monetary policies and full employment*, Berkeley and Los Angeles, 1947, p. 152 ff.

A. G. HART, *Money Debt and Economic Activity*, New York 1953, p. 201 ff.

the proportion in which originally moderate gains may increase in consequence of the improved standing of the firm.

This asymmetry has objective (institutional) roots, but the decisions of the entrepreneur are also influenced by his subjective attitude towards risk. When the entrepreneur has risk aversion, then what is said above is strengthened, when he has a positive preference for risks then it is weakened; but in advanced industrial communities a considerably higher psychological weight is usually attached to the potentiality of a complete loss of assets than to the potentiality of high gains.

To conclude, it seems plausible to state the hypothesis that there will be a lagged adjustment of capacity to cyclical rises in output as a consequence of the uncertainty of expectations.

When at the beginning of the downturn an actual fall of output sets in, there will generally be firms for which, as a consequence of the lagged adjustment in the boom, the available capacity is still below the optimum level for the current output. When the belief in secular growth in the market is strong and the fall in output is only moderate, it is quite possible that some decisions to expand capacity are still made. When the fall in output holds on however, expansion of capacity will of course come to a standstill, available capacity being higher than optimal for the current output. Then excess capacity arises which is undesired from the economic point of view of the entrepreneur.

III.11. DISTRIBUTED LAGS IN CAPACITY ADJUSTMENT; A FINAL REFUTABLE HYPOTHESIS

The considerations of the previous sections will now be summarised in a final hypothesis, which can be submitted to statistical tests.

At a given level of output and considering the expected secular course in sales, the capacity on hand of a firm may be just sufficient or either too large or too small.

In the case of excess capacity there will generally be a tendency to get rid of the excess, but there are checks on a downward adjustment of capacity. First, if the low level of sales is expected to last for a short period only, the willingness not to replace what is scrapped may be absent or very small. The larger the excess capacity and the longer it lasts however, the stronger the tendency to scrap without replacement will be. Second, beyond a certain level scrapping is checked by the durability of machines.

In a situation of short capacity a tendency of plant expansion may be expected. But an immediate full adjustment is neither technically necessary nor considered possible or advisable from an economic point of view. There may be checks from the side of finance or the lumpiness of investment goods. Moreover, the high level of output may be expected to be temporary, in which case a "wait and see" policy will be followed. It seems reasonable to suppose that the incentive to take a decision on capacity expansion will be the stronger the larger the shortage of available capacity is and the longer the high level of output lasts.

Apart from these factors, causing lags between changes in output and decisions to adjust capacity, there is a lag between a decision to expand capacity and the actual enlargement of the productive capacity of a plant.

The adjustment of the capacity of a single plant will be more or less discontinuous, especially in the case of indivisibility of investment goods, the curve describing the adjustment path being a "stepped" one. When an industry with a large number of plants is considered, the adjustment of the total capacity to the total output of the industry will be much more a gradual process, described by a continuous adjustment curve.

New difficulties crop up however in the aggregation procedure, mainly as a consequence of shifts in sales between the firms in the industry and the entry of new firms. The speed

of adjustment of the capacity of an industry to aggregate output may not be invariant against changes in the distribution of output over the firms, new entries included. Little can be said on aprioristic grounds however about this effect, it will be neglected for the time being.

When considering the capacity-output relation in the industry the lag between the adjustment of capacity to output will be a distributed lag. For an individual plant the lag may be discrete, as was pointed out before. But when a large number of plants is considered, the speed of adjustment will be different among the plants, slower for some plants and faster for others. Hence for a whole industry a more continuous process of adjustment will hold.

Now the theory on the adjustment path and distributed lags developed in Chapter II and the ideas on the capacity-output relation of this chapter can be integrated. This will be done in formulating a final refutable hypothesis.

Let, in accordance with the notation which will be used in the empirical investigations of the next chapter, Y_t be the output in a certain period (say a year) t and K_{t-1} the capacity available at the beginning of period t . Further, $\lg Y_t = y_t$ and $\lg K_t = k_t$.

The final hypothesis alluded to is:

$$(III.11.1) \quad K_t = c Y_t^\alpha Y_{t-1}^\beta Y_{t-2}^{\lambda\beta} Y_{t-3}^{\lambda^2\beta} \dots e^{\gamma t}$$

where c is a constant, $0 \leq \lambda < 1$ and e the base of natural logarithms: $e = 2,718$.

After the exposition of chapter II the implications of this hypothesis can shortly be formulated. Let this first be done for the special and simpler case $\beta = \lambda\alpha$ in (III.11.1), i.e. for:

$$(III.11.2) \quad K_t = c Y_t^\alpha Y_{t-1}^{\lambda\alpha} Y_{t-2}^{\lambda^2\alpha} \dots e^{\gamma t}.$$

The most important implications of (III.11.2) are as follows.

- a. The percentage influence of lagged outputs on capacity decreases proportionately with increasing lags.

b. It follows from (III.11.2) that:

$$(III.11.3) \quad \frac{K_t}{K_{t-1}} = \left[\frac{dY_t^{\alpha/(1-\lambda)} e^{\gamma t}}{K_{t-1}} \right]^{1-\lambda}, \quad \text{where } d = ce^{\gamma\lambda/(1-\lambda)}.$$

There is optimum capacity $\bar{K}_t$ in period t for an output y_t when there is no change in the total capacity in period t , i.e. when $K_t = K_{t-1} = \bar{K}_t$. From (III.11.3) follows that:

$$(III.11.4) \quad \bar{K}_t = dY_t^{\alpha/(1-\lambda)} e^{\gamma t}.$$

The coefficient $\alpha/(1-\lambda)$ in (III.11.4) is the long-run capacity output elasticity $\alpha + \alpha\lambda + \alpha\lambda^2 + \dots$ of (III.11.2), which will be denoted by ε so that the equation for the optimum capacity $\bar{K}_t$ can also be written as:

$$(III.11.5) \quad \bar{K}_t = dY_t^{\varepsilon} e^{\gamma t}.$$

The ratio between optimum capacity and output is:

$$(III.11.6) \quad \frac{\bar{K}_t}{Y_t} = dY_t^{\varepsilon-1} e^{\gamma t}.$$

When there are no economies of large scale production for the degree of capacity utilization, ε will be equal to 1. When for larger levels of output the amount of desired reserve capacity is not proportionately larger (e.g. when seasonal fluctuations in output are additive), there are economies of scale for the degree of capacity utilization, and $\varepsilon < 1$.

The factor $e^{\gamma t}$ in (III.11.6) represents gradual changes that may occur in the optimum capacity for a given level of output. Concrete examples will be given in the empirical investigations.

c. Using the equation for $\bar{K}$, equation (III.11.3) can be written as:

$$(III.11.7) \quad \frac{K_t}{K_{t-1}} = \left[\frac{\bar{K}_t}{K_{t-1}} \right]^{1-\lambda},$$

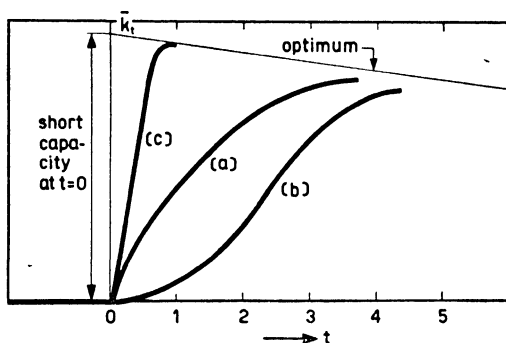
or taking natural logarithms:

$$(III.11.8) \quad \Delta k_{t-1} = (1-\lambda) [\bar{k}_t - k_{t-1}].$$

This means that the relative change of capacity in a unit period is a fraction $1-\lambda$ of the excess or short capacity $\bar{k}_t - k_{t-1}$ at the beginning of that period (λ may have different values for $\bar{k}_t > k_{t-1}$ and $\bar{k}_t < k_{t-1}$). It should be remembered that $\bar{K}_t$ is an economic optimum, it includes desired reserve capacity. The difference between optimum and actual capacity $\bar{k}_t - k_{t-1}$ is undesired excess or short capacity, which gives rise to changes in actual capacity, eliminating the undesired excess or shortage after a certain period of which the length is determined by λ .

d. Let us now return to the more general hypothesis (III.11.1).

The difference with the special case can best be illustrated by drawing the adjustment paths (as exposed in chapter II) for both cases.



Graph 18

In graph 18 the curve (a) is the adjustment path of k_t for the special case and (b) the path for the general case, the final hypothesis (III.11.1). (The optimum $\bar{k}$ for a given y is supposed to show a falling trend, i.e. $\gamma < 0$; when $\gamma = 0$ it should be constant and represented by a horizontal line).

In the general case the relative change of capacity in case of disequilibrium may be small in the beginning of the process of adjustment, e.g. as a consequence of lags in delivery of

machines etc., while in the special case the change in capacity is largest during the first period.

To finish this chapter, considering the theory summed up in the final equation (III.11.1) in the light of the recent discussions in the *Economic Journal* on the investment function ¹⁾, may have a clarifying effect.

We said before already, that the double inter-relation between investment and income plays an essential role in modern economic theory. The discussions referred to above deal with the dependence of investment on income (or output). In some models, such as those of HARROD, SAMUELSON and HICKS, investment is regarded as a function of the *change* in output, and not of output itself. In other models, such as those of KALECKI and KALDOR, investment is assumed to be an increasing function of the *level* of output (and a decreasing function of the stock of capital).

The position of the theory developed in this chapter can be given as follows.

1. From (III.11.5) $\bar{K}_t = dY_t^e e^{\gamma t}$ it follows that between alternative positions of long-period equilibrium of capacity, the change in capacity (i.e. net investment) is a function of the *change* in output.
2. From (III.11.3), which can be written as:

$$(III.11.9) \quad \Delta k_{t-1} = \alpha y_t - (1-\lambda)k_{t-1} + (1-\lambda)\gamma t + d',$$

a short-run relation for net investment in capacity can be derived.

In the short-run, total capacity k_{t-1} in this equation can approximately be treated as a constant when λ is considerably different from zero, i.e. when the adjustment of total capacity to output is slow as a consequence of the distributed lag. This is the usual assumption of the Marshallian (and Keynesian) short-run. When the influence of k_{t-1} and t is neglected

¹⁾ See especially N. KALDOR, Hicks on the trade cycle, *Economic Journal*, 1951, pp. 833-847.

for short-run changes, the short-run relation resulting from (III.11.9) is:

$$(III.11.10) \quad \Delta k_{t-1} = \alpha y_t + d',$$

net investment in capacity is a function of the *level* of output.

In the special case of (III.11.9) $\lambda = 0$, i.e. when an immediate full adjustment of capacity to its equilibrium level in the current period is assumed (see the adjustment path (c) in graph 18) the relation resulting from (III.11.9) is:

$$(III.11.11) \quad \Delta k_{t-1} = \alpha y_t - k_{t-1} + \gamma t + d',$$

or

$$(III.11.12) \quad k_t = \alpha y_t + \gamma t + d',$$

or

$$(III.11.13) \quad \Delta k_t = \alpha \Delta y_t + \gamma.$$

This special case is easily recognised as the acceleration principle in its rigorous form.

The theory of capacity-output adjustment given in this chapter makes no a priori choice in the KALDOR-HICKS discussion. If λ should be found to be near zero in an empirical case, the Hicksian theory, (and generally the acceleration principle) is approached; if on the other hand λ should be found to be considerably different from zero the KALDOR-theory is approached.

CHAPTER IV

EMPIRICAL INVESTIGATIONS ON THE TIME-SHAPE OF THE REACTION OF CAPACITY TO CHANGES IN OUTPUT

IV.1. RAILWAY FREIGHT TRAFFIC

A. Capacity and output in railway freight traffic

The theoretical analysis on distributed lags and adjustment of capacity to output will now be applied and submitted to statistical verification and estimation of coefficients.

The first industry to be studied is railway freight traffic and more particularly the adjustment of capacity, represented by the stock of freight cars available, to output, represented by ton miles of freight carried by the railway companies of the United States of America in the period 1894–1940.

The reasons for this choice are not very profound. The availability of statistical data over a long period is an obvious reason. In this period expansions of capacity not only occurred but also several years of contraction, so that a comparison between the time-shapes of upward and downward adjustments will be possible. The fact that freight cars are a clearly defined and easily measurable kind of capital good, that technological progress with respect to freight cars has not been sudden in the period considered (the size and quality of cars and the speed of freight trains have increased gradually) are favourable facts for a statistical analysis of our theoretically derived relation.

The fact, lastly, that we know two econometric studies in this field, TINBERGEN's study of "net investment, in railway rolling-stock" and KLEIN's study of "a demand equation

for plant and equipment of railroads”¹⁾, seemed a suitable starting point.

Our problem is slightly different from TINBERGEN’s and KLEIN’s. TINBERGEN dealt with net-investment, computed as a weighted arithmetic average of the per cent increases in the number of locomotives, freight cars and passenger cars, the weights being 20, 1 and 10 respectively. KLEIN gives an explanation of “net-investment of railroads, measured in billions of 1934 dollars”, where net-investment is computed as the difference between “expenditures for plant and equipment” and “depreciation charges”, both deflated by a price index number 1934 = 100.

In calculating real net-investment KLEIN had to divide the money expenditures for plant and equipment by a computed index number of the prices of plant and equipment, and these are procedures in which errors of aggregation are very likely. This will also be the case in deflating depreciation charges, and most of all perhaps in measuring depreciation itself. Of course KLEIN has realised this. About measuring depreciation from the accounting estimates of business firms, he says: “This is admittedly unsatisfactory because business records are not intended to show what the econometrician desires to measure. Depreciation policies of business firms may be influenced by tax considerations and other matters which have nothing to do with the consumption of fixed capital”. “It must be pointed out to the reader, though, that our estimates of net-investment are no better than the business records on depreciation”²⁾.

Errors of measurement and aggregation are much less probable in TINBERGEN’s study. Locomotives, freight cars

¹⁾ J. TINBERGEN, *Statistical testing of business-cycle theories I*, 1939 Geneva. Chapter VI.

L. R. KLEIN, *Economic Fluctuations in the United States 1921–1941*, New York 1950, p. 115 ff.

²⁾ L. R. KLEIN, *op.cit.* p. 125.

and passenger cars can easily be "counted". The computation of an aggregate index of investment by means of a weighted average of locomotives, freight cars and passenger cars probably does not entail large aggregation errors. As a consequence of the high correlation between per cent increases in locomotives, passenger cars and freight cars, the influence of the weights, and hence of errors in these weights, on the investment index for a rather short period is only small.

Nevertheless, to avoid these errors as far as possible—the period to be investigated by us being more than twice as long as TINBERGEN's—we restricted ourselves to the relation between freight cars and freight traffic.

As to the period to be investigated, statistical material goes even further back than 1894, the beginning of our period. In 1894 the great expansion of railways is over. The times when railway investments were mainly autonomous and speculative, when it was not demand for traffic that determined capacity but just the other way round, when it was much more the supply of traffic facilities (opening up new lines) that created demand, these times were past. The fact that 1894 not only showed a fall in ton miles produced but even a slight contraction in the total number of freight cars available is already an indication of the statement made above.

In a discussion of railroadization, SCHUMPETER says about the period 1878 till 1887 a.o.: "... the more an innovation becomes established, the more it loses the character of an innovation and the more it begins to follow impulses, instead of giving them". "Accordingly, railroad construction, increasingly setting into a predetermined frame work and exploiting pre-existing investment opportunities, became during the period under discussion much more (though not yet entirely) a function of railroad business and, hence, of the rest of the business organism than it had been before"¹).

¹) J. A. SCHUMPETER, *Business Cycles*, vol. I. New York, London 1939, p. 339.

So it will be safe to assume that after 1894 investment in freight cars will be basically induced by demand, so that the theory of the previous chapter can be used in the statistical analysis which now follows. The period 1916–1919 will be excluded however from the calculations; as a consequence of government interference with the railway companies and short supply of cars owing to large deliveries to allied forces in Europe, the theoretical framework of our investigations will probably not fit this period.

B. Statistical analysis

1. *Introduction.* In this analysis the following notation will be used:

K_t — number of freight cars available half way year t ;
unit 10^4 .

Y_t — number of ton miles of freight carried during year t ;
unit 10^6 .

U_t — disturbance.

t — years; $t = 0$ in 1917.

The theory to be submitted to statistical analysis can, in accordance with the previous chapters, be given as:

$$(IV.1.1) \quad K_{t+1} = k Y_t^\alpha Y_{t-1}^\beta Y_{t-2}^{\lambda\beta} Y_{t-3}^{\lambda^2\beta} \dots e^{\gamma t} \cdot U_t$$

or

$$(IV.1.2) \quad \frac{K_{t+1}}{K_t} = k' \cdot Y_t^\alpha Y_{t-1}^{\beta-\alpha\lambda} e^{\gamma(1-\lambda)t} K_t^{\lambda-1} \cdot U_t U_{t-1}^{-\lambda}.$$

Taking natural logarithms, and writing $\lg K_t = k_t$, $\lg Y_t = y_t$, $\lg U_t = u_t$ and further:

$$\begin{aligned} \varepsilon_1 &= \alpha \\ \varepsilon_2 &= \beta - \alpha\lambda \\ \varepsilon_3 &= \gamma(1 - \lambda) \\ \varepsilon_4 &= \lambda - 1, \end{aligned}$$

this gives:

$$(IV.1.3) \quad \Delta k_t = \varepsilon_1 y_t + \varepsilon_2 y_{t-1} + \varepsilon_3 t + \varepsilon_4 k_t + u_t - (\varepsilon_4 + 1) u_{t-1}$$

(all variables in deviations from their averages).

The exponential trend in (IV.1.1) represents technological progress in the productivity of freight cars: gradual improvements in quality and increases in tonnage of cars, increase in speed of freight trains, improvements in the efficiency and speed of loading and unloading of cars and in the switching process. These factors will cause a decrease in the optimum number of cars required for a given level of freight traffic.

2. *First analysis for the whole period.* In accordance with the exposition of II.9, the "direct" estimates of the coefficients ε of (IV.1.3) will be indicated by e and the consistent estimates, computed in accordance with the method developed in II.9, by $\bar{\varepsilon}$. Further let $a = e_1$

$$b - al = e_2$$

$$c(1 - l) = e_3$$

$$l - 1 = e_4$$

and

$$\bar{\alpha} = \bar{\varepsilon}_1$$

$$\bar{\beta} - \bar{\alpha} \bar{\lambda} = \bar{\varepsilon}_2$$

$$\bar{\gamma}(1 - \bar{\lambda}) = \bar{\varepsilon}_3$$

$$\bar{\lambda} - 1 = \bar{\varepsilon}_4.$$

To begin with the "direct" estimates are calculated. The result is:

$$\left\{ \begin{array}{l} e_1 = 0,077 \\ e_2 = 0,017 \\ e_3 = -0,0033 \\ e_4 = -0,110. \end{array} \right.$$

This gives the regression equation:

$$(IV.1.4) \quad \left\{ \begin{array}{l} \Delta k_t^* = 0,077 y_t + 0,017 y_{t-1} - 0,0033 t - 0,110 k_t \\ R^2 = 0,8480. \end{array} \right.$$

From the e 's the "direct" estimates a , b , c and l can easily be derived:

$$\begin{cases} a = 0,077 \\ b = 0,086 \\ c = -0,030 \\ l = 0,890. \end{cases}$$

The next step is to calculate $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, $\bar{\varepsilon}_3$ and $\bar{\varepsilon}_4$ from the following system of equations (confer (II.9.29) and (II.9.30)).

$$(IV.1.5) \quad \begin{cases} (\bar{\varepsilon}_1 - e_1) \sum y^2 + (\bar{\varepsilon}_2 - e_2) \sum yy_{-1} + (\bar{\varepsilon}_3 - e_3) \sum yt + (\bar{\varepsilon}_4 - e_4) \sum yk = 0 \\ (\bar{\varepsilon}_1 - e_1) \sum yy_{-1} + (\bar{\varepsilon}_2 - e_2) \sum y^2_{-1} + (\bar{\varepsilon}_3 - e_3) \sum yt_{-1} + (\bar{\varepsilon}_4 - e_4) \sum y_{-1}k = 0 \\ (\bar{\varepsilon}_1 - e_1) \sum yt + (\bar{\varepsilon}_2 - e_2) \sum yt_{-1} + (\bar{\varepsilon}_3 - e_3) \sum t^2 + (\bar{\varepsilon}_4 - e_4) \sum tk = 0 \\ (\bar{\varepsilon}_1 - e_1) \sum yk + (\bar{\varepsilon}_2 - e_2) \sum yk_{-1} + (\bar{\varepsilon}_3 - e_3) \sum tk + (\bar{\varepsilon}_4 - e_4) \sum k^2 = \end{cases}$$

$$= \frac{(\bar{\lambda} - \xi) \sum z^2}{1 - \xi \bar{\lambda} + l(\bar{\lambda} - \xi)}$$

where $\bar{\lambda} = 1 + \bar{\varepsilon}_4$ and z represents the residuals of (IV.1.4): $z_t = \Delta k_t - \Delta k_t^*$; for ξ see (II.9.16).

For alternative values of ξ in the interval $0 \leq \xi < 1$ the values of $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, $\bar{\varepsilon}_3$ and $\bar{\varepsilon}_4$ are calculated. From the $\bar{\varepsilon}$'s the estimates $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ and $\bar{\lambda}$ are easily derived. The result is given in the following table IV.1.1. The long-run elasticity of capacity with respect to output $\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$ is also given in table IV.1.1.

TABLE IV.1.1

	$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$\xi = 0$	0,89	0,077	0,087	- 0,030	0,86
$\xi = 0,4$	0,89	0,077	0,086	- 0,030	0,86
$\xi = 0,8$	0,89	0,077	0,086	- 0,030	0,86
$\xi = 0,95$	0,88	0,078	0,081	- 0,028	0,82

3. *Analysis for the whole period with a linear restriction on the coefficients.* There is some intercorrelation between k and t ($r_{k,t} = 0,62$), so it may be that especially $\bar{\gamma}$ and $\bar{\lambda}$ are unreliable. Therefore another calculation is made, where an a priori relation between α , β and γ is introduced: $\alpha + (\beta/1 - \lambda) = 1$.

This means that the long-run capacity output elasticity is supposed to be equal to one. In other words, it is assumed that there are no economies of large scale with respect to the capacity-output ratio, a rise in output of say ten per cent causing a rise of ten per cent in the long-run equilibrium number of freight cars.

When $\alpha + (\beta/1 - \lambda) = 1$ a relation between ε_1 , ε_2 and ε_3 in $\Delta k_t = \varepsilon_1 y_t + \varepsilon_2 y_{t-1} + \varepsilon_3 t + \varepsilon_4 k_t + u_t - (\varepsilon_4 + 1)u_{t-1}$ will exist.

Because:

$$\varepsilon_1 = \alpha$$

$$\varepsilon_2 = \beta - \alpha\lambda$$

$$\varepsilon_4 = \lambda - 1, \text{ this relation is: } \varepsilon_1 + \varepsilon_2 + \varepsilon_4 = 0.$$

After substitution of $\varepsilon_1 + \varepsilon_2 + \varepsilon_4 = 0$ we have:

$$(IV.1.6) \quad \Delta k_t = \varepsilon_1(y_t - k_t) + \varepsilon_2(y_{t-1} - k_t) + \varepsilon_3 t + u_t + (\varepsilon_1 + \varepsilon_2 - 1)u_{t-1}.$$

Again applying the method expounded in II.9 first the "direct" estimates e_1 , e_2 and e_3 are calculated by minimizing $\sum [\Delta k_t - e_1(y_t - k_t) - e_2(y_{t-1} - k_t) - e_3 t]^2$ for variations in e_1 , e_2 and e_3 . This gives:

$$\begin{cases} e_1 = 0,077 \\ e_2 = 0,008 \\ e_3 = -0,0033 \\ e_4 = -(e_1 + e_2) = -0,085 \end{cases} \quad \text{or} \quad \begin{cases} a = 0,077 \\ b = 0,079 \\ c = -0,039 \\ l = 0,915 \end{cases}$$

$$R^2 = 0,8231.$$

Let us now investigate how the consistent estimates $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, $\bar{\varepsilon}_3$ and $\bar{\varepsilon}_4$ are to be found when a linear restriction on the $\bar{\varepsilon}$'s is assumed. For convenience's sake $y_t - k_t$ is denoted by w_t and $y_{t-1} - k_t$ by w_{t-1} , (IV.1.6) can then be written as:

$$(IV.1.7) \quad \Delta k_t = \varepsilon_1 w_t + \varepsilon_2 w_{t-1} + \varepsilon_3 t + u_t - (\varepsilon_4 + 1)u_{t-1}.$$

From (IV.1.7) follows that:

(IV.1.8)

$$\begin{cases} \varepsilon_1 E w_t^2 & + \varepsilon_2 E w_t w_{t-1} + \varepsilon_3 E w_t t & + E u_t w_t & - (\varepsilon_4 + 1) E u_{t-1} w_t & = E w_t \Delta k_t \\ \varepsilon_1 E w_t w_{t-1} & + \varepsilon_2 E w_{t-1}^2 & + \varepsilon_3 E w_{t-1} t & + E u_t w_{t-1} & - (\varepsilon_4 + 1) E u_{t-1} w_{t-1} & = E w_{t-1} \Delta k_t \\ \varepsilon_1 E w_t t & + \varepsilon_2 E w_{t-1} t & + \varepsilon_3 E t^2 & + E u_t t & - (\varepsilon_4 + 1) E u_{t-1} t & = E t \Delta k_t \\ \varepsilon_1 & + \varepsilon_2 & & & + \varepsilon_4 & = 0. \end{cases}$$

Because $k_{t+1} = \alpha y_t + \beta y_{t-1} + \beta \lambda y_{t-2} + \beta \lambda^2 y_{t-3} \dots + \gamma t + u_t$ and $u_t = \xi u_{t-1} + \varepsilon_t$ we have in the first equation of (IV.1.8):

$$\begin{aligned} Eu_t w_t &= Eu_t(y_t - k_t) = -Eu_t k_t = -Eu_t u_{t-1} = -\xi Eu_{t-1}^2 \\ Eu_{t-1} w_t &= Eu_{t-1}(y_{t-1} - k_t) = -Eu_{t-1} k_t = -Eu_{t-1}^2. \end{aligned}$$

Similarly, in the second equation of (IV.1.8):

$$\begin{aligned} Eu_t w_{t-1} &= -Eu_t k_t = -\xi Eu_{t-1}^2 \quad \text{and} \\ Eu_{t-1} w_t &= -Eu_{t-1} k_t = -Eu_{t-1}^2. \end{aligned}$$

Further $Eu_t t = 0$ and $Eu_{t-1} t = 0$.

Substituting these results in (IV.1.8) gives:

$$(IV.1.9) \quad \left\{ \begin{array}{llll} \varepsilon_1 Ew_t^2 & + \varepsilon_2 Ew_t w_{t-1} + \varepsilon_3 Ew_t t & + (\varepsilon_4 + 1 - \xi) Eu_{t-1}^2 & = Ew_t \Delta k_t \\ \varepsilon_1 Ew_t w_{t-1} + \varepsilon_2 Ew_{t-1}^2 & + \varepsilon_3 Ew_{t-1} t + (\varepsilon_4 + 1 - \xi) Eu_{t-1}^2 & = Ew_{t-1} \Delta k_t \\ \varepsilon_1 Ew_t t & + \varepsilon_2 Ew_{t-1} t + \varepsilon_3 Et^2 & & = Et \Delta k_t \\ \varepsilon_1 & + \varepsilon_2 & + \varepsilon_4 & = 0. \end{array} \right.$$

In quite the same way as (II.9.26) was derived, it can be shown that:

$$(IV.1.10) \quad Eu_{t-1}^2 = \frac{Ez_t^2}{1 + (1 + \varepsilon_4)^2 - 2\xi(1 + \varepsilon_4) + (1 + \varepsilon_4 - \xi) \text{plim}(\varepsilon_4 - e_4)}.$$

It can easily be seen from (IV.1.9) and (IV.1.10) that consistent estimates $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, $\bar{\varepsilon}_3$ and $\bar{\varepsilon}_4$ can be defined by the following system of equations.

$$(IV.1.11) \quad \left\{ \begin{array}{llll} \bar{\varepsilon}_1 \sum w_t^2 & + \bar{\varepsilon}_2 \sum w_t w_{t-1} + \bar{\varepsilon}_3 \sum w_t t & + \frac{(\bar{\varepsilon}_4 + 1 - \xi) \sum z_t^2}{1 - \xi(\bar{\varepsilon}_4 + 1) + (1 + \bar{\varepsilon}_4 - \xi)(e_4 + 1)} & = \sum \Delta k_t w_t \\ \bar{\varepsilon}_1 \sum w_t w_{t-1} + \bar{\varepsilon}_2 \sum w_{t-1}^2 & + \bar{\varepsilon}_3 \sum w_{t-1} t + \frac{(\bar{\varepsilon}_4 + 1 - \xi) \sum z_t^2}{1 - \xi(\bar{\varepsilon}_4 + 1) + (1 + \bar{\varepsilon}_4 - \xi)(e_4 + 1)} & = \sum \Delta k_t w_{t-1} \\ \bar{\varepsilon}_1 \sum w_t t & + \bar{\varepsilon}_2 \sum w_{t-1} t + \bar{\varepsilon}_3 \sum t^2 & & = \sum \Delta k_t t \\ \bar{\varepsilon}_1 & + \bar{\varepsilon}_2 & + \bar{\varepsilon}_4 & = 0. \end{array} \right.$$

Remembering that e_1, e_2, e_3 and e_4 were calculated from:

$$(IV.1.12) \quad \begin{cases} e_1 \sum w_i^2 & + e_2 \sum w_i w_{i-1} + e_3 \sum w_i t & = \sum \Delta k_i w_i \\ e_1 \sum w_i w_{i-1} + e_2 \sum w_{i-1}^2 & + e_3 \sum w_{i-1} t & = \sum \Delta k_i w_{i-1} \\ e_1 \sum w_i t & + e_2 \sum w_{i-1} t + e_3 \sum t^2 & = \sum \Delta k_i t \\ e_1 & + e_2 & + e_3 & = 0 \end{cases}$$

and subtracting (IV.1.12) from (IV.1.11) we get, (writing

$$\kappa(\bar{e}_4, \xi) \text{ for } \frac{(\bar{e}_4 + 1 - \xi) \sum z_i^2}{1 - \xi(\bar{e}_4 + 1) + (1 + \bar{e}_4 - \xi)(e_4 + 1)}):$$

$$(IV.1.13) \quad \begin{cases} (\bar{e}_1 - e_1) \sum w_i^2 & + (\bar{e}_2 - e_2) \sum w_i w_{i-1} + (\bar{e}_3 - e_3) \sum w_i t & + \kappa(\bar{e}_4, \xi) = 0 \\ (\bar{e}_1 - e_1) \sum w_i w_{i-1} + (\bar{e}_2 - e_2) \sum w_{i-1}^2 & + (\bar{e}_3 - e_3) \sum w_{i-1} t + \kappa(\bar{e}_4, \xi) = 0 \\ (\bar{e}_1 - e_1) \sum w_i t & + (\bar{e}_2 - e_2) \sum w_{i-1} t + (\bar{e}_3 - e_3) \sum t^2 & = 0 \\ (\bar{e}_1 - e_1) & + (\bar{e}_2 - e_2) & + (\bar{e}_3 - e_3) & + (\bar{e}_4 - e_4) = 0. \end{cases}$$

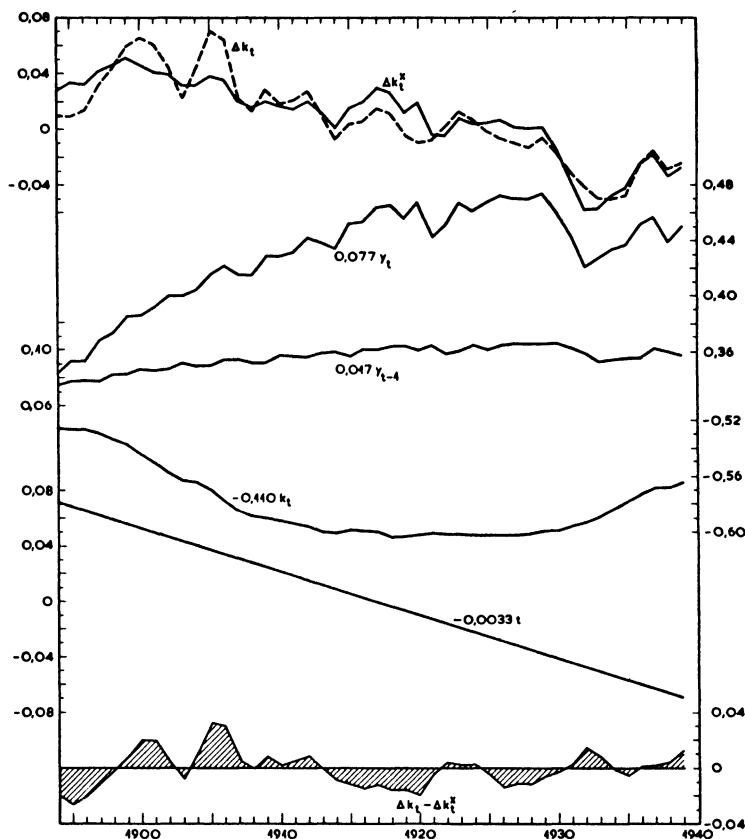
From (IV.1.13) an equation in $\bar{e}_4$ and ξ can be derived. The values of $\bar{e}_4$ for alternative values of ξ in the interval $0 \leq \xi < 1$ can then be calculated. As $\bar{e}_4 = \bar{\lambda} - 1$ and $0 \leq \bar{\lambda} < 1$ only the values of $\bar{e}_4$ between zero and minus one are relevant. After substitution of ξ and the corresponding value of $\bar{e}_4$ just calculated, in three equations of (IV.1.13) the estimates $\bar{e}_1, \bar{e}_2$ and $\bar{e}_3$ are easily found. Then $\bar{\alpha}, \bar{\beta}, \bar{\gamma}$ and $\bar{\lambda}$ can again be derived from $\bar{\alpha} = \bar{e}_1, \bar{\beta} - \bar{\alpha}\bar{\lambda} = \bar{e}_2, \bar{\gamma}(1 - \bar{\lambda}) = \bar{e}_3$ and $\bar{\lambda} - 1 = \bar{e}_4$.

The results of the calculations are given in table IV.1.2 below.

TABLE IV.1.2

	$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{e}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$\xi = 0$	0,920	0,077	0,076	- 0,040	a priori = 1
$\xi = 0,4$	0,920	0,077	0,076	- 0,040	a priori = 1
$\xi = 0,8$	0,918	0,077	0,077	- 0,040	a priori = 1
$\xi = 0,9$	0,915	0,077	0,078	- 0,039	a priori = 1
$\xi = 0,95$	0,913	0,078	0,081	- 0,039	a priori = 1

The assumption $\alpha + (\beta/1 - \lambda) = 1$ apparently does not lead to great differences with the previous calculations, it only means



Graph 19

a considerable change in $\bar{\gamma}$. Practically for all relevant values of ξ the consistent estimates are again not significantly different from the direct least-squares estimates.

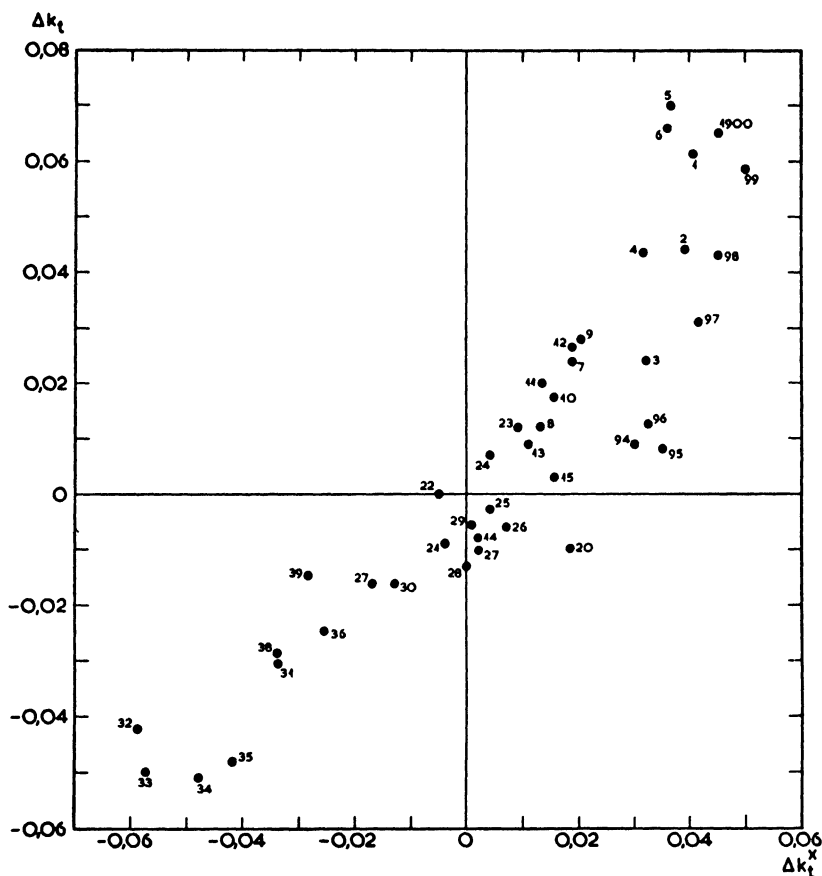
The results of the first calculation, where the following estimates were found for $0 \leq \xi \leq 0,8$:

$$\bar{\alpha} = 0,077$$

$$\bar{\beta} = 0,086$$

$$\bar{\gamma} = -0,030$$

$$\bar{\lambda} = 0,89 \text{ (see table IV.1.1),}$$



Graph 20

corresponding with the regression equation

$$(IV.1.4) \quad \Delta k_t = 0,077 y_t + 0,017 y_{t-1} - 0,0033 t - 0,110 k_t,$$

is illustrated in graphs 19 and 20.

4. *Analysis for the expansion period 1894–1915.* It was said in chapters II and III that there may be a difference in the speed of adjustment between upward and downward adjustments. Roughly speaking the period 1894–1915 in the above analysis on railway freight traffic is one of expansion in

the stock of freight cars, while the years 1920–1939 are years of contraction. An inspection of graph 20 indeed suggests a structural difference between prewar expansion and postwar contraction, graph 19 moreover shows that the correlation for the period 1894–1915 is much lower than that for the period 1920–1939. Separate statistical investigations for the period 1894–1915 and for the period 1920–1939 would seem worth while.

First the period of expansion 1894–1915 will be considered. As a consequence of the high intercorrelation between k_t and t a restriction between the coefficients is assumed, viz. $\alpha + \beta/(1 - \lambda) = \varepsilon_L$ corresponding with $\varepsilon_1 + \varepsilon_2 + \varepsilon_L \varepsilon_4 = 0$. In these relations ε_L , the long-run capacity-output elasticity, is given a priori values $\varepsilon_L = 1$ and $\varepsilon_L = 0,8$.

As discussed in details before, first the direct estimates e_1, e_2, e_3, e_4 and the corresponding direct estimates a, b, c and l are calculated. The results are given in table IV.1.3.

TABLE IV.1.3

$\varepsilon_L = 1$				$\varepsilon_L = 0,8$			
$e_1 =$	0,1760	$a =$	0,1760	$e_1 =$	0,1567	$a =$	0,157
$e_2 =$	0,0824	$b =$	0,213	$e_2 =$	0,0843	$b =$	0,194
$e_3 =$	- 0,0075	$c =$	- 0,029	$e_3 =$	- 0,0048	$c =$	- 0,016
$e_4 =$	- 0,2556	$l =$	0,744	$e_4 =$	- 0,3002	$l =$	0,700
$R^2 = 0,8158$				$R^2 = 0,7600$			

The next step again consists of calculating the consistent estimates $\bar{\varepsilon}_1, \bar{\varepsilon}_2, \bar{\varepsilon}_3, \bar{\varepsilon}_4$ and the corresponding $\bar{\alpha}, \bar{\beta}, \bar{\gamma}$ and $\bar{\lambda}$ using the linear restrictions $\bar{\varepsilon}_1 + \bar{\varepsilon}_2 + \bar{\varepsilon}_4 = 0$ and $\bar{\varepsilon}_1 + \bar{\varepsilon}_2 + 0,8\bar{\varepsilon}_4 = 0$ respectively. In the tables IV.1.4 and IV.1.5 the results of the calculations are given.

Here again the influence of the unknown coefficient ξ on the regression coefficients proves to be low, and the estimates $\bar{\alpha}, \bar{\beta}, \bar{\gamma}$ and $\bar{\lambda}$ are only slightly different from a, b, c and l .

TABLE IV.1.4
PERIOD 1894-1915

	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\lambda}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1-\bar{\lambda})$
$\xi = 0$	0,173	0,206	- 0,029	0,753	a priori = 1
$\xi = 0,4$	0,174	0,208	- 0,029	0,751	a priori = 1
$\xi = 0,7$	0,176	0,213	- 0,029	0,744	a priori = 1
$\xi = 0,8$	0,177	0,214	- 0,029	0,741	a priori = 1
$\xi = 0,9$	0,180	0,223	- 0,029	0,730	a priori = 1
$\xi = 0,95$	0,185	0,227	- 0,029	0,716	a priori = 1

TABLE IV.1.5

	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\lambda}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1-\bar{\lambda})$
$\xi = 0$	0,152	0,184	- 0,016	0,718	a priori = 0,8
$\xi = 0,4$	0,154	0,187	- 0,016	0,712	a priori = 0,8
$\xi = 0,7$	0,157	0,194	- 0,016	0,700	a priori = 0,8
$\xi = 0,8$	0,159	0,200	- 0,016	0,689	a priori = 0,8
$\xi = 0,9$	0,165	0,215	- 0,016	0,662	a priori = 0,8
$\xi = 0,95$	0,172	0,232	- 0,016	0,632	a priori = 0,8

5. *Analysis for the contraction-period 1920-1939.* The years 1920-1939 were a period of contraction in the stock of freight cars, with a few small exceptions in the beginning of the twenties. This period will now be submitted to statistical analysis.

Three types of calculations are made. In the first no restrictions on the coefficients are assumed, in the second it is assumed that $\varepsilon_L=1$ or in other words, that $\varepsilon_1+\varepsilon_2+\varepsilon_4=0$, while it is assumed that $\varepsilon_L=0,8$ or in other words that $\varepsilon_1+\varepsilon_2+0,8\varepsilon_4=0$ in the third type of calculations.

After the detailed exposition of the methods of calculation in the previous investigations, these methods will not be discussed again; only the results of the analysis will be given. In tables IV.1.6, IV.1.7 and IV.1.8 the results of the first, second and third type of calculations mentioned above are shown.

TABLE IV.1.6

Direct estimates	Consistent estimates					
		$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\lambda}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,066$	$\xi = 0$	0,067	0,052	- 0,036	0,936	0,88
$b = 0,053$	$\xi = 0,4$	0,067	0,052	- 0,036	0,936	0,88
$c = - 0,032$	$\xi = 0,7$	0,067	0,052	- 0,035	0,934	0,86
$l = 0,922$	$\xi = 0,8$	0,067	0,053	- 0,035	0,931	0,84
$\varepsilon_L = a + (b/1 - l) = 0,74$	$\xi = 0,9$	0,066	0,053	- 0,033	0,924	0,76
$R^2 = 0,7918$	$\xi = 0,95$	0,065	0,054	- 0,030	0,911	0,67

TABLE IV.1.7

Direct estimates	Consistent estimates					
		$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\lambda}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,069$	$0 \leq \xi \leq 0,9$	0,069	0,053	- 0,037	0,943	a priori = 1
$b = 0,054$						
$c = - 0,037$	$\xi = 0,95$	0,069	0,054	- 0,037	0,942	a priori = 1
$l = 0,942$						
$R^2 = 0,7845$						

TABLE IV.1.8

Direct estimates	Consistent estimates					
		$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\lambda}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,067$	$\xi = 0$	0,066	0,051	- 0,033	0,930	a priori = 0,8
$b = 0,053$	$\xi = 0,4$	0,066	0,051	- 0,033	0,930	a priori = 0,8
$c = - 0,033$	$\xi = 0,7$	0,066	0,051	- 0,033	0,930	a priori = 0,8
$l = 0,928$	$\xi = 0,8$	0,066	0,051	- 0,033	0,930	a priori = 0,8
$R^2 = 0,7892$	$\xi = 0,9$	0,067	0,053	- 0,033	0,928	a priori = 0,8
	$\xi = 0,95$	0,067	0,054	- 0,033	0,927	a priori = 0,8

6. *Summary of results.* A short summary and comparison of the results of the various statistical calculations will now be given. The equation underlying all investigations is $k_{t+1} = \alpha y_t + \beta y_{t-1} + \lambda \beta y_{t-2} + \lambda^2 \beta y_{t-3} \dots + \gamma t + u_t$, where u_t is supposed to behave according to $u_t = \xi u_{t-1} + \varepsilon_t$.

The difference between the least-squares estimates a, b, c, l

and the consistent estimates as calculated according to the method developed in II.9 proved to be small in nearly all cases. This is especially so when the autoregression of the disturbances is supposed to have low or moderate values. When ξ , the coefficient of the autoregression of the disturbances, approaches one, the difference becomes larger in all cases.

The time-shape of the reaction of k on y can be described by the coefficients $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\lambda}$. As exposed in the second chapter (cf. II.6 and graph 7) the coefficient $\bar{\lambda}$ has to do with the speed of adjustment of k to y . The lower $\bar{\lambda}$ the higher the speed of adjustment.

The coefficients $\bar{\alpha}$ and $\bar{\beta}$ have to do with the short-term reactions of k to y , $\bar{\alpha}$ represents the elasticity of k with respect to y during the first year of the adjustment period and $\bar{\alpha} + \bar{\beta}$ the elasticity of the reaction during the first and second year together.

The lower $\bar{\lambda}$ and the higher $\bar{\alpha}$ and $\bar{\beta}$ the shorter the period of adjustment of k to changes in y will be; or in other words, the more the changes in k will be determined by the changes in the current values of y (which, by the way, underlies the acceleration principle).

In all investigation $\bar{\lambda}$ proved to be high and $\bar{\alpha}$ and $\bar{\beta}$ relatively low, which means that there is a considerable difference between the short-run and the long-run reaction. Changes in current output y are only slightly followed by changes in current capacity k . There is a considerable "toughness" in the behaviour of k over time which is indicated by the high values of $\bar{\lambda}$.

There is also a strong indication of a structural difference between the period before the first world war, a period characterized among others, by an expansion in the stock of freight cars, and the period of contraction in the stock of cars between the first and second world war.

In table IV.1.9 the values of $\bar{\alpha} + \bar{\beta}$ and $\bar{\lambda}$ are summarized to illustrate what is said above. The intervals of $\bar{\alpha} + \bar{\beta}$ and $\bar{\lambda}$ in this table, are corresponding with the interval $0 \leq \xi \leq 0,95$.

TABLE IV.1.9

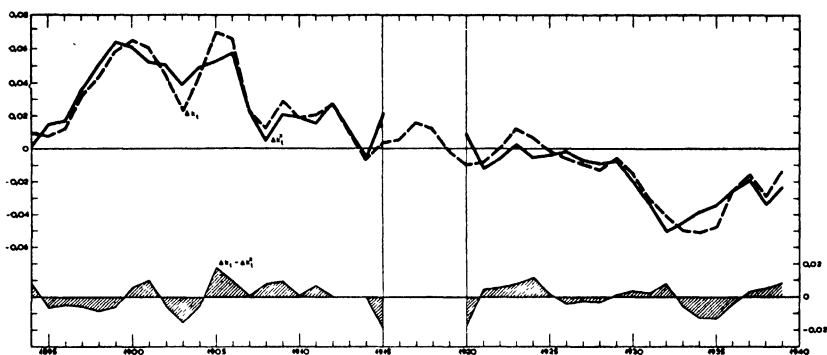
SUMMARY TABLE

Period of analysis			
	whole period	expansion period	contraction period
no restriction on the coeffi- cients	$\bar{\alpha} + \bar{\beta} = 0,164-0,159$ $\bar{\lambda} = 0,89 -0,88$	not calculated ($\bar{\lambda}$ probably unreliable)	$\bar{\alpha} + \bar{\beta} = 0,117-0,121$ $\bar{\lambda} = 0,94 -0,91$
assumption: $\varepsilon_L = 0,8$	practically the same as above: $\bar{\alpha} + \bar{\beta} = 0,164-0,159$ $\bar{\lambda} = 0,89 -0,88$	$\bar{\alpha} + \bar{\beta} = 0,336-0,404$ $\bar{\lambda} = 0,72 -0,63$	$\bar{\alpha} + \bar{\beta} = 0,117-0,121$ $\bar{\lambda} = 0,93 -0,93$
assumption: $\varepsilon_L = 1$	$\bar{\alpha} + \bar{\beta} = 0,153-0,159$ $\bar{\lambda} = 0,92 -0,91$	$\bar{\alpha} + \bar{\beta} = 0,379-0,412$ $\bar{\lambda} = 0,75 -0,72$	$\bar{\alpha} + \bar{\beta} = 0,112-0,123$ $\bar{\lambda} = 0,94 -0,94$

As can easily be seen from this summary table, $\bar{\alpha} + \bar{\beta}$ is considerably higher and $\bar{\lambda}$ lower for the expansion period than for the contraction period, or in other words, the speed of adjustment is higher when the adjustment is an expansion in the stock of cars than in the case when the adjustment is a process of contraction. Hence, fluctuations in recent output are more reflected in changes in the stock of cars in the expansion period before the first world war than is the case in the period 1920–1939. Though $\bar{\lambda}$ is still high for the period 1894–1915 and considerably different from zero, which is the assumption underlying the acceleration principle in its rigid form (cf. III.11), the expansion period would fit this principle better than the contraction period between the wars.

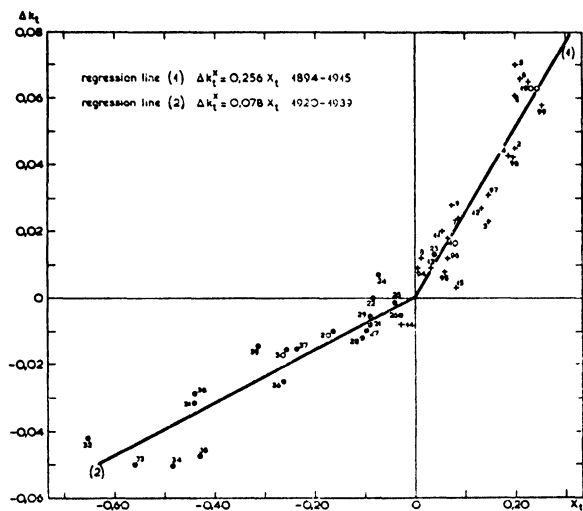
Finally a graphical illustration of the statistical analysis is given in graphs 21 and 22.

In graph 21 Δk_t and Δk_t^* are shown, where Δk_t^* is calculated separately for the period 1894–1915 and 1920–1939. For the period 1894–1915 the coefficients $\bar{\alpha}=0,176$, $\bar{\beta}=0,213$, $\bar{\gamma}=-0,029$ and $\bar{\lambda}=0,744$ are chosen (cf. table IV.1.3) and for the period 1920–1939 Δk_t^* is calculated with the coeffi-



Graph 21

icients $\bar{\alpha} = 0,066$, $\bar{\beta} = 0,053$, $\bar{\gamma} = -0,032$ and $\bar{\lambda} = 0,922$ (cf. Table IV.1.6).



Graph 22

To give a graphical representation of the difference between $\bar{\lambda}$ for 1894 – 1915 and $\bar{\lambda}$ for 1920 – 1939 the equation for Δk_t^* is written somewhat differently than before. We had:

$$\Delta k_t = \varepsilon_1 y_t + \varepsilon_2 y_{t-1} + \varepsilon_3 t + \varepsilon_4 k_t,$$

or

$$\Delta k_t = \alpha y_t + (\beta - \alpha \lambda) y_{t-1} + \gamma (1 - \lambda) t + (\lambda - 1) k_t.$$

This can be written as:

$$\Delta k_t = (1-\lambda) \left\{ \frac{\alpha}{1-\lambda} y_t + \frac{\beta-\alpha\lambda}{1-\lambda} y_{t-1} + \gamma t - k_t \right\}.$$

When the expression between $\{ \}$ is denoted by x_t we have: $\Delta k_t = (1-\lambda)x_t$. Now x_t is first calculated for 1894–1915 using $\bar{\alpha}=0,176$, $\bar{\beta}=0,213$, $\bar{\gamma}=-0,029$ and $\bar{\lambda}=0,744$ and then for the period 1920–1939 with $\bar{\alpha}=0,066$, $\bar{\beta}=0,053$, $\bar{\gamma}=-0,032$ and $\bar{\lambda}=0,922$. The regression equation for the first period then becomes $\Delta k_t^* = 0,256 x_t$ and for the second period $\Delta k_t^* = 0,078 x_t$.

This is illustrated by the regression lines (1) and (2) in graph 22. In this way the difference between the $\bar{\lambda}$'s is indicated by the difference in the slopes of the regression lines in the scatter diagram.

IV.2. THE ELECTRIC LIGHT AND POWER INDUSTRY

A. *Capacity and output in electric utilities*

The capacity of electric utilities may best be measured at the generating plant. From 1920 onwards, the industry provides annual statistical data on rated capacity of generators in kilowatts and output in kilowatt hours generated, both for fuel-burning and hydro-electric plants. Incidentally, it may be noted that expenditure for generating plant commonly exceeds that for either transmission or distribution and ranges from less than 30 to 50 % of the total investment. The ratio between output and capacity is referred to by the Federal Power Commission as the "plant factor", it is defined as the quotient of current generated and the product of the midyear kilowatt capacity and 8760, the number of hours in the year.

For hydro-electric plants the plant factor is considerably higher than that for fuel-burning plants; as a consequence of the high fixed costs of hydro-electric stations, water power generation is generally regarded as suitable only for areas affording a high load factor. In order to avoid the effect of

relative changes in output between hydro-electric and fuel-burning plants, although these changes are very small in practice, separate investigations for both parts of the industry will be made.

Although the electric power industry is a fast growing industry, the rate of growth is gradually declining. In a study by GOULD¹⁾ the trend in output is given by a logarithmic parabola, with annual rates of growth of

17,8	%	in	1900
14,4	,,	,,	1910
11,1	,,	,,	1920
7,9	,,	,,	1930
4,7	,,	,,	1940.

The industry is further characterized by highly indivisible equipment and economies in the use of large units. "Large generators involve lower unit investment both in turbine and related equipment. They enjoy slightly lower fuel consumption and lower operating and maintenance costs"²⁾.

The high rate of growth in demand and the economies in installing large generators gave rise to a policy of building ahead of demand, so that a large degree of overcapacity is usual for this industry. This is strengthened by the fact that the product cannot be stocked while, on the other hand, there are large daily and seasonal fluctuations in demand. The overcapacity may be, as explained in Ch. III, wholly or to a large extent an optimum, a desired, reserve capacity.

The records, however, give the impression that the degree of optimum reserve capacity is gradually falling, or in other words, that there is a secular rise in the optimum load factor. There are several reasons for this. There is a relative increase

¹⁾ J. M. GOULD, *Output and productivity in the electric and gas utilities 1899-1942*, Nat. Bureau of Econ. Research: 1946.

²⁾ E. G. NOURSE and associates, *America's capacity to produce*, Washington, D.C. 1934, p. 330.

in so called off-peak consumption, "demand has been considerably diversified as a result of a host of institutional changes, inclusive rate adjustment and "education" ¹⁾. The progress of interconnections also made it possible to have less reserve capacity. On the other hand, the improvements in the standards of service require a larger reserve for emergencies, but in view of the facts it is likely that this factor is outweighed by the others, so that on balance there is a gradual fall in the optimum capacity-output ratio.

B. Statistical investigations

1. *Introduction.* The theory to be submitted to statistical analysis is again given by

$$(IV.2.1) \quad k_{t+1} = \alpha y_t + \beta y_{t-1} + \lambda \beta y_{t-2} + \lambda^2 \beta y_{t-3} \dots + \gamma t + u_t$$

or

$$(IV.2.2) \quad \Delta k_t = \alpha y_t + (\beta - \alpha \lambda) y_{t-1} + \gamma(1 - \lambda)t + (\lambda - 1)k_t + u_t - \lambda u_{t-1},$$

where again

$$k_t = \lg K_t$$

$$y_t = \lg Y_t$$

$$K_t = \text{midyear capacity in mil. kw. in year } t$$

$$Y_t = \text{bil. kwh. generated in year } t.$$

The gradual fall in the optimum capacity-output ratio (or rise in the optimum load factor), which is considered very likely in the exposition just given, is represented in the theory above by the trend factor γt , where $\gamma < 0$.

The methods of estimation are the same as in the railway investigations, so they will not be discussed in detail again. Only the results of the calculations will be given, first for hydro-electric and then for fuel-burning plants. Again, as in the railway analysis, alternative a priori assumptions of the

¹⁾ GOULD, *op.cit.* p. 69.

long-run capacity-output elasticity $\varepsilon_L = \alpha + (\beta/1 - \lambda)$ are made, viz. $\varepsilon_L = 1$ and $\varepsilon_L = 0,8$. The assumption $\varepsilon_L = 1$ implies that a change of say 10 % in output causes an equal percentage change in optimum capacity; $\varepsilon_L = 0,8$ implies an 8 % change in optimum capacity as a consequence of a 10 % change in output, there are economies of large scale with respect to the optimum capacity-output ratio in this case. The value 0,8 is chosen on account of the low or moderate economies of scale in the capacity-output ratio, which is suggested by technical literature; further, lower values for ε_L , which were also tried, generally gave results which were unsatisfactory on economic grounds (e.g. a positive value for γ).

For hydro-electric plants the results of the statistical analysis are rather unsatisfactory. The correlation is low and,

TABLE IV.2.1
Hydro-electric plants, a priori $\varepsilon_L = 0,8$

"Direct" estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,082$	$\xi = 0$	0,782	0,077	0,157	-0,0101	a priori = 0,8
$b = 0,188$	$\xi = 0,4$	0,773	0,078	0,164	-0,0093	a priori = 0,8
$c = -0,0069$	$\xi = 0,8$	0,716	0,085	0,204	-0,0057	a priori = 0,8
$l = 0,737$	$\xi = 0,95$	0,561	0,107	0,306	-0,0008	a priori = 0,8
$R^2 = 0,6880$						

TABLE IV.2.2
Hydro-electric plants, a priori $\varepsilon_L = 1$

"Direct" estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,117$	$\xi = 0$	0,806	0,104	0,174	-0,023	a priori = 1
$b = 0,210$	$\xi = 0,4$	0,800	0,106	0,179	-0,022	a priori = 1
$c = -0,0190$	$\xi = 0,8$	0,760	0,120	0,214	-0,018	a priori = 1
$l = 0,762$	$\xi = 0,95$	0,553	0,176	0,368	-0,011	a priori = 1
$R^2 = 0,6660$						

TABLE IV.2.3
Fuel-burning plants, a priori $\varepsilon_L = 0,8$

"Direct" estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,073$	$\xi = 0$	0,778	0,073	0,160	- 0,018	a priori = 0,8
$b = 0,168$	$\xi = 0,4$	0,775	0,073	0,163	- 0,018	a priori = 0,8
$c = - 0,0043$	$\xi = 0,8$	0,768	0,073	0,168	- 0,018	a priori = 0,8
$l = 0,768$	$\xi = 0,95$	0,750	0,072	0,180	- 0,013	a priori = 0,8
$R^2 = 0,8681$						

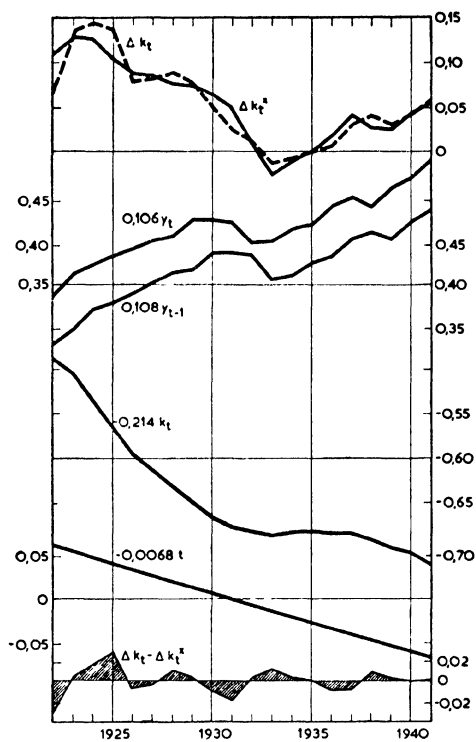
TABLE IV.2.4
Fuel-burning plants, a priori $\varepsilon_L = 1$

"Direct" estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,106$	$\xi = 0$	0,793	0,106	0,185	- 0,033	a priori = 1
$b = 0,191$	$\xi = 0,4$	0,791	0,106	0,186	- 0,033	a priori = 1
$c = - 0,032$	$\xi = 0,8$	0,786	0,106	0,191	- 0,032	a priori = 1
$l = 0,786$	$\xi = 0,95$	0,770	0,105	0,206	- 0,028	a priori = 1
$R^2 = 0,8790$						

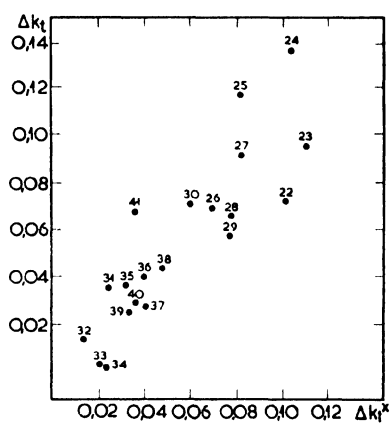
at the same time, the intervals for the consistent estimates admitted by the hypothesis $0 \leq \xi < 1$ are large.

The results of the investigations are more satisfactory in the case of fuel-burning plants. The correlation is higher and the intervals for the consistent estimates are narrow.

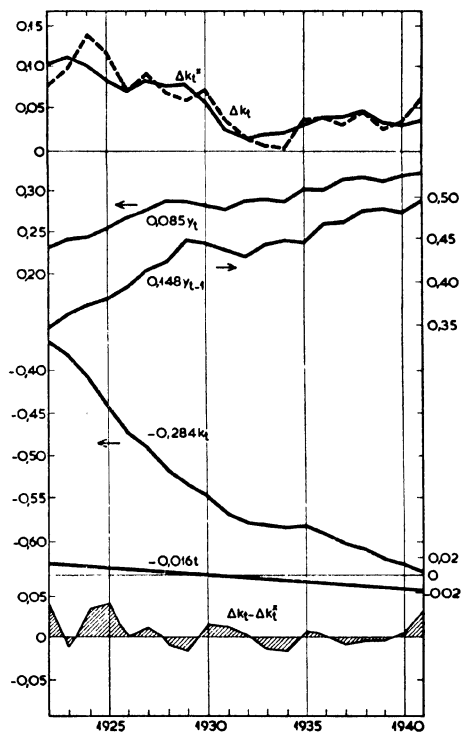
A graphical illustration of the statistical investigations is given in graphs 23 and 24 for hydro-electric plants, and in graphs 25 and 26 for fuel-burning plants. In the case of hydro-electric plants Δk_t^* is calculated for $\bar{\alpha} = 0,085$, $\bar{\beta} = 0,204$, $\bar{\gamma} = - 0,0057$, $\bar{\lambda} = 0,716$, corresponding with the regression equation: $\Delta k_t^* = 0,085 y_t + 0,143 y_{t-1} - 0,0016 t - 0,284 k_t$. For fuel-burning plants, Δk_t^* as plotted in graphs 25 and 26 is calculated with the use of $\bar{\alpha} = 0,106$, $\bar{\beta} = 0,191$, $\bar{\gamma} = - 0,032$, $\bar{\lambda} = 0,786$, corresponding with the equation $\Delta k_t^* = 0,106 y_t + 0,108 y_{t-1} - 0,0068 t - 0,214 k_t$.



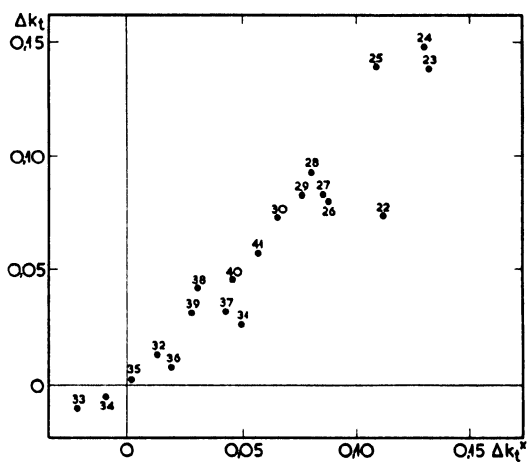
Graph 23



Graph 24



Graph 25



Graph 26

IV.3. PORTLAND CEMENT INDUSTRY

A. Capacity and output

There are two measures of capacity in use in this industry, viz. kiln capacity and grinding capacity ¹⁾).

In producing portland cement, the raw materials, once tested and properly apportioned, pass into giant rotary kilns where they are burned and fused into clinker. This is an intermediate product, which is ground into a highly standardized finished cement.

Demand for cement is fluctuating heavily. The main sources of demand, construction and road building, show large seasonal and cyclical fluctuations. The huge rotary kilns must always operate continuously if they operate at all. The rate of operation in the production of clinker is determined by the number of kilns in operation. Consequently, in a five-kiln mill only five different levels of operation are possible: 20, 40, 60, 80, and 100 %. Therefore, sufficient equipment must be installed to meet the peak loads. This gives rise to reserve capacity and since the average fixed investment per barrel of clinker capacity is relatively very high, the financial burden that must be carried in periods of low activity is apparent. The fixed investment per barrel of grinding capacity is much smaller. Hence, to minimize the costs, the industry irons out some of the seasonality by storing the intermediate product, clinker, in the slack months for grinding in the busy grinding season.

This makes kiln capacity less than grinding capacity. For the industry as a whole the difference is about 10 %. The changes over time of both capacities are practically speaking the same. For our statistical investigation it does not matter very much which of the two is chosen, the im-

¹⁾ For this and the following, confer E. B. ALDERFER and H. E. MICHL, *Economics of American industry*, New York and London, 1942, p. 174 ff.

portant requirement for us is that the capacity be reported on the same basis from year to year. This is so with the figures of the United States Bureau of Mines, which reports the grinding capacity of the finishing mills. These are therefore used in the statistical analysis, where K_t refers to the mid-year capacity in millions of barrels of finished cement and Y_t to the output in millions of barrels of finished cement.

B. Statistical investigations

Applying the same theory and using the same notation as before, the results of the analysis are at once given in tables IV.3.1 and IV.3.2.

TABLE IV.3.1

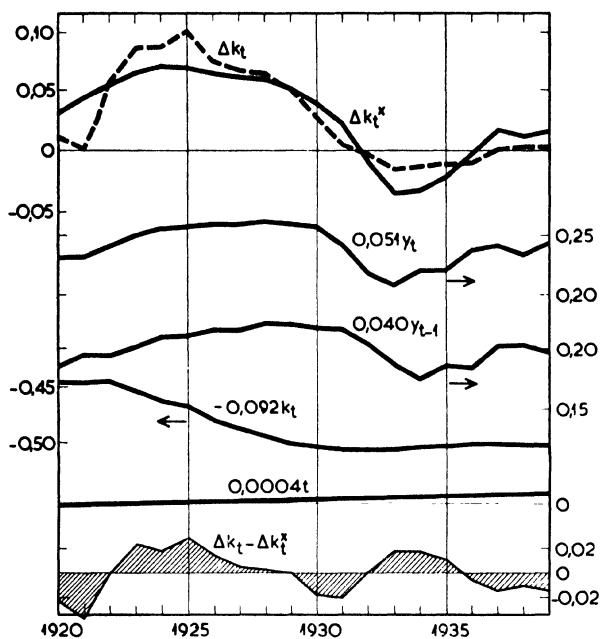
Direct estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,051$	$\xi = 0$	0,910	0,050	0,086	0,003	a priori = 1
$b = 0,089$	$\xi = 0,4$	0,910	0,050	0,086	0,003	a priori = 1
$c = 0,005$	$\xi = 0,8$	0,908	0,051	0,086	0,004	a priori = 1
$l = 0,907$	$\xi = 0,95$	0,905	0,051	0,090	0,005	a priori = 1
$R^2 = 0,7514$						

TABLE IV.3.2

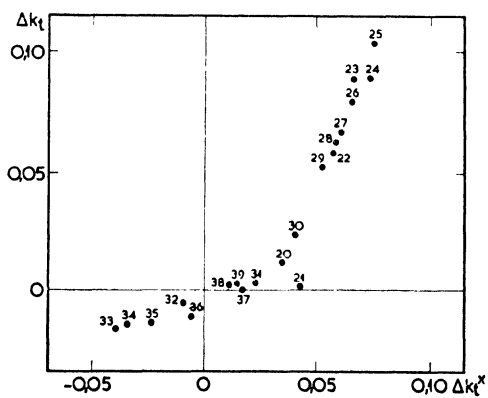
Direct estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,040$	$\xi = 0$	0,907	0,036	0,073	0,005	a priori = 0,8
$b = 0,076$	$\xi = 0,4$	0,906	0,037	0,074	0,005	a priori = 0,8
$c = 0,004$	$\xi = 0,8$	0,904	0,038	0,075	0,005	a priori = 0,8
$l = 0,901$	$\xi = 0,95$	0,895	0,042	0,077	0,004	a priori = 0,8
$R^2 = 0,6913$						

A graphical illustration of the investigations is given in graphs 27 and 28, where Δk_t^* is calculated for $\bar{\alpha} = 0,051$, $\bar{\beta} = 0,086$, $\bar{\gamma} = 0,004$ and $\bar{\lambda} = 0,907$, corresponding with the equation:

$$\Delta k_t^* = 0,051 y_t + 0,040 y_{t-1} + 0,0004 t - 0,092 k_t.$$



Graph 27



Graph 28

Just as in the railway investigations, graph 28 suggests that there is a structural difference between expansion and contraction of capacity. A separate investigation, excluding the years 1932–1936, seems worth while. (The limited number of contraction observations does not permit the determination of a separate regression for the contraction behaviour). The result is given in tables IV.3.3 and IV.3.4.

TABLE IV.3.3

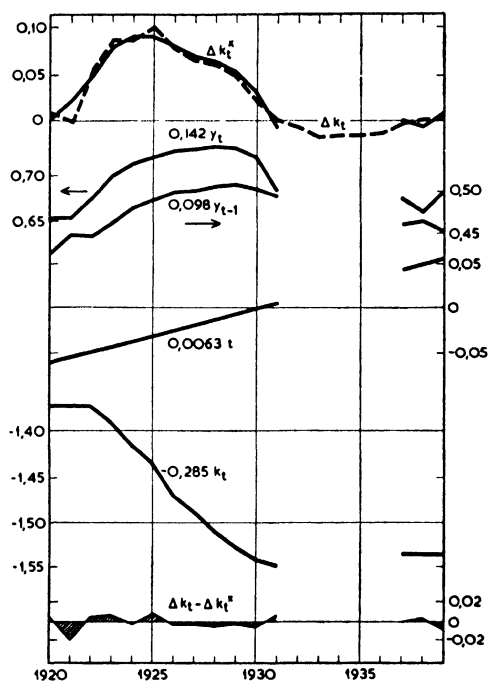
$a = 0,166$	$\bar{\alpha} = a$ for $0 \leq \xi < 1$
$b = 0,186$	$\bar{\beta} = b$ for $0 \leq \xi < 1$
$c = 0,020$	$\bar{\gamma} = c$ for $0 \leq \xi < 1$
$l = 0,776$	$\bar{\lambda} = l$ for $0 \leq \xi < 1$
$R^2 = 0,9224$	$\bar{\varepsilon}_L$ a priori = 0,8

TABLE IV.3.4

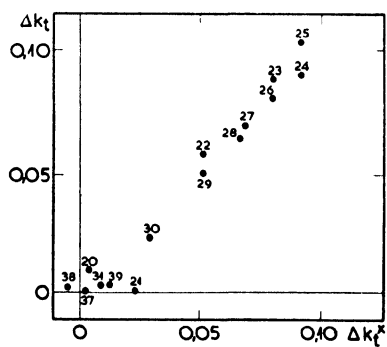
$a = 0,142$	$\bar{\alpha} = a$ for $0 \leq \xi < 1$
$b = 0,199$	$\bar{\beta} = b$ for $0 \leq \xi < 1$
$c = 0,022$	$\bar{\gamma} = c$ for $0 \leq \xi < 1$
$l = 0,715$	$\bar{\lambda} = l$ for $0 \leq \xi < 1$
$R^2 = 0,9313$	$\bar{\varepsilon}_L$ a priori = 1

There is indeed a considerable difference between these estimates and those for the whole period inclusive the period of disinvestment in capacity 1932–1936. The speed-of-adjustment-coefficient $\bar{\lambda}$ is considerably lower and $\bar{\alpha} + \bar{\beta}$ considerably higher when the contraction period is excluded. This means that the adjustment of capacity to changes in output is faster.

A graphical illustration of this result is given in graphs 29 and 30, where Δk_t^* is calculated for $\bar{\alpha} = 0,142$, $\bar{\beta} = 0,199$, $\bar{\gamma} = 0,022$, $\bar{\lambda} = 0,715$, corresponding with $\Delta k_t^* = 0,142 y_t + 0,098 y_{t-1} + 0,0027 t - 0,285 k_t$.



Graph 29



Graph 30

IV.4. STEEL INDUSTRY

A. *Capacity and output*

The steel industry is one of the oldest industries providing annual capacity data. Our statistical analysis refers to capacity K and output Y of open-hearth steel, which is about 90 % of the total capacity and output of steel. Capacity of open-hearth steel can easily be measured, it is determined by the number of open-hearth furnaces available and the number of tons with which each can be charged.

The American Iron and Steel Institute provides the total steel capacity measured by the tons of steel which can be produced by the open-hearth process and that which can be produced by the Bessemer process. In 1925 however, the definition of capacity was changed. Before 1925 the capacity figures included the capacity of plants which had long been idle. After 1925 the capacity of steel plants which had been long idle and were not likely to resume operations again were excluded. This caused a discontinuity in the data.

An inspection of the data, however, shows that this probably did not materially affect the series of open-hearth capacity. It did affect the Bessemer-steel capacity much more, unlike the data on open-hearth capacity, those of Bessemer capacity showed a strong fall in 1926 as a consequence of excluding plants which had been long idle. In restricting the analysis to open-hearth capacity and output, therefore a closed series of figures 1921 – 1941 could be used.

B. *Statistical investigations*

The results of the calculations is given in tables IV.4.1 and IV.4.2.

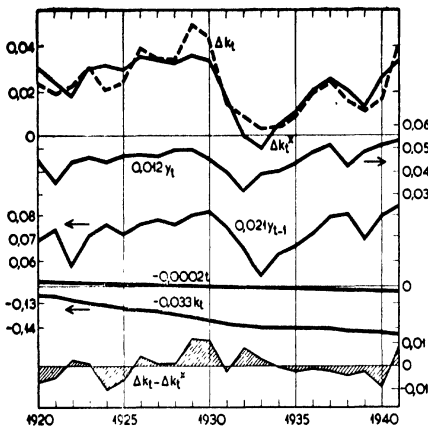
Graphs 31 and 32 give an illustration of the results; Δk_i^* was calculated for $\bar{\alpha}=0,012$, $\bar{\beta}=0,033$, $\bar{\gamma}=-0,007$ and $\bar{\lambda}=0,967$, corresponding with the equation $\Delta k_i^*=0,012 y_t + 0,021 y_{t-1} - 0,0002 t - 0,033 k_t$.

TABLE IV.4.1

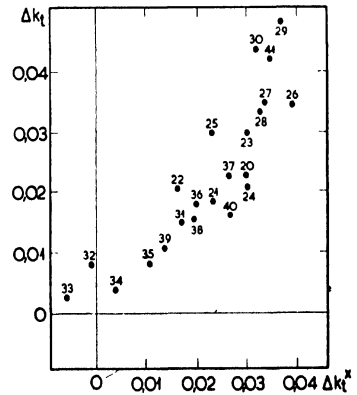
$a = 0,012$	$\bar{\alpha} = a$ for all values of ξ
$b = 0,033$	$\bar{\beta} = b$ for all values of ξ
$c = -0,007$	$\bar{\gamma} = c$ for all values of ξ
$l = 0,967$	$\bar{\lambda} = l$ for all values of ξ
$R^2 = 0,7208$	$\bar{\varepsilon}_L$ a priori = 1

TABLE IV.4.2

Direct estimates	Consistent estimates					
		$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,012$	$\xi = 0$	0,964	0,013	0,032	0,000	a priori = 0,8
$b = 0,031$	$\xi = 0,4$	0,964	0,012	0,031	0,000	a priori = 0,8
$c = 0,000$	$\xi = 0,8$	0,963	0,012	0,031	0,000	a priori = 0,8
$l = 0,962$	$\xi = 0,95$	0,962	0,012	0,031	0,000	a priori = 0,8
$R^2 = 0,6814$						



Graph 31



Graph 32

An inspection of graph 32 again suggests a structural difference between expansion and contraction. A separate calculation therefore is made, excluding the "contraction" years 1932 and 1933. The number of observations of contraction behaviour is far too small for a more advanced statistical test of the suggested structural difference.

The result of this calculation is as follows:

$$\left. \begin{array}{ll} \bar{\alpha} = 0,020 & \text{for } 0 \leq \xi < 1 \\ \bar{\beta} = 0,044 & \text{,,} \quad \text{,,} \\ \bar{\gamma} = -0,005 & \text{,,} \quad \text{,,} \\ \bar{\lambda} = 0,955 & \text{,,} \quad \text{,,} \end{array} \right\} \bar{\varepsilon}_L \text{ a priori} = 1.$$

$$R^2 = 0,7589$$

As was expected, $\bar{\alpha}$ and $\bar{\beta}$ are higher and $\bar{\lambda}$ is somewhat lower when compared with the results given in table IV.4.1. This is in accordance with the suggestion that the speed of adjustment of capacity to fluctuation in output is higher in the period of expansion of capacity.

IV.5. PETROLEUM REFINING

A. Capacity and output

Data on the capacity of petroleum refining are available. Capacity and production of refineries is measured by their input of crude oil. The rated capacity is reported by the barrels of crude oil that can be processed per day, both for operating plants and for shut-down but not yet abandoned or dismantled plants.

In the statistical analysis, K_t represents the mid-year total capacity, as the mean of capacity at the beginning and end of the year, and this including the capacity of idle plants. Part of the idle plants are, probably, so obsolete that they will never be used again ¹⁾. The data, however, do not allow the exclusion of this part of idle capacity.

It will be evident that Y_t in this section, does not refer to the heterogeneous output of refineries; it represents the input in barrels of crude oil run to stills.

B. Statistical investigations

The results of the computations are given in tables IV.5.1 and IV.5.2.

¹⁾ Cf. ALDERFER and MICHL, *op.cit.* pp. 240-245.

TABLE IV.5.1

Direct estimates	Consistent estimates					
	ξ	$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,200$	$\xi = 0$	0,862	0,199	0,110	- 0,018	a priori = 1
$b = 0,122$	$\xi = 0,4$	0,860	0,200	0,112	- 0,018	a priori = 1
$c = - 0,018$	$\xi = 0,8$	0,848	0,200	0,121	- 0,018	a priori = 1
$l = 0,848$	$\xi = 0,95$	0,800	0,202	0,159	- 0,018	a priori = 1
$R^2 = 0,7114$						

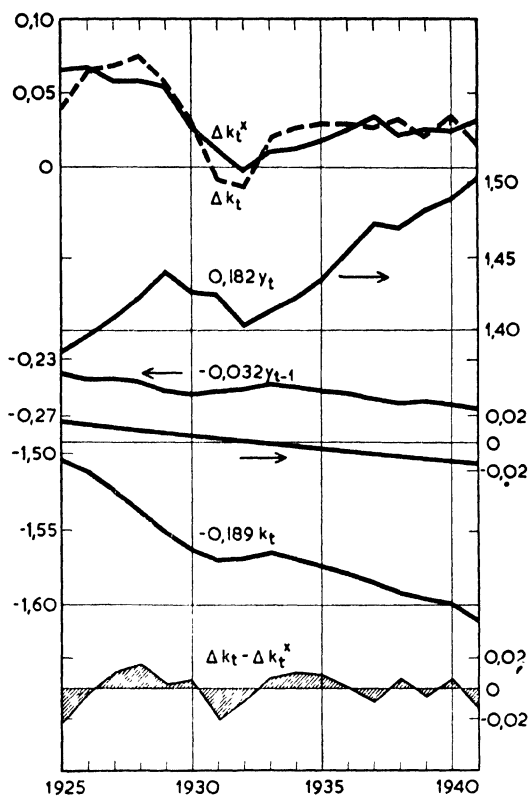
TABLE IV.5.2

Direct estimates	Consistent estimates					
	ξ	$\bar{\lambda}$	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\varepsilon}_L = \bar{\alpha} + (\bar{\beta}/1 - \bar{\lambda})$
$a = 0,182$	$\xi = 0$	0,833	0,183	0,102	- 0,011	a priori = 0,8
$b = 0,116$	$\xi = 0,4$	0,830	0,183	0,104	- 0,010	a priori = 0,8
$c = - 0,009$	$\xi = 0,8$	0,811	0,182	0,116	- 0,009	a priori = 0,8
$l = 0,811$	$\xi = 0,95$	0,740	0,179	0,161	- 0,006	a priori = 0,8
$R^2 = 0,7159$						

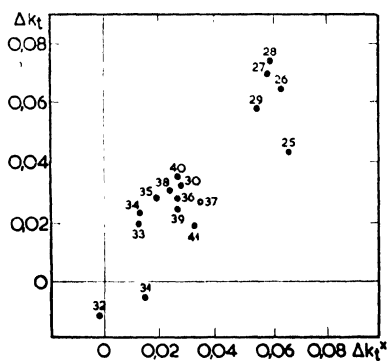
In graph 33 and 34 an illustration of the investigation is given for $\bar{\alpha} = 0,182$, $\bar{\beta} = 0,116$, $\bar{\gamma} = - 0,009$, $\bar{\lambda} = 0,811$, resulting in: $\Delta k_t^* = 0,182 y_t - 0,032 y_{t-1} - 0,0002 t - 0,189 k_t$.

IV.6. SUMMARY OF RESULTS

The investigations of this chapter have the purpose of illustrating the ideas on distributed lags and speed of adjustment as developed in Ch. II. As an application of these ideas the reaction of capacity to changes in output for a number of industries was chosen. The choice of these industries is far from a random sample of American industries, the selection was mainly determined by the easy availability of statistical data on capacity. All are industries in the early stages of production, for industries in the later stages capacity data are either non-existent or hard to interpret because of the varieties in the process of production and in the products made.



Graph 33



Graph 34

Generally, the selected industries have high capital intensities and are selling in oligopolistic markets. Expansion of capacity is made mainly by firms already in the industry, the entry of new firms being hampered by large capital requirements and institutional factors.

As to the theory about the reactions and the statistical investigations thereof, it was thought desirable to use as uniform a procedure as possible rather than to try to improve the results for each case by individual variations. The theory about the reaction of capacity to output, applied in the empirical investigations, no doubt was a simple one. As the main purpose was an application of the ideas on estimating time-shapes, no further refinement of the theory by introducing additional explanatory variables in the equations was made. The correlation was fairly high in most cases, even without the introduction of additional variables. The main influences represented in the disturbances are probably: geographical shifts in demand; the influence of profits and price expectations, in so far as these are not correlated with output, on the timing of capacity expansion; the rate of interest and availability of funds. Further the fact that the statistical figures do not exactly represent what they should, is of course, also a cause of disturbances.

The time-shapes of the reaction of capacity to changes in output, investigated in this chapter, are determined by the coefficients $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\lambda}$; $\bar{\lambda}$ is an inverted index of the speed of adjustment, if $\bar{\lambda}$ should prove to be near zero this would mean nearly an immediate adjustment of capacity to fluctuations in output, if $\bar{\lambda}$ should prove to be near 1 the speed of adjustment to fluctuations in output would be very low.

The values found for $\bar{\lambda}$ are high in all cases; $\bar{\lambda}$ is found to be in the range 0,7–0,8. For steel and for railways in the inter-war period of contraction even higher values are found, $\bar{\lambda}$ being in the range 0,90–0,95. For the coefficients of the short-run influence of fluctuations in output on capacity,

$\bar{\alpha} + \bar{\beta}$ (i.e. the percentage change in capacity during the first and second year after a one per cent change in output), correspondingly low values are found. The range for $\bar{\alpha} + \bar{\beta}$ is about 0,25 – 0,30; for steel and for railways in the inter-war period even lower values are estimated, $\bar{\alpha} + \bar{\beta}$ being about 0,05 – 0,10.

In those cases where observations of capacity-contraction were available, there was an indication of asymmetry between the time-shape of a rise in capacity and that of a decrease in capacity. The values for $\bar{\lambda}$ being higher and correspondingly $\bar{\alpha} + \bar{\beta}$ lower in the case of a decrease in total capacity. This is best shown in the case of railway freight traffic, where a reasonable number of observations of decreasing capacity were on hand; for the period of increasing capacity we found $\bar{\lambda} = 0,65 - 0,70$ and $\bar{\alpha} + \bar{\beta} = 0,35 - 0,40$; for the contraction years the result of the investigations was $\bar{\lambda} = 0,90 - 0,95$ and $\bar{\alpha} + \bar{\beta} = 0,10 - 0,15$.

As discussed in III.11 the more $\bar{\lambda}$ approaches zero the more the rigorous form of the acceleration principle will fit the facts, net-investment in capacity being related to changes in output (and not to output itself). On the other end, the more $\bar{\lambda}$ approaches 1 the more net-investment in capacity is related to the level of output itself (and not to changes in output), as supposed in the model of KALDOR. The results for the industries investigated are not favourable for the acceleration principle but are much more in accordance with the theory of KALDOR.

The lowest value for $\bar{\lambda}$ and hence the most favourable case for the acceleration principle (though not very favourable after all) is that of railway freight traffic before the first world war. This, by the way, was first given as an illustration of the acceleration principle by one of those who established the principle, J. M. CLARK.

As to the more statistical aspects of the investigations, the difference between the direct least-squares estimates a, b, c, l

and the consistent estimates $\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\lambda}$ was small for all relevant values of ξ , the autoregression coefficient of the disturbances. Especially when the correlation was high, the range for $\bar{\alpha}, \bar{\beta}, \bar{\gamma}$ and $\bar{\lambda}$ corresponding with the admitted interval $0 \leq \xi < 1$ was very narrow.

To conclude, it might be useful to remember our statement in the first chapter. The scope of this study and the purpose of the empirical investigations is to deal with one difficulty in the attempts to extend our empirical knowledge of economic reactions, viz. that of the time-shape of reactions, especially when a distributed lag of considerable length may be expected. Many other difficulties in the struggle for empirical knowledge of structural economic relations were purposely left out of consideration.

APPENDIX

SOURCES OF DATA

1. Railway Freight Traffic.

Historical Statistics of the United States 1789—1945,
A Supplement to the Statistical Abstracts of the United
States, 1949.

2. The Electric Light and Power Industry.

J. M. Gould, Output and Productivity in the Electric and
Gas Utilities 1899—1943, National Bureau of Economic
Research, 1946.

3. Portland Cement Industry.

Monthly Labor Review, October 1941.

4. Steel Industry.

Annual Statistical Report of American Iron and Steel In-
stitute.

5. Petroleum Refining.

Petroleum Facts and Figures, American Petroleum Institute.

