Lecture Notes
in Economics and
Mathematical Systems

Managing Editors: M. Beckmann and W. Krelle

340

T. Vasko R. Ayres L. Fontvieille (Eds.)

Life Cycles and Long Waves

Springer-Verlag

Vol. 236: G. Gandolfo, P.C. Padoan, A Disequilibrium Model of Real and Financial Accumulation in an Open Economy. VI, 172 pages. 1984.

Vol. 237: Misspecification Analysis. Proceedings, 1983. Edited by T.K. Dijkstra. V, 129 pages. 1984.

Vol. 238: W. Domschke, A. Drexl, Location and Layout Planning. IV, 134 pages. 1985.

Vol. 239: Microeconomic Models of Housing Markets. Edited by K. Stahl. VII, 197 pages. 1985.

Vol. 240: Contributions to Operations Research. Proceedings, 1984. Edited by K. Neumann and D. Pallaschke. V, 190 pages. 1985.

Vol. 241: U. Wittmann, Das Konzept rationaler Preiserwartungen. XI, 310 Seiten. 1985.

Vol. 242: Decision Making with Multiple Objectives. Proceedings, 1984. Edited by Y.Y. Haimes and V. Chankong. XI, 571 pages. 1985.

Vol. 243: Integer Programming and Related Areas. A Classified Bibliography 1981–1984. Edited by R. von Randow. XX, 386 pages. 1985.

Vol. 244: Advances in Equilibrium Theory. Proceedings, 1984. Edited by C.D. Aliprantis, O. Burkinshaw and N.J. Rothman. II, 235 pages. 1985.

Vol. 245: J.E.M. Wilhelm, Arbitrage Theory. VII, 114 pages. 1985.

Vol. 246: P.W. Otter, Dynamic Feature Space Modelling, Filtering and Self-Tuning Control of Stochastic Systems. XIV, 177 pages. 1985.

Vol. 247: Optimization and Discrete Choice in Urban Systems. Proceedings, 1983. Edited by B.G. Hutchinson, P. Nijkamp and M. Batty. VI, 371 pages. 1985.

Vol. 248: Plural Rationality and Interactive Decision Processes. Proceedings, 1984. Edited by M. Grauer, M. Thompson and A.P. Wierzbicki. VI, 354 pages. 1985.

Vol. 249: Spatial Price Equilibrium: Advances in Theory, Computation and Application. Proceedings, 1984. Edited by P. T. Harker. VII, 277 pages. 1985.

Vol. 250: M. Roubens, Ph. Vincke, Preference Modelling. VIII, 94 pages. 1985.

Vol. 251: Input-Output Modeling. Proceedings, 1984. Edited by A. Smyshlyaev. VI, 261 pages. 1985.

Vol. 252: A. Birolini, On the Use of Stochastic Processes in Modeling Reliability Problems. VI, 105 pages. 1985.

Vol. 253: C. Withagen, Economic Theory and International Trade in Natural Exhaustible Resources. VI, 172 pages. 1985.

Vol. 254: S. Müller, Arbitrage Pricing of Contingent Claims. VIII, 151 pages. 1985.

Vol. 255: Nondifferentiable Optimization: Motivations and Applications. Proceedings, 1984. Edited by V.F. Demyanov and D. Pallaschke. VI, 350 pages. 1985.

Vol. 256: Convexity and Duality in Optimization. Proceedings, 1984. Edited by J. Ponstein. V, 142 pages. 1985.

Vol. 257: Dynamics of Macrosystems. Proceedings, 1984. Edited by J.-P. Aubin, D. Saari and K. Sigmund. VI, 280 pages. 1985.

Vol. 258: H. Funke, Eine allgemeine Theorie der Polypol- und Oligopolpreisbildung. III, 237 pages. 1985.

Vol. 259: Infinite Programming. Proceedings, 1984. Edited by E.J. Anderson and A.B. Philpott. XIV, 244 pages. 1985.

Vol. 260: H.-J. Kruse, Degeneracy Graphs and the Neighbourhood Problem. VIII, 128 pages. 1986.

Vol. 261: Th.R. Gulledge, Jr., N.K. Womer, The Economics of Made-to-Order Production. VI, 134 pages. 1986.

Vol. 262: H.U. Buhl, A Neo-Classical Theory of Distribution and Wealth. V, 146 pages. 1986.

Vol. 263: M. Schäfer, Resource Extraction and Market Structure. XI, 154 pages. 1986.

Vol. 264: Models of Economic Dynamics. Proceedings, 1983. Edited by H.F. Sonnenschein. VII, 212 pages. 1986.

Vol. 265: Dynamic Games and Applications in Economics. Edited by T. Başar. IX, 288 pages. 1986.

Vol. 266: Multi-Stage Production Planning and Inventory Control. Edited by S. Axsäter, Ch. Schneeweiss and E. Silver. V, 264 pages. 1986.

Vol. 267: R. Bemelmans, The Capacity Aspect of Inventories. IX, 165 pages. 1986.

Vol. 268: V. Firchau, Information Evaluation in Capital Markets. VII, 103 pages. 1986.

Vol. 269: A. Borglin, H. Keiding, Optimality in Infinite Horizon Economies. VI, 180 pages. 1986.

Vol. 270: Technological Change, Employment and Spatial Dynamics. Proceedings 1985. Edited by P. Nijkamp. VII, 466 pages. 1986.

Vol. 271: C. Hildreth, The Cowles Commission in Chicago, 1939–1955. V, 176 pages. 1986.

Vol. 272: G. Clemenz, Credit Markets with Asymmetric Information. VIII, 212 pages. 1986.

Vol. 273: Large-Scale Modelling and Interactive Decision Analysis. Proceedings, 1985. Edited by G. Fandel, M. Grauer, A. Kurzhanski and A.P. Wierzbicki. VII, 363 pages. 1986.

Vol. 274: W.K. Klein Haneveld, Duality in Stochastic Linear and Dynamic Programming. VII, 295 pages. 1986.

Vol. 275: Competition, Instability, and Nonlinear Cycles. Proceedings, 1985. Edited by W. Semmler. XII, 340 pages. 1986.

Vol. 276: M.R. Baye, D.A. Black, Consumer Behavior, Cost of Living Measures, and the Income Tax. VII, 119 pages. 1986.

Vol. 277: Studies in Austrian Capital Theory, Investment and Time. Edited by M. Faber. VI, 317 pages. 1986.

Vol. 278: W.E. Diewert, The Measurement of the Economic Benefits of Infrastructure Services. V, 202 pages. 1986.

Vol. 279: H.-J. Büttler, G. Frei and B. Schips, Estimation of Disequilibrium Models. VI, 114 pages. 1986.

Vol. 280: H.T. Lau, Combinatorial Heuristic Algorithms with FORTRAN. VII, 126 pages. 1986.

Vol. 281: Ch.-L. Hwang, M.-J. Lin, Group Decision Making under Multiple Criteria. XI, 400 pages. 1987.

Vol. 282: K. Schittkowski, More Test Examples for Nonlinear Programming Codes. V, 261 pages. 1987.

Vol. 283: G. Gabisch, H.-W. Lorenz, Business Cycle Theory. VII, 229 pages. 1987.

Vol. 284: H. Lütkepohl, Forecasting Aggregated Vector ARMA Processes. X, 323 pages. 1987.

Vol. 285: Toward Interactive and Intelligent Decision Support Systems. Volume 1. Proceedings, 1986. Edited by Y. Sawaragi, K. Inoue and H. Nakayama. XII, 445 pages. 1987.

Vol. 286: Toward Interactive and Intelligent Decision Support Systems. Volume 2. Proceedings, 1986. Edited by Y. Sawaragi, K. Inoue and H. Nakayama. XII, 450 pages. 1987.

Vol. 287: Dynamical Systems. Proceedings, 1985. Edited by A.B. Kurzhanski and K. Sigmund. VI, 215 pages. 1987.

Vol. 288: G.D. Rudebusch, The Estimation of Macroeconomic Disequilibrium Models with Regime Classification Information. VII, 128 pages. 1987.

Vol. 289: B.R. Meijboom, Planning in Decentralized Firms. X, 168 pages. 1987.

Vol. 290: D.A. Carlson, A. Haurie, Infinite Horizon Optimal Control. XI, 254 pages. 1987.

Vol. 291: N. Takahashi, Design of Adaptive Organizations. VI, 140 pages. 1987.

Lecture Notes in Economics and Mathematical Systems

Managing Editors: M. Beckmann and W. Krelle

340

T. Vasko R. Ayres L. Fontvieille (Eds.)

Life Cycles and Long Waves

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo Hong Kong

Editors

Prof. Dr. Tibor Vasko
Prof. Dr. Robert Ayres
International Institute for Applied Systems Analysis (IIASA)
Schloßplatz 1, A-2361 Laxenburg, Austria

Prof. Dr. Louis Fontvieille
Director of Research
Centre National de Recherche Scientifique, University of Montpellier
39, rue de l'Université, F-34060 Montpellier, France

ISBN 3-540-52473-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-52473-8 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2142/3140-543210 – Printed on acid-free paper

Foreword

Life cycles and long waves are based on two different concepts. In spite of this, they both imply periodic – even harmonic – motions or changes. Both "models" imply that fluctuations in economic growth are not merely random individual incidents in history, but components of the development of the system as a whole.

The life-cycle concept states that a system (a product) tends to follow the same phases as that of the life of an individual: birth, growth, maturity, decline, and death. Biology and ecology provide numerous examples of this type of development. In social sciences the concept was first applied at the product level and later to firms and industries. Following that, it was applied to the behavior of individuals whose expenditure and saving patterns varied from generation to generation (Modigliani).

Research on long waves dates back to the nineteenth century. Since that time there has been repeated interest in the subject. This interest received a new impetus from the work of Schumpeter which attributed a key role to radical technological innovations. Further research focused on the frequency of innovations and their relationship to economic growth, investment, and employment.

Several different approaches make up the various growth theories that exist, of which some run parallel to each other while others are incompatible. However, an examination of these theories can contribute to a better understanding of the processes that generate economic change. This was the underlying idea on which a workshop was based that was held in Montpellier, France, in July 1987. It was jointly organized by the Centre Régional de la Productivité et des Etudes Economiques (CRPEE) and IIASA. This workshop provided a forum for an interesting exchange of views between scientists from several disciplines. This volume contains a selection of the papers presented at the meeting that illustrates both concepts in various forms.

R.H. Pry
Laxenburg, Austria
December 1989

Introduction

The selected papers presented in this volume are from a workshop on long waves and life cycles held in Montpellier, France, July 1987. It was intended to focus on possible relationships between so-called "long waves" in the economy and (shorter) "life cycles" in industry and technology. This book is of course only a small "window" into the wide and interesting range of long-term dynamic processes discussed at the workshop. Other topics presented at the workshop included cultural developments (Korpinen), national evolution (Wibe), education (Conus), ecological systems (Godron), social identity (Mallman), and problems of technology and finance (Boccara). There were also interesting methodological papers presented on general systems theory (Miller and Passet).

Concepts of Long Waves and Life Cycles

When one studies the relation of two phenomena, especially such loosely defined ones as long waves and life cycles, it is necessary to define them first or at least to describe the current knowledge on them. There are several papers devoted to this task.

It is possible to consider the paper by Yakovets as devoted to long waves in spite of the fact that he does not explicitly use this term. Instead, he speaks only of cycles and revolutions. He distinguishes between scientific, invention, innovation, technological, investment, and educational cycles. Some could be loosely related to cycles known from the works of Schumpeter, Kuznets, Juglar, and of course Kondratieff. Yakovets discusses three degrees of impact of these cycles: generations of the same technology; new scientific and technological directions; and, most important, scientific–technological revolutions. This interpretation of events helps to accommodate some of the theories developed in the late sixties and the seventies in the USSR on the scientific–technological revolution. Yakovets stresses specifically the human factor – education. After reading his paper one could almost say that Soviet social scientists have come full circle since Kondratieff. Once again, they are emphasizing dynamic effects of economic and social development processes.

The paper by Escudier illustrates the methamorphosis of the notion of long waves (in the economy) in Western economic thought. His presentation is valuable because it shows what semantic evolutions can occur (and did actually happen) when notions such as "long cycle" diffuse across various cultural and language borders and through different disciplines. He ends by identifying basic attributes of long economic cycles, namely, the regular repetition of certain processes (albeit not as deterministic as in astronomy, for example, as stressed already by Kondratieff) and the endogenous character of these repetitive processes. Escudier (in agreement with Boccara) concludes that the long waves are transformations rather than repetitions of previous stages. This brings into the play various influences which in their turn strip the long economic cycles (according to Boyer) of any deterministic character. In other words the descending phase is not to be considered as an unavoidable consequence of the growth phase and vice versa.

The Tylecote paper continues with this train of thought. Tylecote argues that long waves in the economy evoke various social processes and these in their turn may enhance the changes (pro-cyclical) or weaken them (contra-cyclical). He describes some possible causal relations of population growth and savings. He also discussed the formation of policymakers' attitudes and their direct or mediated feedback effect on the economy. These effects may not be significant enough to cause a downswing or upswing when considered alone, but create an interesting "fine structure" of forces influencing in their totality the dynamics of economic processes. These generational factors are important in explaining differences and peculiarities among regions and national economies.

To a certain extent Sinibaldi's paper can be considered as an illustration of the generational ideas expressed in Tylecote's paper. Sinibaldi tries to interpret the events of the last two decades, especially deficit spending, inflation, and interest rates, in terms of long-wave theory. He argues, for example, that the relation of inflation and interest rates shows a "hysteresis effect" because in the inflation-growth phase the nominal interest rates lag behind inflation while in decline they stay higher for a longer period of time. This seems to be confirmed by economic data from the seventies and early eighties. Sinibaldi compares the measures taken by decision makers during the late twenties and early thirties with the handling of crises in the seventies. He argues that the Keynesian solution worked much better than the deflationary policies practiced 50 years ago. These measures were sometimes taken by policymakers without help from economists. According to Sinibaldi this proves that, when it comes to practice, mainstream economic theory does not always have guaranteed solutions.

This is in agreement with Rosegger's paper. Rosegger presents a detailed analysis of life-cycle theory in industry and trade with special regard to both mainstream, neoclassical theory of technological change and various other theories including long economic waves. He concludes (in line with Nelson and Winter) that the technological gap between them remains in spite of many attempts to bridge it. Rosegger continues with a thorough and well-documented description of how the concept of life cycle in industry developed and discovers that researchers first identified the decisive factors of life cycle (technology, resources, inter-industry linkages) 60 years ago. These factors are still the subject of many studies. He also points out the contradictions one runs into when studying real data. When real-life data is aggregated and smoothed to eliminate "noise," one can also eliminate critical phenomena. When detailed case studies are performed, on the other hand, one faces the problem of generalization. Rosegger addresses also the main topic of the workshop, cross-fertilization among disciplines, and finds that simplistic analogues might run from dangerous to irrelevant but also that several disciplines such as psychology and organizational behavior might bring rewarding inputs into life-cycle research.

In the paper by Menshikov and Klimenko an attempt is made at finding simple formal interpretation of the life-cycle concept and its relation to economic variables like return on the capital stock per unit of time and expected accumulated returns. They show that even simple models can bring some insight to the dynamic of processes. They also introduce a more complex set of eight matrix equations which, when tested on US input-output coefficients, generated cycles with various periods. The equations do not yet allow for manipulation to select cycles with particular periods.

Another way to describe the dynamic process in the economy is by a system of differential equations for growth rate of labor productivity and growth rate of capital intensity (capital stock per manhour). Solutions to these equations also show cyclic behavior with damping determined by structural coefficients. It is possible to use these simple equations to test some hypotheses.

In real life the economy can also manifest very abrupt changes which can be described by the instruments of catastrophe theory based on nonlinear differential equations. Menshikov and Klimenko describe equations which manifest both static and dynamic bifurcations and can be used to model abrupt changes like "switching" between two equilibria or explosive shift. They work with an equation of the fourth order and show that the results change qualitatively with the change of coefficient at the highest order of the variable (a_1) and discuss the fitting of this model to real data.

Applications of Life-Cycles Concept

In their paper, Lee and Nakicenovic argue that life cycles might be a solution to the old question "How to manage technologies?" They illustrate the point using civil aviation, railroads, and primary energy substitution as examples. They state that in spite of the fact that some general development patterns can be identified, their use in business decisions is not simple and needs careful analysis. This paper illustrates Rosegger's view that even when life-cycle theories are not part of the mainstream economic theory, they play an important role in business decisions.

A logical continuation of these ideas is presented in Brooks's paper which reports on an interesting meeting in Washington, DC. He describes the views of three well-known researchers in the field – Utterback, Teece, and Ergas – on the technology life-cycle concept and its use on the level of productive unit, firms, and national economy. Utterback introduces finer categories of life-cycle phases (growth, maturity, decline) when he distinguishes fluid pattern, transitional pattern, dominant design, and specific pattern when the product line in question becomes highly standardized. This is the time when the interest shifts from product design to process design. He discusses the peculiarities of each phase.

Brooks further elaborates on Teece's ideas dealing with the question of why many innovators are pushed out of the market by skillful imitators. Here availability of the so-called complementary assets plays a significant role. These assets help to bring the innovative product efficiently and quickly to the market. These assets are manufacturing capabilities and effective distribution-maintenance-service channels. The latter asset is an important component of national strategies as Ergas sees them. He distinguishes three types of strategies:

(1) Heroic, high-tech, government supported.
(2) Diffusion strategy which emphasizes the high value-added of product spectrum.
(3) Combination of both strategies.

Users of these strategies include the USA, the Federal Republic of Germany, and Japan.

Antonelli makes use of data from the archives of the Italian Association of cotton weavers to test an interesting hypothesis. He argues that the S-shaped curve of innovation diffusion is in reality the result of more subtle processes than the "profitability pull" of standard theory. During the diffusion, the innovative product (and or process) is adopted by more and more firms and while the early adopters can guess on the profitability, the later entrants can base their decision on accumulated experience. So "learning by doing" enhanced by "learning by using" play a joint role. Those entering the process late may face the early adopters who managed to capitalize on the innovation and may force latecomers out of the market. In this situation, to be a small firm (and therefore more adaptive) does not seem to be a particular advantage, at least as the data from the diffusion of shuttle-less looms in Italy confirms. This case study shows that simple epidemic model of innovation may ignore events important for decision making at the firm level.

In their paper, Yu and He try to identify factors that cause substitution of one energy technology for another based on analysis of cost and benefits to both business and individual users of this technology. To get the whole picture one has to consider not only direct and indirect costs, but also changing user preferences. This cannot be done without understanding the development of consumption as a function of the affluence of the population. The authors illustrate this development on US data (measured by per capita GNP). They try to explain, for example, why the transitions in residential heating show similar patterns for many countries and why the preferences of users went from primary energy sources with lower direct cost to those with higher cost.

A similar situation occurs in industry. Yu and He present a textbook example of technology life cycle and transitions in industrial motive power. Again, secondary and indirect benefits played major roles in the transition from steam power to electric motors. This all takes place, of course, against the background of general economic development.

An analysis of similar transitions documented on data from steel and coal industries is offered by Grübler. He concentrates on turning points in economic developments and correlates them with large technological changes in two key industries. Grübler demonstrates his ideas quantitatively, for example, through the saturation phase of total output, market share of given technology and/or state, productivity indicators. He also analyzes the share of individual countries in the total world production. His approach presents a mature methodology, part of which was developed and tested at IIASA.

His idea of clustering of saturations of leading technologies, markets, and products is interesting. This clustering has its impact on other economic variables (prices) and also on social processes. These ideas are an extension of existing hypotheses. Together they may better explain the relations between technological change, life cycles of key technologies, and long-term tendencies of economic development.

Whereas Grübler tries to identify turning points in economic development by analyzing empirical data series on steel and coal, Ayres constructs an analytical, formal model of optimal growth based on the nontraditional premise that production factors can be measured by their information content (defined by information theory). He defines an optimal control model for a utility function of aggregate consumption where control variables are two kinds of investments and technology (knowledge) growth rate. The optimal sequence of events exhibits four phases where the "turning points" coincide with transition from one kind of capital to the other. Price-productivity product manifests life-cycle-type pattern. This proves that the cyclical character of economic processes is quite robust, and relatively simple structures can manifest these processes.

Tchijov and Sytchova analyze input–output tables from Japan for the period 1951–1980 to find the long-term changes in total labor requirements. They assume that impact of technological development can be identified through resource saving and work with three components – labor, fixed capital, and material. They accordingly identify three types of technical progress. They also propose the hypothesis that in the past few decades the differences in total labor requirements among industries has diminished and that in this process of "equalization" those industries with higher labor requirements were reducing this requirement faster.

Changes in total labor requirements as a function of industrial output show certain trend irregularities. Labor requirements declined most in the period between 1951–1955 and in the following period somewhat less. They rose slightly between 1970–1975 and declined significantly between 1975–1980. In each period different types of technological progress were dominant starting with labor and material saving and ending with fixed capital savings (1975–1980). The data allow analysts to speculate on the causes of this development. They seem to be mainly price increases in the case of labor and material in the 1960–1980 period and fixed capital in the late seventies.

Intuitively life cycles and long waves could be more easily related through innovation-induced industries and their impact on economic growth. This was described almost 50 years ago by Schumpeter on the basis of innovation clustering. The idea evoked immediate criticism and discussion. This discussion continues today, and the state of the art is presented by Kleinknecht. [He compares and integrates the technological innovation databases of several leading scholars and concludes that there are identifiable periods with both above and below average rates of innovation.] In spite of this fact several interesting areas are left for further research. Mechanism of restructuring of technological base of capital accumulation triggered by depression, and linking innovation waves to long-term profit rate changes are examples.

A detailed view of innovation case study in the long-wave context is presented by Fontvieille and Prigent. They argue that conventional innovation studies of both new products and processes are isolated from the whole productive process: a kind of *in vitro* instead of *in vivo* process. In the real economy new fixed capital used (operated) by labor brings high profit rate, and the situation favors capital rather than workers. Workers respond by various ways that reduce the profit rate and ultimately increase the need for innovation. This creates a sequence of substitutions for new technology generations, each one increasing the profit rate in comparison with the current practice. In the paper the authors present the results of detailed research on underground transport in French mines and describe different systems of transportation and their effect on the workers.

In his well-documented paper, Nakicenovic tries to prove his dictum that economic development is a sequence of substitutions and that these substitutions for as well as saturations of old technologies can cluster. Nakicenovic tests these hypotheses on several important areas like civil aviation, cars, infrastructures, and energy. To detect ups and downs of economic development and saturation of given technologies, he analyzes several indicators like market share, total output measured in monetary units, and wholesale price indexes. Only by using this "multiple perspective" mediated by several indicators can one gain the full picture of the situation, because one cannot expect perfect synchronization of the saturation effect of individual key technologies. Nakicenovic's paper as well as that of Grübler have a strong methodological message for those who want to study long-term economic phenomena.

An interesting integrative view on the main points of the workshop are presented by Brooks. His long-term practice and deep insight helped him to formulate a scenario on how technology life cycles could cause long-waves in economic performance. A new technology after a formative stage is gaining momentum through several economic phenomena (economy of scale, learning, and a set of indirect benefits) which create a positive (reinforcing) feedback. This feedback is tight enough to counteract any interference by other technologies (even better ones) until the saturation and "externalities" do not weaken this feedback. Brooks also elaborates on boundary condition for the validity of this scenario. Brooks's idea is an interesting methodological guideline for further studies.

In this introduction we have only scanned individual papers and presented their main highlights and as mutual relationships we have probably missed many others, which readers can find for themselves. We hope to have provided readers with some help to find those papers of most interest.

Tibor Vasko
Robert U. Ayres
IIASA
Laxenburg, Austria

Papers presented at the workshop and not included in the selection include the following: P. Korpinen, "Long Cycles in Fine Arts: Towards a New Theory of Style"; S. Wibe, "Empirical Evidence of Mancur Olson's Theory of the Rise and Decline of Nations"; M.F. Conus, "An Approach to Long-Term Fluctuations in Labor Force Qualifications: The Development of the French Educational System"; M. Godron, "The Evolution of Ecological Systems and Periodicity"; C.A. Mallman, "Long Waves and Tempos of Social Identity Life Cycles"; P. Boccara, "The Original Aspects of the Present Long Phase in a Systematic and Historical Analysis of the Long Cycle: Technological and Financial Problems"; R. Passet, "From Fluctuations to a New Future in the Making: Long-Term Economic Forecasting"; A.V. Poletayev, "Profits and Long Waves"; I.M. Saveljeva, "Long Waves and Labor Movement"; J.E. Cassiolato and R.R. Sampaio, "High Technologies, Long Waves, and the Newly Industrialized Countries"; J. Georgelin, "French Inventions in the 19th Century: Their Relation to the International Expositions"; S. Rasmussen, E. Mosekilde, and E.R. Larsen, "Innovation Bunching and the Socio-Economic Life Cycles."

Contents

CHAPTER 1

Technology Life Cycles and Business Decisions

Thomas Lee and Nebojsa Nakicenovic

1.1. Introduction

Solutions and problems are two inseparable partners in real life, in personal affairs, in business, in science and technology, and in national and international politics. Most of the time, problems search for solutions. Occasionally, a solution may be searching for problems.

In business, life is filled with problems searching for solutions. Businessmen are so eager for new promising solutions that they often overreact. When they find that the promise is not there, they drop the "solution" like a hot potato. The rise and fall of "Operational Research" is a shining example.

A major and perhaps the most difficult problem faced by business executives is: how to manage technological change? This problem has been searching for a solution since the industrial revolution; but none has been found. However, general patterns of regularities have been observed, e.g., production capacity for a mature product tends to move gradually from high-wage to low-wage countries, which eventually become exporters to the high-wage countries.

A closely related phenomenon associated with mature industries is the excess capacity problem. Today, there is excess capacity in petroleum refining, steel, aluminum, automobiles, traction equipment, farm equipment, steam turbines, generators, transformers, switchgear, shipbuilding, textiles, commodity chemicals, semiconductors – the list is long. How did business executives let that happen? Threats of excess capacity are on the horizon even for high-tech products such as commercial airliners, telecommunication equipment, and computers

Reprinted from *International Journal of Technology Management*, 3(4):411–426, 1988. Published by Inderscience Enterprises Ltd., World trade Center, 110 Avenue Louis Casii, Case Postel 306, CH 1215, Geneva Airport, Switzerland.

in all sizes except the very largest. There is no indication today that industries not yet affected are aware of this danger enough to take actions before it is too late.

When technologies compete for market share, there seems to be also a regular pattern, i.e., the process follows an S-curve (often approximately the mathematical convenient logistic curve). Fisher and Pry (1971) at General Electric Co. demonstrated this 26 years ago for binary competition (two technologies competing against each other); Marchetti and others (1979) extended the concept to multi-component systems and successfully applied the methodology to energy substitution and a number of other cases.

These "regularities" have led to the suggestion that there may be indeed a solution to the difficult question "how to manage technologies," i.e., the concept of technology life cycles. The underlying belief is that technologies behave somewhat like biological systems. There are embryonic, adolescent, and mature stages (*Figure 1.1*). Behavior or characteristics, such as total market, technical performance, cost, market share, and industry structure, vary from stage to stage along the S-curve. Recognizing the need to manage differently for different stages can be a key to success while ignoring it may invite disaster. Foster (1986) has given a number of case histories of business failures due to either ignoring or disbelieving the S-curve phenomenon in technological innovation, e.g., how American Viscose lost out to DuPont in the competition between rayon and nylon for tire cords and how DuPont subsequently lost out to Celanese in the competition between nylon and polyesters for the same market.

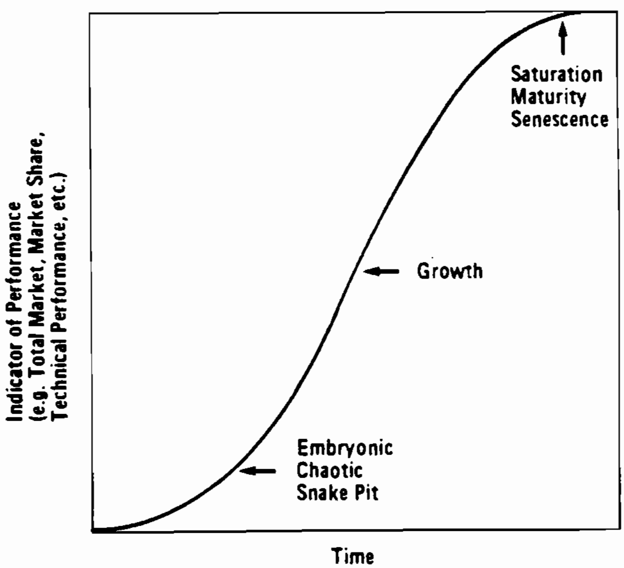


Figure 1.1. Improvement of technological performance in time shown as an S-shaped growth path indicating three life cycle phases – the embryonic introduction, growth, and saturation phase.

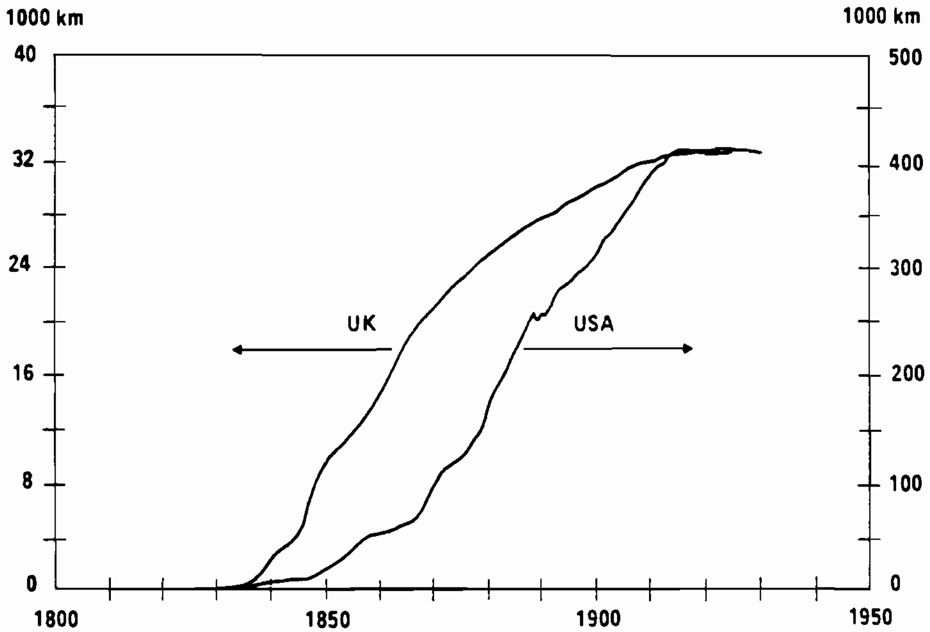


Figure 1.2. Growth of railway network length in the UK (left) and the USA (right). While the growth path follows an asymmetric S-shaped curve in the UK, in the USA it is more symmetric with a gradual increase especially between the 1830s and the 1870s. (Sources: Grübler, 1987; Nakicenovic, 1987.)

Nevertheless, reactions to the S-curve concept are mixed. Believers tend to consider it as a forecasting tool. When they apply this concept to the case of technological diffusion, e.g., the competition between technologies for market share, they tend to leave a fatalistic impression, i.e., once a technology starts to lose market share to another technology, nothing can be done to reverse that trend. This fatalistic prediction, true or not, is hard for operating businessmen to accept. No one wants to lie down and play dead. And in history, there are enough examples to show that the trends can be reversed by product or process innovation. Foster makes that point by using the S-curve for technical performance. Whenever the rate of progress slows down, reflected in the sharp increase in the cost-for-unit progress, one should be alert to new technological advances.

However, if one follows this concept blindly, disasters can happen too. Take the large steam turbine generator business: the rate of technical progress slowed down after Philadelphia Electric's Eddy Stone experience long before the oil embargo. But if utility and manufacturers continue to pump resources after 1973 to advance that technology, or any other technology for the same market niche, they would be wasting a great deal of money (of course, this did happen) because the market has disappeared. The market saturation shows up in a different S-curve, unrelated to the one for technical performance.

Disbelievers easily argue that there are exceptions to the S-curves. *Figure 1.2* shows such an exception. The growth of the railway network length in the UK follows an S-shaped path, but not a symmetric one. Instead the growth rate increased rapidly until the 1850s and then declined slowly for a period of 60 years. (In fact, the growth of the railway network in the UK follows an asymmetric S-shaped curve called the Gompertz function.) This is an interesting counterexample since the growth of railway networks in the USA and Germany followed symmetric S-shaped paths. *Figure 1.2* also shows a more symmetric growth path for the railways in the USA. Thus, disbelievers can argue that a number of S-curves are not symmetrical; therefore, symmetrical S-curves are useless as forecasting tools. Planners in the electric power industry in the early 1970s certainly did not believe S-curves at all for electricity demand growth. The consequences are well known to us: overcapacity and capital crunch.

In this paper, we address the question of "how to use S-curves" on a broad basis. We will look at S-curves for total market, market share, as well as technical performance. We will show that when we look at a set of curves for the same market niche, a great deal can be learned for business planning purposes. We will discuss three specific uses for S-curves:

- Investment and R&D Decisions.
- Contingency Planning.
- Implications of Multiple S-curves.

1.2. Application to Investment and R&D Decisions

Investment in capacity expansion at the wrong stage of a life cycle can often lead to excess capacity, depressed prices, and reduced profitability. History is filled with such cases. We will illustrate this point with aviation, an industry which may be on the verge of such a mistake.

Figure 1.3 shows the world air transport market, all operations, measured in 10^6 passenger km/hr. On a linear plot it clearly exhibits an S-curve behavior. *Figure 1.4* is the same information on a semilog plot. [We will plot the S-curves on a semilog paper. A simple S-curve has a saturation level (K). For the variable x_1 (be it total market, cost, market share, etc.), we can express it as a fraction of the saturation level: $F = x/K$. If $F/(1\text{-}F)$ is plotted as function of time on a semilog paper, a perfect S-curve will appear as a straight line. Saturation level (K) is sometimes unknown, but it can be estimated from the data.] The fit of the data to a logistic (or S-) curve is obvious.

On *Figure 1.4*, we have also indicated the years of introduction for three landmark aircraft, B377, B707, and B747. If the logistic trend continues, the total market will saturate at about 200×10^6 passenger km/hr. The inflection point apparently occurred in 1977, when the market reached 50% of the saturation value.

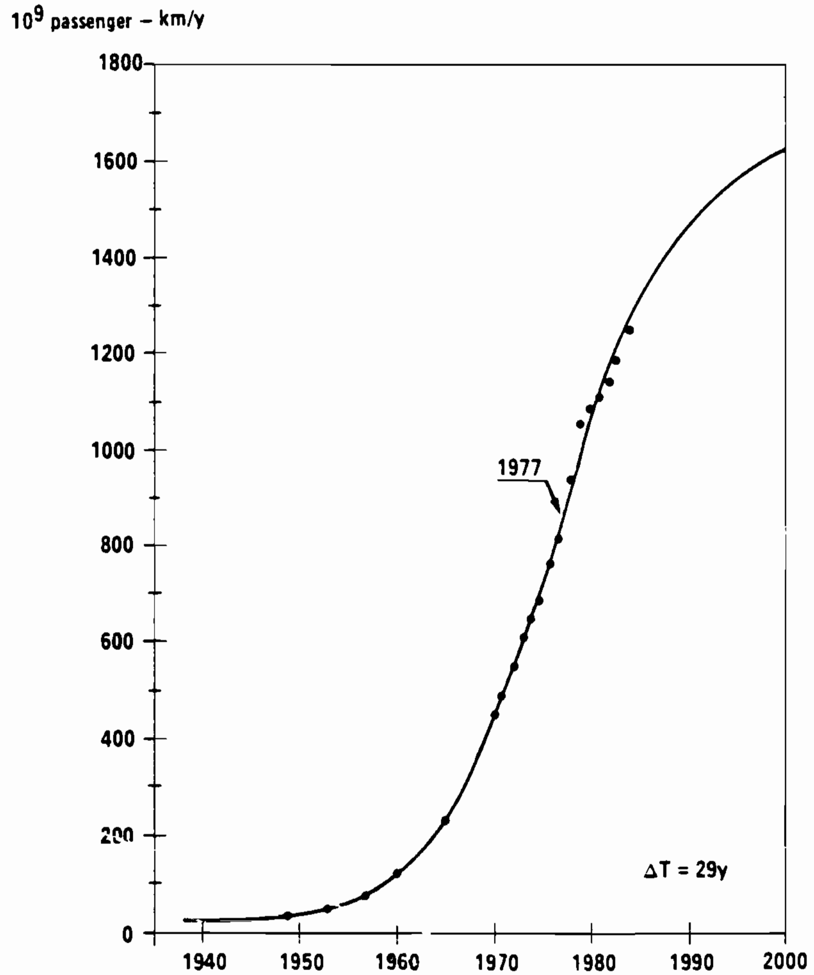


Figure 1.3. Growth of world air transport in billion passenger-km/year. (Source: Nakicenovic, 1987.)

What are the predictions that can be made from this plot?

- The market will reach 90% of its "ultimate" level about the year 2000.
- ΔT, the time required for the market to go from 10% to 90% of the saturation level, for the total market growth is about three decades (29 years).

While these implications are interesting, even more can be extracted if we look at the evolution of the technical performance of passenger aircraft, shown on *Figure 1.5.* Like the air transport market, the performance of individual aircraft can also be measured in passenger km/hr. Instead of following closely a straight line, there is a band, the left line representing the performance of the

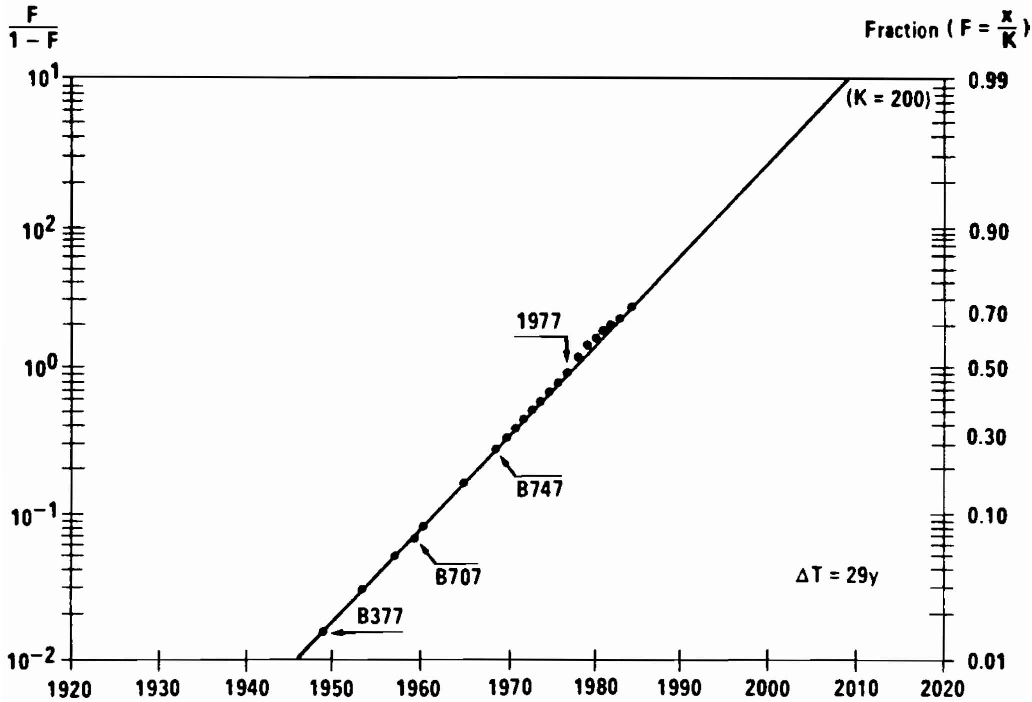


Figure 1.4. Growth of world air transport in million passenger-km/hr (from *Figure 1.3*) plotting F/(1-F) as a function of time on a semilog paper, where F is the fraction of the estimated saturation level K. In this way, results that would otherwise appear as S-shaped curves come out as straight lines, making them easier to interpret. (Source: Nakicenovic, 1987.)

best aircraft. Let us stay with the left (upper) envelope of the band. The ultimate performance appears to be 1.2×10^6 passenger-km/hr. The following are additional implications when one combines the information in *Figure 1.4* and *Figure 1.5*:

- The performance of the present 747 is almost halfway to the "ultimate" level. A stretched 747 (perhaps the planned 500 series) may be all that is needed. Certainly, this has serious implications on R&D decisions.
- The "ultimate" market can be served by 170 stretched 747s or 340 planes of the present vintage, if the aircrafts operate to capacity all the time. Of course, this is not feasible for many reasons. There are about 600 747s in service now, suggesting a 30% capacity utilization (assuming that 170 stretched 747s would be needed). What should the ultimate capacity utilization be, we don't know. But it seems obvious that in the future, the need for additional production capacity is not great. Yet at the present, both

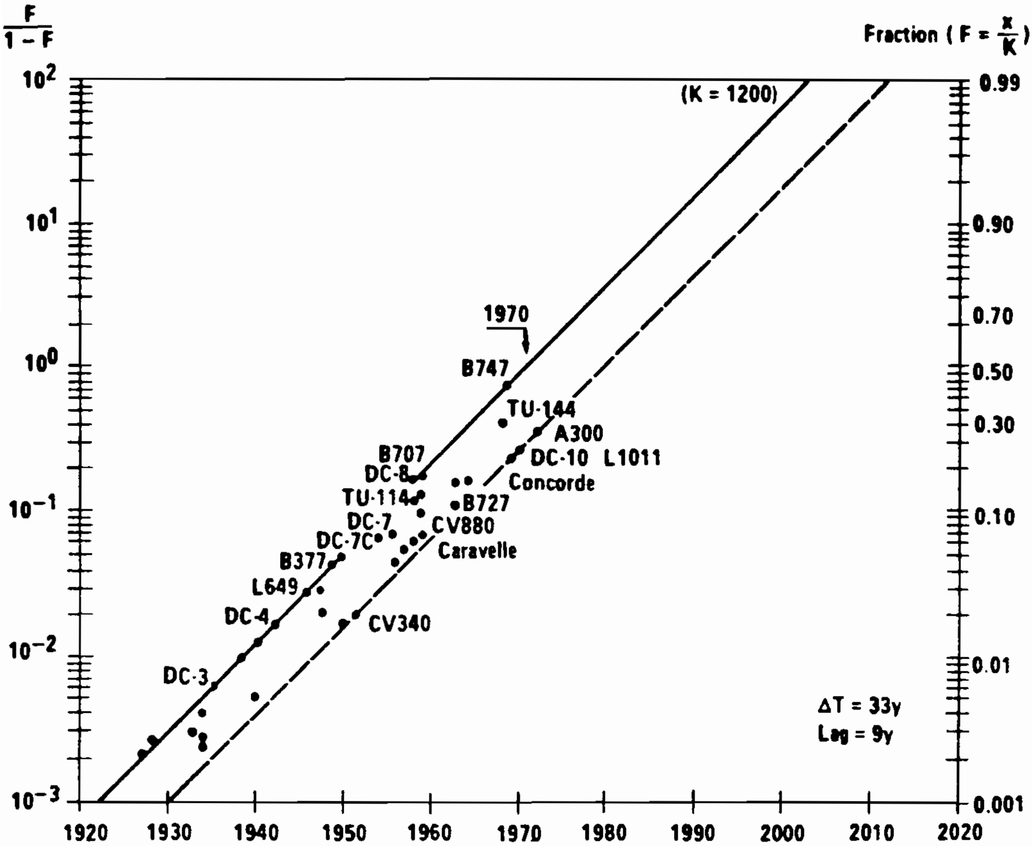


Figure 1.5. Improvement of passenger aircraft productivity in thousand passenger-km/hr plotting F/(1-F) as a function of time on a semilog paper, where F is the fraction of the estimated saturation level K. (Source: Nakicenovic, 1987.)

Japan and the EEC are trying hard to increase their penetration of the 747 market. Although they do not plan to build aircraft as large as the 747, they could increase the volatility and competition in the market by offering competitive smaller aircraft.

The above analysis is, by intention, grossly oversimplified. A more careful analysis would segregate the total market according to distances and the aircraft that serve these sectors. Yet it is not clear how a more disaggregated analysis would alter the basic conclusion that the industry as a whole may be moving into the state of overcapacity.

It is also interesting to note that ΔT for the evolution of technical performance is 33 years, not very different from that for the total market. Associated with the dynamic technical evolution of aircraft is cost improvement, which has a significant effect on market growth. At the embryonic stage, the cost improvement is easy to achieve. The market grows rapidly because of demand elasticity. When a product moves toward the mature stage, improvement in technical performance becomes more difficult or costly, the learning effect on cost slows down and so does the market growth. It is therefore not surprising that ΔT for a purely technological attribute (performance) is not very different from the ΔT for a market attribute. The detailed reasons behind the actual closeness of these two numbers in the example should be the topic of further research. The fact that they are close implies that the total number of aircraft in service has been relatively stable, especially when compared with a factor-100 increase on all operations and aircraft productivity.

If one had to make a choice between different modes of transportation for investment purposes in the 1930s, the only information on air transport was the characteristics of the Douglas DC-3. Comparing the performance, cost, and personal comfort of traveling by DC-3 with travel by railroad, one might easily prefer to invest in the latter – if the relative rates of technical change were not considered. Fifty years later, there is no direct way to travel by rail from coast to coast in the USA. Looking at *Figure 1.5*, the reason for this is clear. The young aviation technology in the 1930s has been improved by more than a factor of 100. The already mature railroad technology showed little, if any, improvement. In fact, the total length of the main tracks and thus the effective size of the network decreased by about 30% ever since the 1930s. The lesson: Neglecting the technology life cycle phenomenon and consequently also the growth potential of embryonic technologies can lead to disastrous investment decisions.

1.3. Contingency Planning: An Example

In 1975, Marchetti of IIASA studied energy substitution processes (*Figure 1.6*) and predicted that after oil, natural gas may become the dominant energy supply in the world (Marchetti and Nakicenovic, 1979). His prediction was made at a time shortly after the oil embargo and the whole world was concerned with the depletion of two valuable energy resources: oil and natural gas. President Carter declared in 1976 that the energy problem was the moral equivalent of war. The Fuel Use Act was passed to forbid the use of natural gas in industrial boilers. The country adopted the energy policy of depending on nuclear and coal to become energy-independent.

A massive energy R&D program was introduced to develop the use of coal: coal gasification, coal liquefaction, fluidized bed combustion, and a host of other technological options. Industries responded well to this call for energy independence; among them, the General Electric Co. (GE). It spent several billion dollars to acquire a major coal company – Utah International.

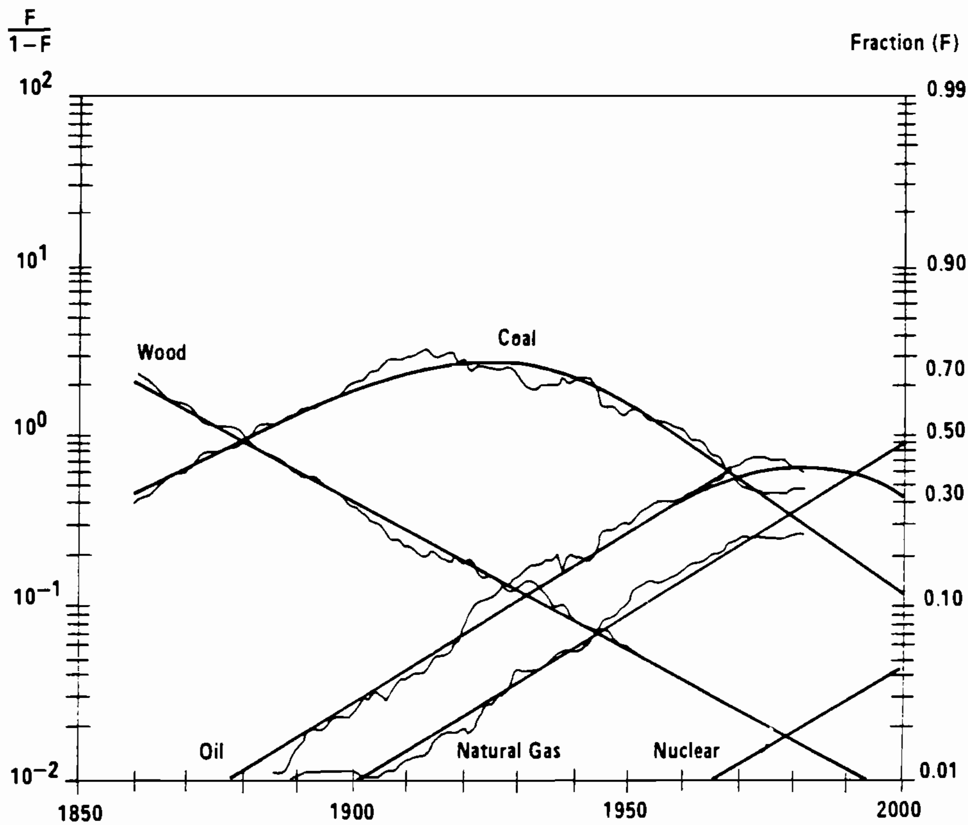


Figure 1.6. The history of global primary energy substitution from 1860 to 1983. F is the fractional market share of a given energy source in global primary energy consumption and data are shown plotting F/(1-F) as a function of time on a semilog paper. In this way, straight lines show the S-shaped growth and decline in market shares. Coal and oil curve through a saturation phase that joins S-shaped growth to S-shape senescence due to the fact that in a multicomponent substitution process not all competitors can follow S-shaped growth and senescence paths at the same time so that one competitor is saturating. Market shares of primary energy sources are projected through 2000 by making explicit assumptions about the market penetration rate of nuclear, a 1% market share in 1965 and a 5% share in 2000. (Source: Marchetti and Nakicenovic, 1979.)

It was against this background that Marchetti was invited to present his prediction to GE's executive officers. Halfway through his presentation, one of the senior vice presidents stormed out of the meeting, angrily stating: "I don't believe a word he said."

One of the authors had a subsequent conversation with that senior vice president to point out that it is futile for business people to engage in a debate on which prediction is correct. It is an endless debate. The proper questions to ask are:

- If the prediction is correct, what might the consequences be to the General Electric Co.?
- If the potential consequences are severe enough, is a contingency plan (or insurance policy) justified?

The answer to the first question was obvious: GE's steam turbine business might be in serious trouble since the bulk of it was in the large steam turbine sector, which supplied machines to large nuclear and coal-fired plants. On the other hand, the stationary gas turbine business might grow.

The answer to the second question was rather obvious too: GE should have acted to protect its stationary gas turbine business by investing in R&D and by developing a creative international strategy.

Whether that discussion had a deciding effect on GE's strategy, we will never know. In any event, GE did pursue gas turbine market leadership aggressively. In 1983, 15 years later, the large steam turbine business in the United States, for all practical purposes, collapsed. The gigantic plant in Schenectady had to depend mostly on spare parts business. At the same time, the gas turbine business prospered, not in the USA, but in international markets. GE's creative gas turbine strategy, the "manufacturing associates" arrangement, propelled GE to be the unquestionable number one supplier in the world. The dollar volume ratio of gas turbine business to steam turbine business within GE reversed almost exactly from that of 10 years ago.

We do not suggest that GE's success in the intervening 10 years in the gas turbine business was due to smart contingency planning, because it was never considered as such within GE. We report this conversation, therefore, only for the purpose of illustrating the potential importance of contingency planning.

Ironically, GE's gas turbine business had a major set back in late 1985 and 1986, due to the drop in oil prices. A significant fraction of GE's gas turbine business was in the Mideast. The drop in oil prices caused cancellations that were not expected. Shouldn't there be a contingency plan for low oil prices in GE's gas turbine division? The answer is (in retrospect) yes. However, we suspect no such plan actually existed.

Let's now return to the aircraft example. Two questions that executives and planners in the aircraft industry might debate are:

- Are *Figures 1.4* and *1.5* good forecasts for the total market?
- Is *Figure 1.5* a reliable forecast for the performance of civil aircraft?

Based on previous arguments, we suggest that, instead, the proper questions to debate are:

- If *Figures 1.4* and *1.5* are correct – knowing the behavior of the industry in general – what is the likelihood that the industry as a whole will march into a situation of excess capacity?
- If that happens, what are the implications for market share, price level, and profitability?

- Should the industry spend money on an aircraft that is more advanced than a stretched 747?

By more advanced aircraft, we mean higher productivity than that of a stretched 747 (say, higher than the planned 500 series). This means more passengers (perhaps more than 1,000 passengers capacity) or higher speeds (supersonic or hyper-sonic) or both. Emergence of more advanced aircraft could imply a new growth curve that would eventually substitute current aircraft technologies with a promise of a much larger growth potential. This situation is similar to the choice between the railroad and the DC-3 in the 1930s.

One of the authors of this paper published a paper in *Harvard Business Review* not long ago on R&D planning, Lee *et al.*, (1986). In it, he proposed the concept of "robustness" for examining the value of an R&D program against possible (even not likely) environmental shocks. Contingency planning is a way to improve the robustness of a strategic plan. Technology life cycle can be a very useful concept for that purpose.

In business, the challenge is always to develop viable management strategies in the face of imperfect information. The S-curve concept helps to reduce the imperfection.

1.4. Multiple S-Curves and Their Implications

Now we examine a few examples where the evolution cannot be described by a single S-curve. These examples shed light on the question: Is the S-curve a reliable forecasting tool?

Figure 1.7 shows the advances of capacity for the best aircraft engines on the market. The fact that there are two straight lines parallel to each other means:

- There are two S-curves (*Figure 1.8*), one sitting almost on top of the other, each describing the evolution of a technology: the bottom one for piston engines and the top one for jet engines.
- The rate of change of the technology and the characteristic time ΔT are the same for both technologies, suggesting the existence of some social and economic environment common to both technologies. In this case they serve common markets, namely, civil and military aviation.

What useful implications can be extracted from *Figure 1.7*? If one believed in the predictive value of the logistic curve, say in the year 1930, one might have concluded that the largest piston engine would saturate with a thrust of about 3,800 HP. Such a conclusion, by itself, would not have been too unreasonable. But if one used it in conjunction with multi-component substitution curves for all modes of transportation (*Figure 1.9*), one might have been led to a disastrous conclusion, based on the following logic.

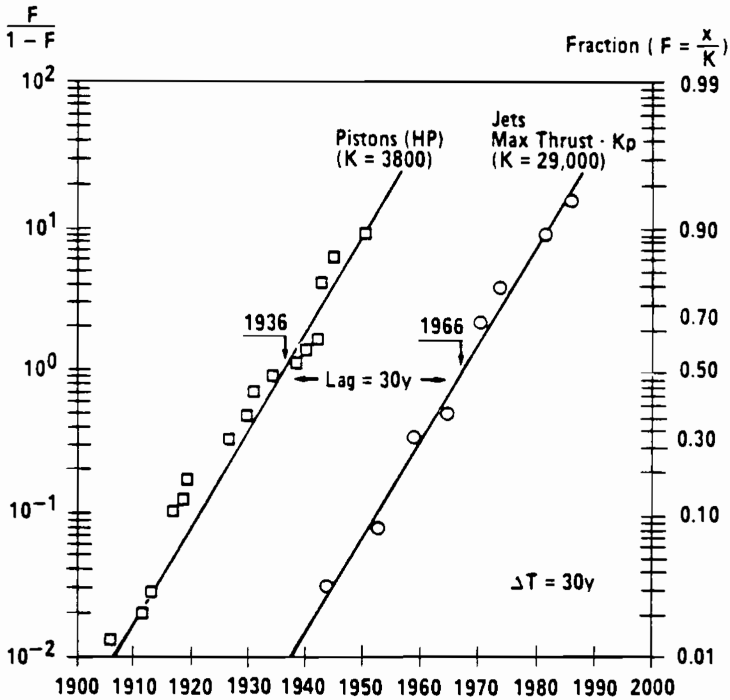


Figure 1.7. Advances of capacity for the best aircraft engines on the market, shown as two S-shaped growth pulses by plotting F/(1-F) as a function of time on a semilog paper, where F is the fraction of the estimated saturation level K. The first pulse (left) shows piston aircraft engines and the second (right) the jet engines. (Source: Grübler, 1987.)

The aviation market in 1940 was rapidly increasing (*Figure 1.9*), but the piston engine was nearing its ultimate technical performance. The "obvious" strategy to capture greater market share is the learning curve strategy. This strategy had not been formally articulated (as far as we know) by 1940, though T.P. Wright formulated in 1936 the so-called labor cost-quantity curve, on the basis of an analysis undertaken in 1922 (Wright, 1936). Actually, the Curtiss Company's price determinations for the decade after 1922 were based on the use of these curves that subsequently became well known as learning curves. To exploit that strategy, one needs to have greater production capacity for piston engines. Following this strategy would have been disastrous because of the imminent appearance of a new technology: jet engines. The fact that saturation of piston engines was approaching in the 1940s while the total market was still growing should have alerted the planners of the possible emergence of new technologies. The job of the planner was made more difficult by combining the dynamic behavior of both the market and the technology, but the risk of wrong decisions can be minimized.

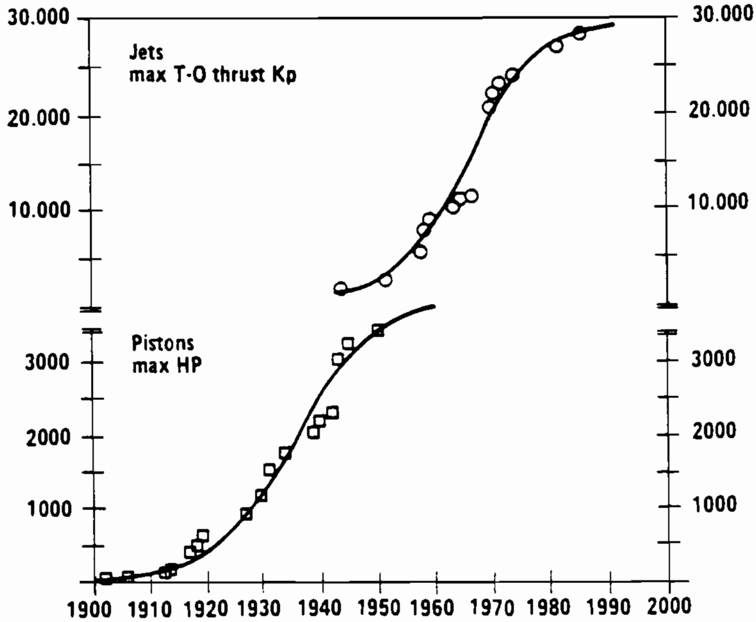


Figure 1.8. Advances of capacity for the best aircraft engines from *Figure 1.7.* The first (lower) S-shaped growth pulse shows the increase in horse power (HP) of piston aircraft engines, while the second (upper) growth pulse shows the increase of take-off thrust (in Kp) of jet engines. (Source: Grübler, 1987.)

A second example is shown in *Figure 1.10,* which shows the evolution of transmission voltages in electric power systems. Again, there are two straight lines, though with different slopes.

If a planner studied the situation in the late 1940s, he would have seen a continuing growth of electricity demand (*Figure 1.11*), at an annual rate of 7% and a "saturated" transmission voltage (Fisher, 1974). The obvious strategy would again be to exploit the learning curve. If this strategy was followed, it would have been disastrous.

What actually happened? In the early 1950s, the electric utility industry discovered that by shifting a greater percentage of its capital resources into transmission *vis-a-vis* generation, the net saving, due to lower generation needs, is quite attractive from an overall viewpoint. This might well have occurred because of saturation in the rate of improvement of power-generating technology. When this resource allocation strategy was discovered, an international competition on "who builds the highest voltage first" began. This significantly altered the economic environment and probably was responsible for the shorter ΔT (19 years) for the second line.

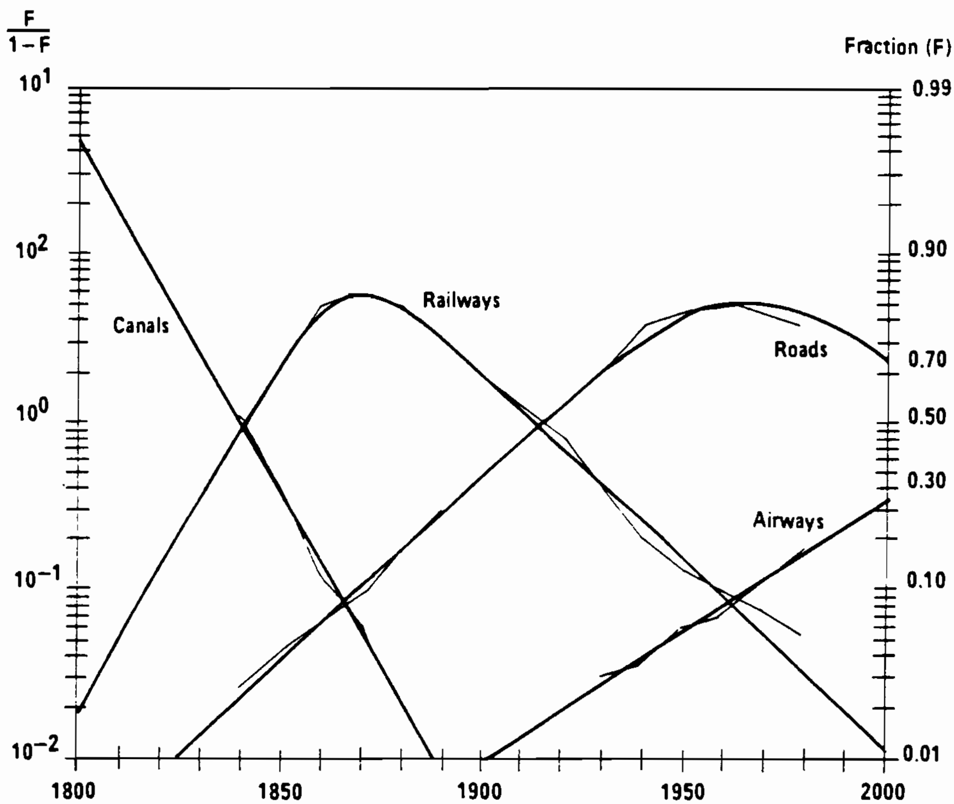


Figure 1.9. The history of transport infra-structure substitution in the United States from 1830 to 1982. Market shares of transport infrastructures are projected through 2000. F is the fractional share of a given transport infra-structure length in total length of all transport networks. (Source: Nakicenovic, 1987.)

It is interesting to note that the maximum voltage predicted by *Figure 1.11*, 819 KV, has not yet been exceeded. Some isolated attempt to develop voltages higher than 1,000 KV have not yet led to commercial application. There are indications in today's environment that power plants may change in the future from large to small, from centralized to dispersed. Thus, the need for higher transmission voltages may not appear for some time, although superconductive transmission may emerge as a new technology thereafter.

Two conclusions can be drawn from the previous examples:

- A single attribute S-curve is not a reliable or useful forecasting tool (and may even be a misleading one). The most important reason is the possibility of a second wave, due to either new technology or a new economic environment.

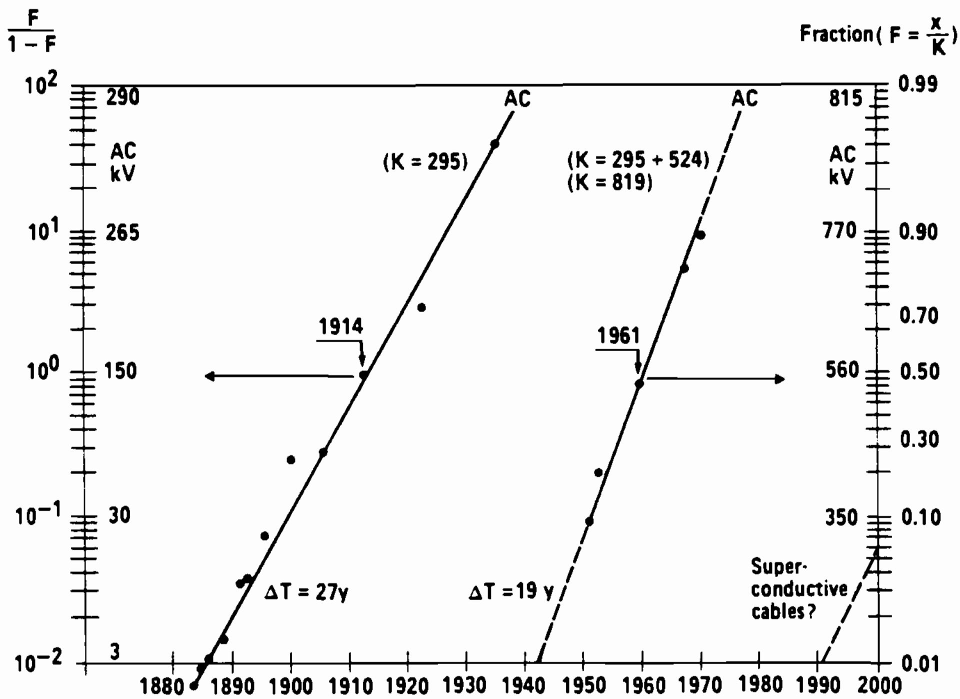


Figure 1.10. Evolution of transmission voltage, shown as two S-shaped growth pulses by plotting F/(1-F) as a function of time on a semilog paper, where F is the function of the estimated saturation level K. The first pulse (left) shows the voltage increase to K = 295 KV and the second (right) the increase to K = 819 KV. We have added a third, hypothetical increase in voltage that could be possible with the advent of superconductor cables. (Source: Grübler, 1987.)

- A multi-attribute approach to technology dynamics can be useful for business planning.

1.5. Conclusions

(1) The S-curve offers a tool for recognizing the period when excess capacity and technology saturation may become probable. This should be factored in the formulation of business strategies, e.g., the learning curve strategy may cease to be feasible.

(2) S-curves may not be reliable forecasting tools; but they can be very useful for contingency planning.

(3) Saturation of a dominating technology in a growing market may signal the emergence of a new competitor or successor technology.

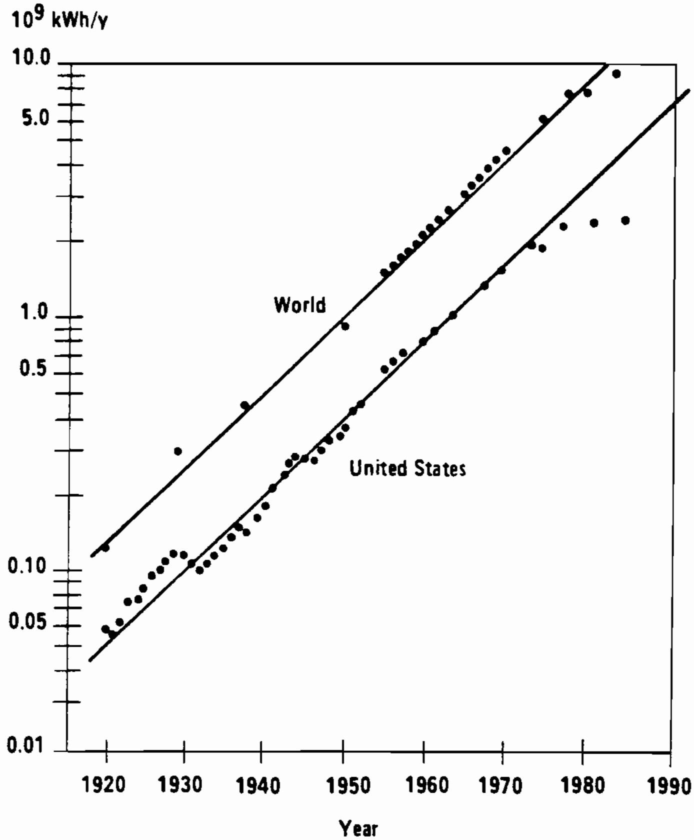


Figure 1.11. Generation of electric power over the past half century, for selected years worldwide and for the United States. The trend lines correspond to uniform growth at 7.2% per year (100% per decade). World electric power generation has been growing at about 8% per year. During the last decade both curves show signs of saturation implying an S-shaped growth process. (Source: Fisher, 1974.)

(4) A single S-curve is not as useful, for planning purposes, as a set of S-curves dealing with different aspects of the same technology and market.

In short, S-curves can be useful, but we must learn how to use them intelligently.

Acknowledgment

We want to thank Professor R.U. Ayres for his comments and discussions relating to this paper.

References

Fisher, J.C., 1974, *Energy Crises in Perspective*, John Wiley & Sons, New York, NY, USA.

Fisher, J.C. and Pry, R.H., 1971, A Simple Substitution Model of Technological Change, *Technological Forecasting and Social Change* 3:75–88.

Foster, R.N., 1986, *Innovation: The Attacker's Advantage*, Summit Books, Melbourne, Australia.

Grübler, A., 1987, private communication.

Lee, T.H., Fisher, J.C., and Yau, T.S., 1986, Is Your R&D on Track? *Harvard Business Review* January–February.

Marchetti, C. and Nakicenovic, N., 1979, *The Dynamics of Energy Systems and the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Nakicenovic, N., 1987, *Transportation and Energy Systems in the US*, WP-87-1, International Institute for Applied Systems Analysis, Laxenburg, Austria.

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

CHAPTER 2

Aspects of the Life Cycle in Industry and Trade

Gerhard Rosegger

2.1. Introduction

The story is told that Nassau Senior, who held the first chair of political economy at Oxford, once dared to offer a talk on the population problem. When Malthus heard about this invasion of his intellectual terrain, he offered the following acid comment: "It is among the disadvantages of public lectures, that the lecturer thought he was called upon to say something new, when there was nothing new to be said." I was immediately reminded of this story when Professor Ayres asked me to talk about the state of the art in life-cycle research. What could I possibly say that is new to this audience, whose collective years of work in the field I could not begin to estimate?

To make the task of saying old things at least manageable, I decided to constrain my problem by sticking to the industrial economist's perspective, instead of ranging into other social-science management concerns with life-cycle phenomena. For the noneconomists in the audience, however, I must start with an awkward confession, albeit one that will cause no surprise among the economists: what I am going to talk about has little to do with the kind of economics we teach our students, at least in the United States.

Take the typical American textbook on elementary economic theory. In the chapters on economic growth, the reader is told that – depending on which of several dozen econometric studies one believes – technological advances account for anywhere between half and three-quarters of all long-term increases in aggregate real output. But then, in the chapters on the economics of the firm, which presumably is the agent responsible for generating these advances, the reader will find scant reference to the crucial role of innovation or, for that matter, of other strategic decisions. At best, there will be a diagram or two that show how some mysterious *deus ex machina*, called "technological change," shifts production functions about until a highly desirable state of affairs, called equilibrium, is once again reached.

Some mention of Schumpeter may be made in connection with all of this, but perceptive readers are bound to be confused about the role that his ideas play in this orderly world of fully informed, optimizing firms. More advanced students of micro-theory are admitted to further secrets. They learn about the Cobb-Douglas and other forms of production function, and they learn what Harrod, Hicks, and Solow had to say about the precise nature of shifts in these functions. They are also told that there are such things as bias and learning – ideas that lend themselves to very elaborate kinds of modeling.

Friedlaender (1986, pp. 328–329) demonstrated the limitations of the theory by posing the question how, five years hence, an econometrician might measure the effects of the introduction of two recent innovations in the US automobile industry – quality-control circles and just-in-time inventories. The first presumable will increase the marginal product of labor relative to that of capital, and the second will show up as a reduction in working-capital investment. Depending on the relative impact of these two quite distinct innovations, the econometrician could end up by describing the combined effects of "technical change" as labor-saving, neutral, or capital-saving. None of these *ex post* descriptions, argues Friedlaender, would tell anything about what really happened.

It all comes down to the fact that human beings are strangely missing from this well-integrated theoretical structure. One never knows *who* makes all these important decisions, or *where* the necessary information comes from. Many economists have, of course, been bothered by this peculiar picture of the world, and so we have a more or less peaceful coexistence between what has become known as the mainstream, neoclassical theory of technological change and a host of heterodox models, among them various life-cycle theories.

This is not the place to deal with all of these, but in a cursory way it seems fair to say that certain notions about the importance of the life cycle have received more widespread attention than others. Thus, for example, Vernon's (1966) model of the product cycle in foreign trade and investment, as modified in his subsequent writings, can be found today in every text on international economics. Nevertheless, life-cycle theories have always carried the stigma of being *ad hoc* and – even more damaging – "partial-equilibrium" constructs.

It seems equally fair to conclude that more encompassing hypotheses, like those dealing with the existence of technology-driven long waves, have run into reactions ranging from agnosticism to rejection. Rosenberg and Frischtak (1983) and Mansfield (1983) may serve as illustrative examples of this attitude. Interestingly, both these critical surveys refer to the possibility that the macroeconomic climate and "those deeper-rooted forces that shape the rhythm of capitalist development" may influence the pace of technological advance. This is, of course, precisely the problem addressed in a very persuasive way by Olson (1982). To put the matter politely: at least in the United States, life-cycle research and related work have been no more than tolerated guests in the house of mainstream economic theory, even while they have been warmly welcomed by other disciplines, such as strategic planning, industrial management, and marketing.

This state of affairs would be perfectly understandable if, despite (or perhaps because of) its very austerity, the neoclassical theory of technological change had been especially helpful in bringing order and understanding to the data of the real world. Such has hardly been the case. Stoneman (1983, p. 62) concludes his excellent review of the theory of invention and innovation as follows:

> The outcome of all this is to argue that there are strong theoretical reasons to believe that from economic analysis we can make predictions as to the determinants of the rate of technological advance and the nature of technological advance in different industries, as well as saying something about the bias of technological advance at the macro-level. However, empirical support for many of these theories is not really as encouraging as one might hope.

Numerous diagnoses of the problem have been offered. These run all the way from the observation that modern economics has become a profoundly ahistorical science to more complete analyses along Kuhnian lines. For our purposes, however, it may be sufficient to accept what Nelson and Winter (1982) have argued most convincingly – that no amount of well-intended attempts at bridge-building can close the epistemological chasm between the neoclassical economics and those models that purport to make technological and other strategic decisions the focus of explanations of the economic process.

This introduction, so brief that it is almost a caricature, must not end on a negative note. First, we have to confess that the apparatus of neoclassical production theory has provided us with a vocabulary and a grammar without which our discourse would be difficult, indeed. Second, I would dissent from a summary judgment against the production–function paradigm, because it can deal very well with *induced innovations*, that is, innovations triggered by changes in relative input and output prices [see, for example, Binswanger and Ruttan (1978)]. These are, of course, at the opposite pole from Schumpeterian innovations. Third, it has provided the conceptual foundations for a host of normative theories concerning the *operational* decisions of firms. And finally, as the recent survey by Link (1987) demonstrates, many economists who in all likelihood consider themselves solidly in the mainstream have made valuable empirical contributions to our understanding of the relationship between technological changes and industrial growth.

In addition, there is some pioneering theoretical work like that of Salter (1960) and others, who attempted, with at least partial success, to incorporate innovative activity into the traditional models of firms and industries. All, of course, had to confront (or avoid) the difficulties created by the underlying assumption that firms optimize on some well-defined objective. The more recent literature on the economics of information, too extensive to cite here, may be considered another thread in these developments.

2.2. Observations

By now, the observation that the difficulties I have just sketched can be traced to the Newtonian roots of modern economic theory has become a platitude. In an early essay on the subject, Kenneth Boulding (1957, p. 9) wrote: "Economics has been, and still largely is, a sunless astronomy of commodities, seeking an ever changing equilibrium in cycles not quite as regular as those of the planets, but still moved by differential equations rather than by men." Despite the great progress of mainstream theory during the past 30 years, that judgment seems as valid now as it was then.

It does not, however, answer the crucial question: what kinds of theoretical conceptions can life-cycle research put in place of the mechanistic ones? Before turning to that question, I want to point out that the first wave of American investigators, if I may call it that, does not seem to have been beset by any doubts about the matter. To them, the driving forces and causal linkages were quite obvious. Let me cite just a few representative examples.

Arthur Burns (1934) stated the logic of the case in elegantly simple terms. He first observed that, "the conception of the indefinite growth of industries can neither be supported by analysis nor by experience," and he also postulated that there is no sound rational basis for "the notion that industries grow until they approximate some maximum size and then maintain a stationary position for an indefinite period" (p. 170). From these two premises he concluded that there is an inherent tendency "for an industry to grow at a declining rate, its rise being eventually followed by a decline" (pp. 171–172). Burns found his propositions confirmed by a heterogeneous sample of some 30 industrial-output time series for the United States. He did not discuss the nature of the processes underlying the rise and decline of individual industries in any great detail, nor did he offer any generalizations about their causes.

Earlier on, however, a set of hypotheses about the nature of these causes had already been proposed by Kuznets (1929, p. 277):

> As an industry starts from small beginnings and develops rapidly to substantial output, it is enabled to do so mainly by progress in the technical conditions of production. But the effects of technical progress show an unmistakable tendency to slacken due either to retardation in technical progress or the pressure of exhaustion of resources or both. Added to that is the check exercised by groups of productive activity whose industrial arts do not improve as rapidly and as significantly as in the industry in question.

That has a very contemporary sound to it. A sympathetic exegesis could read into this paragraph not only a recognition of technological evolution as a driving force of the life cycle, but also a statement on the importance of inter-industry linkages for the cycle. Yet one must note that at least as many theoretical questions are raised as are answered by such an interpretation. Where does the "progress in the technical conditions of production" originate, and how is it transmitted among the members of an industry? What causes the "retardation in technical progress?" Is not "the pressure of exhaustion of resources" an *effect* of a lack of technological advances, rather than an independent cause of industrial retardation? If one sees the interindustry transmission of retardation in terms other than those of simple vertical linkages or technological complementarities, what are the forces at work?

These are, of course, the same questions with which economists interested in the life cycle are concerned even today. I list them here not to engage in a cheap criticism of pioneering efforts, but rather to suggest that the search for a more general theory of the life cycle, one that incorporates explicit answers to these questions, is still going on. I would also propose that a less sympathetic interpreter could say that, from a theoretical viewpoint, the early researchers did no more than to add a few orbits to Boulding's "sunless astronomy of commodities."

This was not, however, the contemporary view in the profession nor that of later interpreters. Consider the following quotes as evidence: "This study (of Burns) has become a classic, and not much can be added to its findings. . . . A declining rate of growth after a certain point is the general law in individual industries, just as it is in organic growth" (Woytinski, 1953); and "exhaustion of the possibilities of innovation contributes to the maturing and eventual decline of industry. Kuznets and Burns used this line of reasoning in their explanation of retardation in the rates of growth of individual industries, and I think they were right to do so" (Fabricant, 1960).

The widespread acceptance of the early empirical results and of their plausible theoretical explanation, together with the fact that the economics profession had turned its attention to other problems, meant that interest in the life cycle *per se* was on the wane for several decades. Therefore, a study by Gold (1964), in which he updated the industry time series of Burns, made no great impression on the received doctrine. Gold concluded on the basis of his empirical results that, far from being preprogrammed by invariant patterns of technological evolution or by other factors, industry growth trajectories could take a number of shapes. These ranged from long-term, steady-state expansion at one end of the spectrum to recovery into rapid growth after prolonged, and seemingly terminal, decline at the other. Gold *et al.*, (1968) further extended this line of research by demonstrating even more heterogeneous patterns among the firms of given industries.

There may be a further reason for the inconclusiveness of all the early statistical evidence. As van Duijn (1983) pointed out, one problem was that these studies "lump together many kinds of industries, from agricultural staples to industrial inputs, from capital to consumer goods, and then look for one growth pattern" (p. 28). His own findings "suggest the existence of an S-shaped growth pattern *up to* the maturity of the industry, *with various possible patterns thereafter*" [emphasis added] (p. 29). Whatever the other implications of van Duijn's observation for the analysis of industry-level cycles, it certainly leads to a methodological conclusion with which no one could disagree – we need to collect and analyze as many data as possible on as many industries as possible.

Equally agreeable is the suggestion of Kleinknecht (1984, p. 147) that "the level of aggregation is very important in the study of changes in the character of innovation." One is tempted to add that the higher the aggregation, the greater the chances that all sorts of noise will drown out the phenomena one is after. It is not surprising that the same point, together with an excursion into the problems of trend-fitting, was made by Edwin Frickey (1934), the scholar who was in all likelihood the statistically most sophisticated contributor to the early life-cycle studies.

I would add one more lesson from the experience of the first wave of research: to the extent that our investigations have to rely on official statistics, bound to unchanging definitions of an industry, we may miss some of the most important longer-term phenomena. In the first place, such statistics do not reveal the full extent of such structural changes in the various sectors of the economy that occur as industries "spin off" or take on whole ranges of activity. The impact of "high technology" is mentioned most frequently in this connection. There are, however, more homely examples: consider the extent to which the hauling of crucial hard-copy information, replacement parts, etc., by company couriers has been taken over, in the United States, by several competing overnight express-delivery services: an industry that did not even exist a decade ago has turned into a multimillion-dollar business. Conversely, the integration of computers into their products has propelled some machine-tool firms to the technological frontier, even though statistically they will remain wedded to an ostensibly "mature" industry.

Second, it may be that what is recorded as the behavior of an industry is in fact driven by the unique effects of technological innovation on a single firm's position in the market. Examples here are the Ford Model T and the Douglas DC-3, both of which accounted for well over 70% of their respective industries' success during the period of most rapid growth. Some economists would argue that this latter possibility is adequately explored by the discipline of industrial organization (I.O.); I would counter, however, that the traditional paradigm of I.O. sees the arrow of causation pointing the other way – from a "given" industry structure to the propensity for innovation.

Finally, an exclusive focus on the data of industry *per se* may cause us to overlook the role of changes in institutional arrangements as crucial elements of the life cycle (Rosegger, 1976). I am thinking here not only of political, legal, and regulatory influences, but also of such factors as the voluntary standardization of products and processes, the organization of capital markets, as well as the frequently observed, more or less formal, interfirm cooperative arrangements in research and development, production, and marketing.

The conclusion from these observations is obvious, and it is largely reflected in the second wave of life-cycle research, which started up about two decades ago: explanations of longer-term fluctuations in the fortunes of industries are not quite as straightforward as the pioneers thought. Guided by Einstein's famous dictum that only theory can decide what one is able to observe, economists and other scholars have ranged widely in formulating and testing more complex hypotheses about the nature of the life cycle.

2.3. Life-cycle Research

What should rightly follow now, in a state-of-the-art survey, is a comprehensive look at the second-wave literature. In view of the flood of recent work this would be a near-impossible task in the time allotted to me. Instead, I want to deal with just a few matters that have aroused some interest in connection with this work. My choice is very idiosyncratic, so it would be better to say that these matters have aroused my interest.

2.3.1. What can we learn from other disciplines?

Turning away from Newtonian mechanics is one thing, finding new paradigms is another. When we used words like *evolution, diffusion, infancy, maturity*, and so on, we are obviously leaning on concepts that have been developed and used in the biological sciences. Today, we could no longer get by without these concepts. I must point out, however, that the longing for a "biological" rather than a "mechanical" economics is nothing new. In fact, even the father of the neoclassical school, Alfred Marshall (1890), expressed this longing in his most famous work (see especially p. 765 and Appendix C).

Recent progress in theoretical biology has made it possible to do more than borrow a language. Increasingly, economic life-cycle research is able to draw on models and hypotheses developed in evolutionary and ecological studies. This has brought some very useful results, but there are hazards as well in such cross-fertilization among sciences. As one of the pioneers in the formal modeling of social processes, Anatol Rapoport (1956), pointed out a long time ago:

> The use of the word "diffusion" in the social science indicates the awareness of some similarity between the spread of, say, a technological artifact and the spread of a solute through a solvent. Of course it may be argued that the use of a metaphor does not establish the reality of the connection between situations compared and, in fact, may be seriously misleading.

Simplistic analogies are always dangerous, or they turn out to be so self-evident as to be irrelevant. What insights, for example, are we meant to derive from propositions like, "no living organism grows indefinitely" or "species can survive only by differentiating themselves from other species?" This kind of bio-talk has crept into some of the more superficial business writing on the life cycle, where it is usually meant to add a "scientific" tone to otherwise trivial exhortations.

Digging a little deeper, however, I believe that we may find fascinating and largely unexploited potentials for the modeling of industrial life-cycle phenomena in the directions of work suggested by R.M. May (1976) [see also May and Seger (1986) and R.V. Jensen (1987)]. Their explorations of the complex dynamics of time trajectories, which can result from apparently random, small-scale fluctuations, are based on very simple equations, but they suggest a range of insights.

These may well be obvious to persons of mathematical sophistication, but if so, I have yet to find them reflected in the life-cycle literature.

If nothing else, the work sheds new light on the possible variations in behavior produced by that mainstay of life-cycle research, the logistic function. In particular, there is the possibility that systematic changes in the coefficient indicating the "potency of spread" (Rapoport, 1956) may allow us to identify separate *strategic regimes* "inside" the seemingly smooth industrial growth curves. Our own recent study on the impact of the entry and exit of makes on the aggregate growth of the American and European automobile industries (Rosegger and Baird, 1987) seems to provide at least some hints in this direction. Similar patterns are suggested in the study of Gort and Klepper (1982).

I want to mention just two other disciplines from which life-cycle research in economics may derive considerable fertilization: psychology and organizational behavior. If we believe that the evolution of firms, and therefore of industries, is driven by the strategic decisions of individuals, we ought to accord an understanding of these decisions high priority. Herbert Simon and the behavioralist school have, of course, provided a bridge between the sciences of man and the science of the firm. Here I am thinking, however, more specifically of the role of *creativity* in invention and innovation. Economists have tended to shy away from the subject, except in some of its most superficial, quantifiable manifestations, such as patent activity. To the best of my knowledge, Haustein and Maier (1985) are the only ones to have probed more deeply into the "economic dimensions of creativity" (pp. 142–148).

In my own review of the factors influencing the supply of inventions (Rosegger, 1986, pp. 112–117), in which I attempted to model inventing as purposeful search, I derived many useful clues from the work of Austin (1978). He proposes the hypothesis that the directions and outcomes of creative activity can be characterized in terms of four different types of "chance," each of which requires different intellectual (and by implication, organizational) approaches to the search process. A suggestion along somewhat similar lines is provided by Burton Klein (1979), when he defines dynamic efficiency as "maximizing the probability of recognizing good luck, while minimizing the consequences of bad luck" (p. 81), a definition that raises the question of how organizations might best go about recognizing "good luck" and capitalizing on it. Answers not only are important from the managerial viewpoint, but also may have a bearing on our ability to deal with another big issue that has occupied recent life-cycle research – the predictability of technological development.

2.3.2. Is technological progress preprogrammed?

Clearly, an ability to say something about the contours of likely technological developments is a necessary, though not a sufficient, condition for assessing the future directions of change in individual industries. The issue is controversial, however, in both theoretical and empirical terms. Certainly the optimism of the

1960s about the potential of technological forecasting, an optimism that was never really shared by the economics profession, has given way to a more cautious attitude.

The prevailing view among economists, including the followers of Schumpeter, is probably best represented by Scherer (1984):

> The distinction in an economic sense between invention and the development process underlying innovation is best summarized in the difference between two words "predictability" and "describability." Basic invention is truly unpredictable ... on the other hand (a scientist) knows in appraising the detailed problems of development that an answer will be obtained and can only not describe what the answer will be [p. 6].

I take this to imply that we can say something about the directions of technological trajectories once they have started, but that their starting points are randomly determined by the randomness of inventions.

An opposing view is taken by Ayres (1987, p. 7), who argues that "scientists and engineers tend to be less pessimistic about the possibility of forecasting what will be produced by R&D." His historical review aims to make several points in support of the argument. For the purpose of analyzing the life cycle in individual industries, the most important is Ayres's discussion of the role of scientific or technological constraints: to the extent that these inhibit further advances in a technology and thus condemn the industries tied up with the technology to maturity or even decline, it is possible to predict "what it would take" to break through the constraint. Ayres's examples provide copious evidence for the proposition, but they are not likely to persuade the agnostic economist, who views history as a series of events that happened with a probability of 1.0 and therefore are no longer of great interest, except on the whole and on average.

Our studies of long-term technological evolution (see, for example, Gold *et al.*, 1984) have persuaded me that it may be quite possible to construct an industry-level model of process innovation based on the proposition that incentives come mainly from the pressure for the successive elimination of *bottlenecks* created by preceding innovations. Needless to say, however, such a hypothesis could tell nothing about the timing of these events or, indeed, about the speed of diffusion after they have occurred (Gold *et al.*, 1970).

More generally acceptable may be the view that we are better at using our current knowledge as a base for forecasting what is *not* likely to happen in the development of products, processes, and industries. For example: 50 years ago, storage batteries had capacities of approximately 30 watt-hours per kg of battery weight; technological progress has pushed this limit to today's 70 watt-hours. Even this remarkable improvement, however, is regarded as insufficient to permit the development of an economically attractive, electricity-powered automobile, government subsidies for R&D not withstanding. Will there be the kind of "breakthrough" that is central to Ayres's argument? The economist is likely to reply yes, if the (relative) price is right.

In his delightful study of the seemingly irrational persistence, across time and industries, of the QWERTY keyboard (a true "bottleneck" if there ever was one) David (1986) proposes another hypothesis that students of the life cycle must take quite seriously: every now and then in economic history,

> "one damn thing follows another" . . . it is sometimes not possible to uncover the logic (or illogic) of the world around us except by understanding how it got that way. A *path-dependent* sequence of economic changes is one in which important influences upon the eventual outcome can be exerted by temporally remote events, including happenings dominated by chance elements rather than systematic forces. Stochastic processes like that do not converge automatically to a fixed-point distribution of incomes. . . . In such circumstances "historical accidents" can neither be ignored, nor neatly quarantined for the purposes of economic analysis; the dynamic process itself takes on an *essentially historical* character [p. 30].

No doubt there are numerous instances where the long-term development trajectories of industries were determined by initial technological events that proved to be what David calls "quasi-irreversible" (p. 44). He draws on the work of Arthur *et al.*, (1985) for an explanation of these trajectories in terms of the formal theory of stochastic processes.

Unfortunately, the case study as a tool of research is held in low esteem by the economics profession. Yet it is only through such detailed investigations of sequences of *particular* technological events that we might learn more about their role in industrial evolution. Modern economists have yielded the field to the historians and, I would argue, have lost out in the resulting division of labor.

In discussing the problem of predictability, I have deliberately avoided the issue that appears to have been discussed most extensively: whether technology-push or demand-pull is the dominant force in the "programming" of markets for new products and processes. I believe that no amount of data-gathering and analysis will settle the argument, because it depends on the assumption that there exists an *a priori* demand function for a commodity about whose specific characteristics and price virtually nothing is known. Certainly the standard demand function of equilibrium theory cannot bear the burden of these assumptions.

The push and the pull occasioned by an innovation depend on how well its characteristics satisfy some technical or other requirements of potential adopters, including requirements about which they had no information until the innovation appeared. These characteristics change rapidly in the early stages of the life cycle and probably more slowly later on. At the same time the key variable in the demand function, price, also changes in the process of diffusion, as do relevant extraneous variables, such as real incomes. Therefore, shifts of a demand function are indistinguishable from movements along a given function. The effect on diffusion measures was demonstrated in a formal, empirical way by Bonus (1973). Some innovations find a long-term niche in the market, and many others do not; to debate *ex post* about the reasons for their success or failure in terms of clearly separable "push" and "pull" forces seems to me a little like arguing which blade of a pair of scissors does the cutting.

2.3.3. Does the capitalist/financier matter?

The diffusion of major technological innovations involves the coordination of a number of firms in different industries – arrayed vertically along the production path, horizontally, as in marketing and distribution, or in more complex networks. Economic studies of the life cycle generally recognize this fact and trace these linkages through the behavior of the relevant markets. Yet there is one market whose coordinative function is hardly ever mentioned, the capital market. We speak of "investment in innovation," but that term seems to be used as a metaphor suggesting a managerial or organizational commitment, rather than the commitment of one's own or someone else's savings.

The omission of financial factors from life-cycle accounts is all the more surprising as Schumpeter himself drew a clear distinction between the roles of the innovator and of the capitalist. Of course, the innovator is the more interesting player in the game, and the financing even of large ventures is frequently internalized in modern corporations. But, as any textbook will tell us, this latter fact in no way reduces the capital market's impact on decisions. In a very real sense, that market's attitudes can be said to be a key argument in the *demand* for innovations.

Perhaps our reluctance to deal with this problem comes from a perception that somehow investment in the radically new is substantively, and not just qualitatively, different from the more routine commitment of financial resources to other types of projects. When Keynes writes of "animal spirits" and when Schumpeter resorts to terms like *Krisenpsychologie* (crisis psychology) in explaining investors' attitudes, there is at least a hint of such a difference. G.L.S. Shackle (1972, p. 27) makes the same point in dramatic brevity: "Time is a denial of the omnipotence of reason." However, von Mises (1966, pp. 112–113) put the matter into what I consider a proper perspective: "Every investment is a form of speculation. There is in the course of human events no stability and consequently no safety. . . . Gambling, engineering and speculating are three different modes of dealing with the future."

We deal with the future by expressing time preferences. This is not the place to investigate the reasonableness of economic theory's standard assumptions about the *structure* of these preferences and their relevance for investment decisions. These assumptions have been questioned a number of times, such as by Gold and Boylan (1975), Gold (1977), and most recently by Ayres and Mori (forthcoming). What matters for my purpose is the question whether longer-term systematic *changes* in time preference do occur and how they might contribute to life-cycle phenomena.

A quick answer is easy: low-time preference strengthens support for saving and hence the accumulation of capital goods incorporating new technology, and *vice versa* for high-time preference. This is a proposition frequently drawn on in recent comparisons between the performance of the American and the Japanese industrial sectors. For a more thorough answer one would have to dig deeper: an era of high-time preference not only produces economic stagnation in the direct sense of my quick answer, but also affects the political, social, and cultural climate, and therefore what Gordon (1969) in his essay on "The feedback between technology and values" has called the "social forces driving research."

Economists are not very comfortable with notions like "values" and "social forces." I would suggest, however, that if one ignores these broader implications of time preference, one may fall into the trap of reifying technology and therefore seeing it as *the* "independent" driving force of economic dynamics. *Time preference, abstinence, patience, waiting, speculation, long horizons*, etc., are words attempting to describe how we fold the future into decisions of the present. In a capitalist economy, they imply that progress rests to a large extent on the willingness of some people to give up certain current gratification (consumption), by setting aside a portion of their incomes for uncertain gratification in the future.

In that sense, saving decisions matter as much as the technical artifacts they help to produce. For purposes of explaining industrial life cycles, one could of course simply say the "macroeconomic conditions" influence changes in demand, investment, and output, but I propose that what really matters lies a bit deeper. How deep, I do not know, nor have I any good ideas on how one would go about probing these matters in an empirical way.

2.3.4. Stagnation vs. "post-industrialism"

The questions I have just raised lead quite naturally to the last issue I want to discuss. Borrowing from yet another discipline, medical science, Kindleberger (1979) asked the question whether there is such a phenomenon as an "economic climacteric," a general, economy-wide transition into senescence. His historical examples are persuasive, and his analysis suggests that there are processes at work, among them a slowdown in the rate of technological advance, which may cause national economies eventually to slide from industrial leadership to relative insignifiance. Olson (1982) provides a much more comprehensive set of explanations of the same phenomenon.

I believe that it is useful to distinguish these views from the more traditional hypotheses about technology-driven long waves. For one thing, long-wave research attempts to find *regularities*, whereas the "sclerosis" theories concern themselves only with the *recurrence* of certain symptoms. And for another, long waves ought by their nature to be global, though perhaps phased nationally, while the focus of scholars like Kindleberger, Olson, and others clearly is on technological and institutional forces at work *inside* individual countries. What both types of hypothesis have in common is that they are testable, at least in principle.

Confusion about the nature and intent of this work, with some misinterpreted data from life-cycle studies thrown in to muddy the waters even further, has given rise to what I consider a most unfortunate set of untestable propositions about the nature of "post-industrialism." Individual contributors to this literature usually are not very careful in defining the concept. They apparently mean to suggest that, at least in the developed parts of the world, a radical change in the bases for economic activity has to occur, if it is not already occurring.

Even a careful reader has difficulty synthesizing some common elements from the many facets proposed for post-industrial society, but I take it that they include at least the following: the accumulation of knowledge in place of the accumulation of physical capital; innovations that improve the "quality of life" rather than innovations that further increase real output; the growth of services, especially in the communications and leisure-time industries, in place of growth in manufacturing and related activities; the development of technological and administrative fixes for social problems; and an attenuation of the work ethic and other bourgeois values.

I expect to be criticized for this simplification of a very complex set of ideas, but it allows me to make my point: because life-cycle researchers' findings could have tempted the post-industrialists into the fallacy of composition, the former may have been unwitting contributors to some fundamental errors. First and foremost, a quasi-moratorium on the accumulation of capital *in the aggregate* would imply not only a halt to growth but a freezing of the existing composition of physical output, i.e., the structure of industries.

Furthermore, if an improvement in the quality of life is to include a more wholesome environment, a goal with which it would be difficult to disagree, the pursuit of that objective involves direct expenditures as well as shifts from more to less environmentally harmful products and processes – in other words, investment.

Most important, however, knowledge accumulation for the economy as a whole is not an alternative to capital accumulation. As Machlup (1962) demonstrated, the two have always been complementary, and it is difficult to see how this symbiotic relationship could be broken.

These points suggest that low rates of time preference are an essential feature of a world that wants to improve the material lot as well as the social and cultural well-being of its people. The relevant arguments are made persuasively, and from a number of different angles, in several collections of studies, such as Olson and Landsberg (1973) and Landau and Rosenberg (1986). The issue is not *whether* societies will continue to experience technological advance and concomitant changes in industrial structure, but what kind of trade-offs they are willing to make in affecting the directions of change.

To gain greater public understanding of this point would seem to be an important social contribution that life-cycle researchers could make, especially at times when an unreasoning hostility toward technology *per se* seems to be fashionable in certain circles. Only cultural historians may be able to explain when and why these waves of public sentiment threaten to swamp more balanced evaluations of the opportunities and risks of innovation. As Brooks (1986) demonstrated, the United States seems to have gone through at least three cycles of changing public attitudes since World War II.

2.4. Conclusion

I want to conclude by suggesting that life-cycle research can make a social contribution, because I believe that a survey of today's state of the art must include a new element: life-cycle research is changing from a rather esoteric field of inquiry, albeit one from which many of us derive great intellectual pleasure, into one that is perceived to have very utilitarian implications.

Helped along, no doubt, by some exuberant writings in the management literature, economists now often are called upon to use their presumed understanding of the life cycle to give all sorts of practical advice – to corporate decision makers, to trade associations, to the designers of industrial policies, and to international organizations. I am not certain that our "art" is quite up to the task of meeting some of these rather high expectations. Yet we seem to be bearing the burden of responsibility with good cheer and with the comforting feeling that no one else is doing much better. Nevertheless, at this stage in our understanding of the life cycle's many facets, it may behoove us to remember Morris Adelman's claim that every now and then, the hardest sentence for an economist to pronounce is: "I don't know."

Acknowledgment

I want to thank my friends and colleagues, Asim Erdilek and William S. Peirce, for their helpful and encouraging criticism of an earlier version of this paper. Having taken some, but by no means all, of their counsel, I cannot blame them for any errors of fact or interpretation.

References

Arthur, W.B., Ermoliev, Y.M., and Kaniovski, Y.M., 1985, Strong Laws for a Class of Path-Dependent Urn Processes, in *Proceedings of the International Conference on Stochastic Optimization, Kiev, 1984*, Springer-Verlag, Berlin, Germany, F.R.

Austin, J.H., 1978, *Chase, Chance, and Creativity: The Lucky Art of Novelty*, Columbia University Press, New York, NY, USA.

Ayres, R.U., 1987, *The Industry-Technology Life Cycle: An Integrating Meta-Model?*, RR-87-3, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Ayres, R.U. and Mori, S., forthcoming, Time Preference and the Life Cycle: The Logic of Long-Term Risk vs. Short-Term Low Risk Investment, *European Journal of Operations Research*.

Binswanger, H.P. and Ruttan, V.W., eds., 1978, *Induced Innovation: Technology, Institutions, and Development*, Johns Hopkins University Press, Baltimore, MD, USA.

Bonus, H., 1973, Quasi-Engel Curves, Diffusion, and the Ownership of Major Consumer Durables, *Journal of Political Economy* **81**:655–677.

Boulding, K.E., 1957, Ecology and Economy, in G.W. Sanford, ed., *Segments of the Economy: A Symposium*, Howard Allen, Cleveland, OH, USA.

Brooks, H., 1986, National Science Policy and Technological Innovation, in R. Landau and N. Rosenberg, *op. cit.*

Burns, A.F., 1934, *Production Trends in the United States since 1870*, National Bureau of Economic Research, New York, NY, USA.

David, P.A., 1986, Understanding the Economics of QWERTY: The Necessity of History, in W.N. Parker, ed., *Economic History and the Modern Economist*, Basil Blackwell, London, UK.

van Duijn, J.J., 1983, *The Long Wave in Economic Life*, George Allen & Unwin, London, UK.

Fabricant, S., 1960, Study of the Size and Efficiency of the American Economy, in E.A.G. Robinson, ed., *Economic Consequences of the Size of Nations*, St. Martin's Press, New York, NY, USA.

Frickey, E., 1934, The Problem of Secular Trend, *Review of Economics and Statistics*, October.

Friedlaender, A.F., 1986, Macroeconomics and Microeconomics on Innovation: The Role of the Technological Environment, in R. Landau and N. Rosenberg, *op. cit.*

Gold, B., 1964, Industry Growth Patterns: Theory and Empirical Results, *Journal of Industrial Economics* November.

Gold, B., 1977, On the Shaky Foundations of Capital Budgeting, *California Management Review* March.

Gold, B. and Boylan, M.G., 1975, Capital Budgeting, Industrial Capacity, and Imports, *Quarterly Review of Economics and Business* Fall.

Gold, B., Huettner, D., Mitchell, R., and Skeddle, R., 1968, Long-Term Growth Patterns of Industries, Firms, and Products, *Proceedings, Business and Economic Statistics Section*, American Statistical Association, Washington, DC, USA.

Gold, B., Peirce, W.S., and Rosegger, G., 1970, Diffusion of Major Innovations in US Iron and Steel Manufacturing, *Journal of Industrial Economics* July.

Gold, B., Peirce, W.S., Rosegger, G., and Perlman, M., 1984, *Technological Progress and Industrial Leadership*, Lexington Books – D.C. Heath, Lexington, MA, USA.

Gordon, T.J., 1969, The Feedback between Technology and Values, in K. Baier and N. Rescher, *Values and the Future: The Impact of Technological Change on American Values*, Free Press, New York, NY, USA.

Gort, M. and Klepper, S., 1982, Time Patterns in the Diffusion of Product Innovations, *Economic Journal* September.

Haustein, H.-D. and Maier, H., 1985, *Innovation and Efficiency*, Akademie-Verlag, Berlin, Germany, F.R.

Jensen, R.V., 1987, Classical Chaos, *American Scientist* March–April.

Kindleberger, C., 1979, An American Economic Climacteric?, *Challenge* January–February.

Klein, B., 1979, The Slowdown in Productivity Advances: A Dynamic Explanation, in C.T. Hill and J.M. Utterback, eds., *Technological Innovation for a Dynamic Economy*, Pergamon Press, New York, NY, USA.

Kleinknecht, A., 1984, *Innovation Patterns in Crisis and Prosperity: Schumpeter's Long Cycle Reconsidered*, Sneldruk Hengelosestr., Enschede, The Netherlands.

Kuznets, S., 1929, Retardation of Industrial Growth, *Journal of Economic and Business History*, August, reprinted in S. Kuznets, 1953, *Economic Change*, Norton, New York, NY, USA.

Landau, R. and Rosenberg, N., eds., 1986, *The Positive Sum Strategy*, National Academy Press, Washington, DC, USA.

Link, A.N., 1987, *Technological Change and Productivity Growth*, Harwood Academic Publishers, New York, NY, USA.

Machlup, F., 1962, Can There be Too Much Research? in E.S. Phelps, ed., *The Goal of Economic Growth*, Norton & Co., New York, NY, USA.

Mansfield, E., 1983, Long Waves and Technological Innovation, *American Economic Review, Papers and Proceedings*, May.

Marshall, A., 1890, *Principles of Economics*, 8th ed., 1920, Macmillan, New York, NY, USA.

May, R.M. 1976, Simple Mathematical Models with Very Complicated Dynamics, *Nature* June 10.

May, R.M. and Seger, J., 1986, Ideas in Ecology, *American Scientist* May–June.

von Mises, L., 1966, *Human Action*, Henry Regnery, Chicago, IL, USA.

Nelson, R.R. and Winter, S.G., 1982, *An Evolutionary Theory of Economic Change*, Harvard University Press, Cambridge, MA, USA.

Olson, M., 1982, *The Rise and Decline of Nations: Economic Growth Stagflation, and Social Rigidities*, Yale University Press, New Haven, CT, USA.

Olson, M. and Landsberg, H., eds., 1973, *The No-Growth Society*, Norton, New York, NY, USA.

Rapoport, A., 1956, The Diffusion Problem in Mass Behavior, *General Systems Yearbook*.

Rosegger, G., 1976, Diffusion Research in the Industrial Setting: Some Conceptual Clarifications, *Technological Forecasting and Social Change* 9:401–410.

Rosegger, G., 1986, *The Economics of Production and Innovation: An Industrial Perspective*, 2nd ed., Pergamon, Oxford, UK.

Rosegger, G. and Baird, R.N., 1987, Entry and Exit of Makes in the Automobile Industry, 1895–1960: An International Comparison, *Omega: International Journal of Management Science* April.

Rosenberg, N. and Frischtak, C.R., 1983, Long Waves and Economic Growth: A Critical Appraisal, *American Economic Review, Papers and Proceedings*, May.

Salter, W.E.G., 1960, *Productivity and Technical Change*, Cambridge University Press, London, UK.

Scherer, F.M., 1984, *Innovation and Growth: Schumpeterian Perspectives*, MIT Press, Cambridge, MA, USA.

Shackle, G.L.S., 1972, *Epistemics and Economics: A Critique of Economic Doctrines*, Cambridge University Press, London, UK.

Stoneman, P., 1983, *The Economic Analysis of Technological Change*, Oxford University Press, Oxford, UK.

Vernon, R., 1966, International Investment and International Trade in the Product Life Cycle, *Quarterly Journal of Economics* May.

Woytinski, W.S., 1953, *Employment and Wages in the United States*, Twentieth Century Fund, New York, NY, USA.

CHAPTER 3

Role of the Technological Life Cycle in Technology and Global Industry

Harvey Brooks

What I am going to do in this paper is to comment on certain issues that arose in a conference I helped to organize in February 1986 for the US National Academy of Engineering (NAE) under the title, "Technology and Global Industry: Companies and Nations in the World Economy," (Guile and Brooks, 1987). In broadcast terms, using language that Jim Miller explains in his paper on general systems theory, we have a world in which two kinds of organizations coexist – economic and political – which are to an increasing degree overlapping in their organizational boundaries. Thus the underlying theme of the NAE symposium was the "economic, social, and industrial organization issues brought forward by confluence of new technologies, a high level of international interdependence, and diverse concerns of nations and companies trading in world markets."

In my observations, however, I will confine my attention to the role of the technology life-cycle concept (TLC) in this discussion. Out of eight major papers presented in the symposium and published by the National Academy Press in June, three – by James Utterback, David Teece, and Henry Ergas – made the most extensive and basic use of the TLC concept. What I would like to do is to comment briefly on some of the issues and questions that arose in relation to the use of this concept in each paper.

First is the question of the nature of the TLC. Most of the illustrations of the TLC were graphs without ordinates which, Professor Miller assures us, would never have been accepted for publication by the editors of the *Journal of Behavioral Science*. What is it that the ordinate in these graphs is supposed to measure? Is it gross sales or market penetration of a new product, or a cluster of closely related products? Or is it some combination of performance parameters of a class of products?

There is also a question of what the unit of analysis is. Utterback (1987) gives the most definite answer to this. In his paper he is primarily concerned with matching organizational capacities and structures within the firm to the evolution of the TLC. The unit of analysis is what Utterback calls the "productive unit," which he defines as a product line and its associated production process. This unit is narrower than a whole firm, or even a division of a firm, but broader than a single innovative product. Thus it is neither a firm-specific nor a technology-specific technology. For example, it would ordinarily include suppliers of components and of capital equipment, especially if the production process requires close collaboration and iterative interaction between the manufacturer of the product line and the designer and supplier of the production machinery. From a technological perspective, therefore, the "productive unit" involves a whole cluster of related innovations both downstream and upstream from the main production process. To an increasing degree the productive unit cannot even be confined within national boundaries, but is integrated multinationally.

Utterback's paper is primarily concerned with the impact of the phase in the TLC on the type of organization or social system best configured to manage that particular phase. He divides the TLC into three main phases that he labels the "fluid pattern," the "transitional pattern" in which the "dominant design" of the product line emerges, and the "specific pattern" in which the product line becomes highly standardized and the emphasis of both innovative effort and management shifts from product design to process design.

In the fluid pattern there are numerous firms with rapid market entry and exit, and competition mainly with respect to the product's technical characteristics and performance. In this phase a highly adaptive, flexible organization is required. Production tends to be labor-intensive with a high requirement for engineers and skilled craftsmen while production equipment tends to be general purpose rather than specialized. The "skunk works" is the optimal type of product development organization for this phase.

As the dominant design begins to emerge, a large number of firms are driven out of the field by competition. Innovative effort becomes more focused on improving a few fundamental performance parameters of the product and on process innovation with emphasis on price and reliability of the product. Production equipment becomes more specialized and the organization more disciplined and hierarchical, with tighter integration both upstream with suppliers and downstream with distributors and service organizations. What evolves is a carefully articulated but increasingly rigid system in which it becomes increasingly difficult to accept other than minor, incremental innovation. Production is no longer by skilled craftsmen but by automated special-purpose equipment operated by semiskilled labor, with increasing separation between brain work and manual work. Further innovation becomes more and more disruptive to the system in that it has repercussions throughout the whole system and not just on a local element of the production process. This describes what Utterback calls the specific pattern in which most economies of scale have been realized and learning curve effects exploited.

One problem raised by Utterback's picture is how many product lines within a corporation in different phases of the TLC just described can be accommodated within a single administrative structure belonging to a common organizational culture. Closely related to this problem is the vulnerability of "mature" industries to innovative products or processes introduced by outsiders whether in the same or another country. The openness of the international economy has increased the probability that this will happen unless the corporation can be organized to respond rapidly.

David Teece (1987) approaches the TLC from a very different standpoint; his paper deals with what he calls "appropriability regimes" – the capacity of a firm or a nation to capture the economic benefits resulting from its ability to innovate technologically, in other words the economic payoffs that a firm or country can expect to derive from its R&D system. Teece stresses the observation that economic benefits do not necessarily accrue to the innovator and cites many examples of major industrial innovations where the innovating firm was subsequently driven out of the market by competition from skillful imitators. Appropriability depends in the first instance on ability to limit the exploitation by others of the idea on which the innovation is based – trade secrets or patent position, for example, or just the difficulty of imitation. The more science-based the technology, the more problematic the appropriability regime, and the more important is access to what Teece calls "complementary assets" – resources necessary for getting the innovation to market earlier and more effectively than potential imitators. The most important complementary assets are excellent manufacturing capabilities and effective distribution-maintenance-service channels. Complementary assets are most valuable when they are specialized to the particular innovation so that they cannot be readily duplicated or simply acquired on the open market, e.g., general-purpose manufacturing equipment available from many different suppliers or distribution channels that handle many related competitive products. Such general-purpose assets can be relatively easily contracted for by a competitor. But, as indicated by Utterback, both manufacturing and distribution tend to become more specialized as one moves from the fluid pattern to the transitional pattern of the TLC, or, as Teece calls it, the "preparadigmatic" to the "paradigmatic" stage. In other words the required complementary assets become more and more specialized to the dominant paradigm of the product line as one progresses through the technology life cycle.

Both Teece (1987) and Ergas (1987) argue that large returns accrue to incremental improvements on a dominant paradigm, with relatively low *average* returns to the totality of firms in the preparadigmatic stage because of the rapid entry and exit of firms and the extreme volatility of relative market shares among the remaining competitors. While Utterback argues that a large advantage accrues to firms that succeed in the preparadigmatic stage and discover the dominant design first, Ergas suggests that, although this might be true for an individual firm, it is probably not true for an economy that follows a general strategy of emphasizing this stage, e.g., relying exclusively on heavy investment in R&D and searching for radical innovations. In other words he argues that such a strategy does not contribute to raising living standards of the work force in the absence of an effective follow-on strategy. The innovator also runs the risk of moving either too slowly or too quickly in acquiring exclusive access to specialized complementary assets. If he moves too quickly – overestimating the stability of the emerging dominant design, for example – his investment may be inappropriate and later have to be scrapped. If he moves too slowly, he may miss the window of opportunity in the market, allowing an imitator to take advantage of complementary assets he controls to take away the market before the innovator can gear up to penetrate it. Teece gives numerous examples of cases where pioneers lost out to imitators because of the latter's superior access to appropriate complementary assets, either because it had them in place in connection with other related products or because it was more confident of the arrival of the dominant design and put the necessary complementary investment in place more rapidly. In this regard Teece particularly emphasizes manufacturing skills, citing Japanese capture of portable radio, TV, auto, and standard RAM markets as examples. But one also has to mention the skillful Japanese exploitation of the efficient US continent-wide distribution systems such as Sears Roebuck as an example of the exploitation of existing complementary assets through producing a product that US manufacturers had neglected. Sears and other mass merchandizers played a key role in the penetration of the US market first by small black-and-white and later color TV sets made in Japan (Peck and Wilson, 1982). The US auto dealer system played a similar role in Japanese penetration of the US small car market. Conversely, the Japanese shrewdly denied US innovators in semiconductor devices and computers access to complementary assets, such as distribution outlets, in the Japanese economy until such time as domestic Japanese producers had learned the technology well enough and realized sufficient scale economies in the domestic market to be able to exploit their manufacturing superiority in the US export market.

Whereas Utterback approaches the technology life cycle from the standpoint of the productive unit, and Teece approaches it from the standpoint of the firm, Ergas (1987) approaches it from the standpoint of the whole national economy. In his chapter he distinguishes three national strategies. One stresses operation in what he labels the "emergent phase" of heroic technologies with heavy emphasis on national sovereignty (prestige or military), centralized decision making by government, and pressing the state of the art in technical performance without too much regard for economics. In the past this has generally represented the military-space-nuclear complex of technologies, with biotechnology, deep-sea technology, and exotic materials being added to the list more recently. According to Ergas the USA, France, and the UK are examples of countries following this technology strategy. He suggests that France has been the most successful and the UK the least successful in this strategy as judged strictly from the standpoint of managerial performance – success in achieving proclaimed technical objectives. On the other hand, while its managerial performance has been middling, the USA has probably been the most successful in deriving economic benefit from this heroic strategy primarily because of the remarkable strength and diversity of its R&D establishment and its ability, because of shear size, to pursue many alternatives in parallel (Ergas, 1987).

The opposite case is represented by the Federal Republic of Germany, Sweden, and Switzerland and is described as a diffusion strategy in which pursuit of the TLC into the mature stage or specific pattern is emphasized. Here the idea is to utilize new technology from whatever source to achieve rapid incremental improvement in existing industries, emphasizing the high value-added end of the possible spectrum of products. For example, in the machine-tool industry the average value-added represented by the West German product mix is about three times the average for all OECD countries (Ergas, 1987). Ergas suggests that this strategy of incremental improvement and high value-added of relatively conventional products may generally be the most successful from the standpoint of improving the living standards of the work force. Its weakness is its vulnerability to attack through the introduction of a brand new paradigm. A good example is what happened to the Swiss mechanical watch industry. It also nearly happened with the West German machine-tool industry with the introduction of electronics, but in both these cases the diffusion strategy, implemented through an unmatched vocational training system, eventually enabled the industry to restore its position – suggesting a high degree of "resilience" in the diffusion strategy.

The diffusion strategy emphasizes wide diffusion of incremental technological improvements throughout the whole industry as soon as they become available and it is administered under the direction of the existing industry, particularly the apprenticeship system.

The third strategy, represented uniquely by Japan, is described by Ergas as a hybrid between the diffusion and the heroic strategies. It consists of a mix of a few heroic high-tech projects, centrally orchestrated by government with strong industry input, and carefully chosen for their long-term commercial leverage combined with an aggressive diffusion strategy over a broad area of more traditional technology and industry. Japan's greatest success has been in what Ergas labels the consolidation phase of the technology cycle – the most rapidly rising part of the curve following the appearance of the dominant design. It has been able to exploit this phase primarily through emphasis on excellence in manufacturing combined with close integration between product design and manufacturing. Japan has been able to enter the market early in the consolidation phase partly through its unexcelled system for monitoring world developments, which has enabled it to move quickly in acquiring key complementary assets as well as protecting its domestic market while the selected technology was still in the emergent phase. One conclusion of Ergas is that both the USA and Europe have tended to stress excellence in R&D somewhat at the expense of diffusion. Even in their heroic R&D projects such as the VLSI program and the fifth generation computer project, Japan seems to treat their joint industry R&D efforts as much as a learning process for the participating people who later return to their own firms as for a source of original knowledge and technique.

The weakness of the Japanese strategy has frequently been pointed out by the Japanese themselves. It is of two kinds. First, the newly emerging industrial countries of Asia, the NICs, can use the same strategy as Japan, taking advantage of their lower labor costs unless Japan can learn to move more rapidly upscale toward the higher value-added end of the product mix – a strategy that may put them even more in confrontation with Europe and the United States and the rising trend of protectionism in these regions.

References

Ergas, H., 1987, Does Technology Policy Matter, in B.R. Guile and H. Brooks, eds., *Technology and Global Industry: Companies and Nations in the World Economy*, National Academy Press, Washington, DC, USA.

Guile, B.R. and Brooks, H., eds., 1987, *Technology and Global Industry: Companies and Nations in the World Economy*, National Academy Press, Washington, DC, USA.

Peck, M.J. and Wilson, R.W., 1982, Innovation, Imitation and Comparative Advantage: The Performance of Japanese Color Television Set Producers in the US Market, in H. Giersch, ed., *Emerging Technologies: Consequences for Economic Growth, Structural Change, and Employment, Symposium 1981*, Institut für Weltwirtschaft an der Universitat Kiel, J.C.B. Mohr (Paul Siebeck), Tübingen, Germany, F.R.

Teece, D., 1987, in B.R. Guile and H. Brooks, eds., *Technology and Global Industry: Companies and Nations in the World Economy*, National Academy Press, Washington, DC, USA.

Utterback, J., 1987, in B.R. Guile and H. Brooks, eds., *Technology and Global Industry: Companies and Nations in the World Economy*, National Academy Press, Washington, DC, USA.

CHAPTER 4

Schumpeterian Waves of Innovation? Summarizing the Evidence

Alfred Kleinknecht

> ***There is one hypothesis, now out of fashion, that I would like to back. That is Schumpeter's theory of bouts of investment induced by major technical discoveries. While the new methods are being installed, there is brisk investment and general prosperity, but, after a time, an overshoot is bound to occur, so that excess capacity emerges and brings investment down. I should be prepared to bet that, when the detailed history of the twenty-five years after 1945 comes to be written, it will be seen to have had the character of a boom ... while now there is a formidable overexpansion*** [Joan Robinson, 1979].

4.1. Introduction: Schumpeter versus Kuznets

In his 1939 book, *Business Cycles*, Schumpeter argued that the long-run development of industrial capitalism is characterized by waves of accelerated and decelerated economic growth of some 50 years each. Schumpeter distinguished three such waves:

- First wave: "Industrial Revolution Kondratieff" with an upswing from 1787 to 1814 and a downswing from 1814 to 1842.
- Second wave: "Bourgeois (or Railway) Kondratieff" with an upswing from 1843 to 1869 and a downswing from 1870 to 1897.
- Third wave: "Neo-mercantilist Kondratieff" with an upswing from 1898 to 1924 and a downswing from 1925 onward.

A bold extrapolation of the above scheme would lead us to consider the period between the two World Wars as well as the 1970s and 1980s as downswings of the third and fourth Kondratieff waves, while the 1940s up to the early 1970s would be regarded as the upswing phase of the fourth Kondratieff. A renewed upswing of the world economy would then have to be expected sometime in the 1990s.

According to Schumpeter (1939), each of the above-named upswings can be linked to the emergence and rapid growth of new industrial activities, which were initiated by radical innovations. The subsequent downswings are due to the exhaustion of innovative growth impulses. In order to produce fluctuations which are visible in macroeconomic data, radical innovations should not be randomly distributed over time but should come about in clusters or waves.

In his famous review of Schumpeter's *Business Cycles*, Kuznets spoke of a "host of crucial questions and disturbing doubts" (Kuznets, 1940, p. 262). His criticism referred to three topics in particular. Firstly, Schumpeter had failed to give evidence that long waves not only are a price phenomenon, but also exist in "real" indicators of general economic activity (see Kuznets, 1940, p. 267); secondly, Schumpeter's explanation of the alleged long waves implied some bunching of radical innovations which still remained to be empirically proved (see Kuznets, 1940, p. 263); thirdly, Schumpeter had also failed to give a convincing explanation of *why* such a bunching should occur (see Kuznets, 1940, p. 262ff.).

In retrospect, it seems fair to admit that Kuznets has been essentially right on all three points of critique. As theorizing on long waves more or less stagnated during the 1950s and 1960s, the critical questions raised by Kuznets have remained unanswered. On the other hand, Schumpeter's theoretical propositions, if correct, are likely to have some obvious and far-reaching consequences for our understanding of long-run economic growth.

In this paper I shall give particular attention to the second point of Kuznets's critique: Is there any evidence of a discontinuous occurrence of radical innovations? Kuznets's first point has been addressed elsewhere, leading to the conclusion that in a number of industrial core countries there is indeed evidence of a significant long-wave pattern in indicators of general economic activity, at least during the last hundred years (Bieshaar and Kleinknecht, 1984; see also the comment by Solomou, 1986b, and the reply by Bieshaar and Kleinknecht, 1986). Moreover, Kuznets's point of how to explain a possible bunching of innovations has been discussed extensively in Kleinknecht (1987). The explanation presented there ("depression-trigger" hypothesis), although still being debated (see, e.g., Coombs, 1987), is beyond the scope of this paper.

4.2. The Debate on Basic Innovation Clusters

In recent years, various attempts have been made to collect long-run historical innovation indicators, and particularly to distinguish a few radical breakthroughs in technology from the large stream of smaller piecemeal changes.

To put it metaphorically: there is a real difference between innovators who introduce improved horse cars and those who abolish horse cars by introducing railways or automobiles. A number of imaginative notions have been introduced to describe this difference in more general terms. For example, Dosi (1982) recommends that innovations that establish new "technological paradigms" be distinguished from innovations that occur within existing paradigms. Others speak of "basic innovations" versus "improvement innovations" (e.g., Mensch, 1975; van Duijn, 1983; Haustein and Neuwirth, 1982) or of "New Technology Systems" (Clark *et al.*, 1983) or of "New Technological Webs" (Roobeek, 1987) or simply of "Major" or "Radical" innovations.

An early attempt by Mensch (1975) to verify the hypothesis that "basic innovations" occur in clusters has been received with skepticism (see, e.g., Scholz, 1976; Mansfield, 1983). In their detailed criticism, Clark *et al.*, (1981) pointed to serious problems in Mensch's database. They refer to topics such as the representativeness of his data source, his selection procedure, and the determination of innovation years (see Clark *et al.*, 1981, p. 148f.).

Their critique has triggered more intense research efforts on long-run innovation patterns, which I have treated elsewhere more exclusively (Kleinknecht, 1987). The results of my examination of various independent sources of long-run innovation indicators eventually confirmed that Clark *et al.* have been right in criticizing the fact that the original Mensch list of "basic innovations" did indeed underestimate the frequency of basic innovations during the "early upswing" phase of the long waves; in other words, the discontinuity in the rate of major innovations does not manifest itself in narrow *clusters* during the depth of the depressions (1880s, 1930s) as hypothesized by Mensch, but in virtual *waves* of major innovations.

A stylized scheme of economic long waves as derived from the econometric test of long waves by Bieshaar and Kleinknecht (1984, 1986) as well as from inspection of innovation data (see Kleinknecht, 1987) is given in *Table 4.1.*

Table 4.1. Scheme of upswings (+++) and downswings (– – –) of long waves.

Economic wave	1873– – –1893+++1913– – –1939+++1974– – –...
Innovative wave (12-year lead)	1861– – –1881+++1901– – –1927+++1962– – –...

In view of the evidence derived from various data sets (including their own data), Clark *et al.* have meanwhile admitted that there might indeed exist a bunching of innovations, while advocating a different causal explanation (see Clark *et al.*, 1983, p. 74f.).[1]

Following that line, emphasis now seems to shift toward how to explain the observed bunching of innovations (see, e.g., the comment by Coombs, 1987). Apart from that development, however, there has recently been a contribution by Solomou (1986a, 1986b) which again radically questions the empirical evidence. The next section will be dedicated to that critique.

4.3. The Solomou Critique

Solomou (1986a, 1986b) examined samples of "basic innovations" by Mensch (1975) and van Duijn (1983) as well as a sample of "important innovations" by Kleinknecht (1981) as derived from Mahdavi (1972). He concluded that these data are compatible with his random-walk (or random-shock) hypothesis rather than with a long-wave perspective. Besides doing some statistical explorations, which will be dealt with further below, Solomou makes several critical remarks on the nature of the data. These can be summarized as follows:

(1) In assembling data on basic innovations one is adding up cases of different importance; certainly, some cases are more "basic" than others and hence some weighting procedure would be desirable.
(2) The randomness of Mensch's selection procedure may be doubted (see, e.g., the critique by Clark *et al.*, 1981, p. 148f.).
(3) If the argument about a relationship between market structure and innovation is valid, then market structure changes between the nineteenth and the twentieth century would make any intertemporal comparison of innovation rates a problematic exercise.
(4) Since the majority of innovation cases had its origin in the USA, world innovation rates should be linked to the alleged Kuznets-cycle pattern in American economic growth.

Before responding in more detail to points (1) and (2) (which appear to be reasonable points of critique), a few remarks need to be made on points (3) and (4).

As to market structure and R&D activity, the classical survey by Kamien and Schwartz concludes that empirical studies (being based on shaky data, of course) give only little support to a positive relationship (1983, p. 104). Moreover, "Investigation of the supposition that large firms have the best innovative talent have disclosed almost the exact opposite. The largest firms appear to be far less efficient innovators than smaller rivals" (Kamien and Schwartz, 1983).

But even if valid, in a long-run historical perspective, changes in market structure would probably have to be conceived as a rather continuous and irreversible process. Consequently, the argument could probably explain a trend increase in innovation rates rather than the type of wave pattern that will show up in our data further below [2] – except if one would argue that market structure changes occur in long waves (this would indeed be a remarkable contribution to the current long-wave debate!).

Solomou's argument about linking world innovation rates to the Kuznets-cycle pattern in American economic growth [point (4)] is misleading in at least two respects. Firstly, there are reasons to believe that the Kuznets cycle is a statistical artifact, due to problematic filtering effects that result from the use of first differences in detrending economic time series (see Bieshaar and Kleinknecht, 1986, p. 190f.). Secondly, provided that the Kuznets cycle exists at all, there seems to be some agreement that it is restricted to the period before World War I (see, e.g., the discussion in Rostow, 1975), while US world market hegemony has emerged during the twentieth century only and is most obvious after World War II.

While rejecting points (3) and (4), the first and second point of the critique should be taken seriously. It is a problematic exercise to add up innovations of quite different importance and complexity, and the rate of innovation observed may be biased by the personal whims and preferences of the compiler. For example: a compiler may include cases of "basic" innovations that other compilers would classify as "minor" cases; or, a researcher may use problematic sources and investigate certain historical periods more carefully than other periods. On the other hand, trusting the personal integrity of researchers, one might hope that such biases (although unavoidable) will remain within acceptable limits. In the following, I shall add up the sets of innovation data by Mensch (1975) [3] and van Duijn (1983), adding another set of basic innovation data by Haustein and Neuwirth (1982), which has not been considered by Solomou (1986a, 1986b).

In doing so, it is hoped that a possible bias from personal judgment by an individual compiler will be reduced. The adding up of *all* cases from the three samples implies some weighting procedure, since cases that are included in all three samples (and which can therefore most confidently be considered as "basic" innovations, since all three authors agree upon these cases) are counted three times. Cases that are included in two out of the three sources (which might still be considered as relatively "safe" cases of basic innovations) are counted twice. The category of basic innovations that is reported by one of the three sources only (and which is most likely to cover a number of doubtful cases) is counted only once. Because of the implicit weighting procedure, we would expect the resulting "supersample" to give a more reliable indication of long-run innovation patterns than the isolated consideration of an individual source could do. The supersample is displayed in *Figure 4.1.*

To get an idea about the reliability of the supersample, it is interesting to see how far the three underlying sources overlap. A schematic presentation of

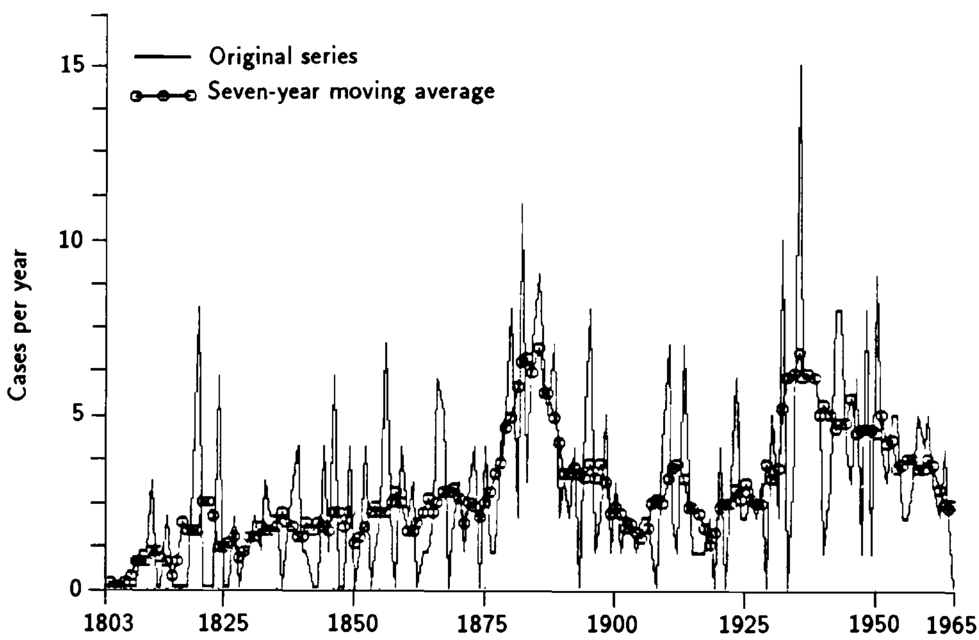


Figure 4.1. All basic innovations from three sources (supersample) from 1800 to 1968.

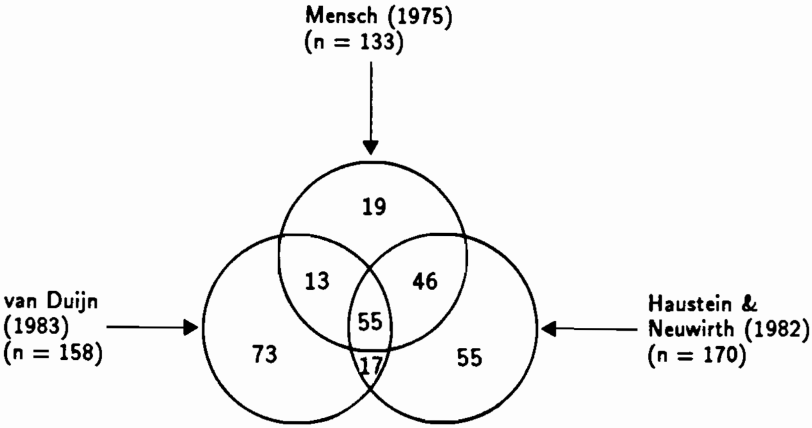


Figure 4.2. **Overlap between three samples of basic innovations from 1800 to 1968.**

overlaps is given in *Figure 4.2.* It should be noted that the numbers in *Figure 4.2* may be subject to some counting errors, which are due to the nature of the data: quite frequently, different sources use a slightly different description of the same innovation case; besides, counting is sometimes complicated because two sources may have considered a different aspect of the same type of innovation (e.g., one source is covering the first commercially successful steamship, while the other source takes the year of the first Atlantic crossing on a steamship). Moreover, even for identical events, often diverging innovation years are given (fortunately, most differences in innovation years remain within the range of a few years). In spite of such problems, *Figure 4.2* may give at least a rough indication of the overlap between the three sources.

It can be seen in *Figure 4.2* that the Mensch (1975) sample shows strong overlap with the other two samples, while the van Duijn (1983) and the Haustein and Neuwirth (1982) samples have only a modest overlap. This can be explained by the fact that the Mensch sample (being published earlier) has been known to van Duijn and to Haustein and Neuwirth, while the latter two have been compiling their samples independently of each other. It is remarkable to see that a number of the Mensch cases have not been included in the samples of the other two compilers, which indicates that they must have examined the Mensch sample quite critically.

It should be noted that when forming the supersample, I deliberately did not interfere with the data, which means that no case was added or omitted; even in the case of diverging innovation years, no innovation year was changed. Besides the above-described supersample, other exercises were done that are not documented here. For example, when adding up all the cases from the three sources and omitting those cases that are named in one source only, a pattern similar to that in *Figure 4.1* was obtained.

While the supersample certainly is an improvement as compared with the individual sources, it should be noted that the wave pattern in the time distribution of basic innovations does not depend on weighting. This will become clear from our test on the significance of differences in mean innovation rates for various *a priori* periods, which brings us to another point of the Solomou critique.

Solomou is right in arguing that for testing the significance of long-run innovation patterns, a test on differences in means between certain *a priori* periods is more appropriate than the runs test as applied by Mensch (1975). It is also correct, that the *cyclicity* of innovation waves (i.e., their endogenously caused regular recurrence) cannot be proved by any quantitative test, simply because of the low number of waves observed (any proof of cyclicity being left to a theoretically convincing endogenous explanation of the turning points). As Solomou rightly points out, however, one can test a "weak" Kondratieff hypothesis, testing whether observable innovation patterns behave according to what one would expect from a long-wave view (ibid., p. 102).

In doing so, I shall apply a one-sided t-test, testing whether the mean number of innovations during the upswing (+++) periods in *Table 4.1* is significantly higher than during the downswing (– – –) periods (and vice versa). The t-test (which is not *exactly* a student t) is defined as follows:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\delta_1^2}{N_1} + \frac{\delta_2^2}{N_2}}}$$

where: $\bar{x}_1$ and $\bar{x}_2$ are the sample means, δ_1^2 and δ_2^2 are the sample variances, and N_1 and N_2 are the sample sizes.

Because of the smaller sample sizes, the use of a t-test for this statistic is more cautious (giving lower levels of significance) than the use of a z-test (as has been done by Solomou, 1986a, p. 108). Moreover, since the hypothesized *direction* of the differences is clearly determined, a one-sided test will be applied. As in Solomou's test it is assumed that the variance during the subsequent upswings and downswings is *not* equal. In the case of the t-test, this assumption implies a considerable loss of degrees of freedom, following the "safe rule" as outlined in Wonnacott and Wonnacott (1977, p. 214).

Table 4.2 documents the results from application of the t-test to the supersample. The test was also applied to the three sources individually, the results being reported in *Table 4.A1* of the Appendix. Documentation in this paper is restricted to the results that were achieved when handling a 12-year lead of the innovation wave over the economic wave as hypothesized in *Table 4.1*. To test the robustness of the results with regard to slight variations in lead times, a 10- and a 15-year lead was also tried. The results differed only slightly, so that the same conclusions could have been drawn, using a slightly different periodization. *Table 4.2* confirms that the fluctuations observed in *Figure 4.1* above can clearly be distinguished from statistical random fluctuations.

Table 4.2. Calculations of t-test for upswings and downswings of long waves: the supersample.

Period	*Mean*	*SD*	*SE*	*t-values*	*d.f.*	*Prob.*
1861–1881	2.6667	2.1525				
			0.8006	1.843	20	0.039
1881–1901	4.1428	2.9712				
			0.7441	2.481	20	0.011
1901–1927	2.2962	1.8976				
			0.6291	3.370	26	0.001
1927–1962	4.4166	3.0740				
			0.7979	2.491	6	0.024
1962–1968	2.4285	1.6183				

In interpreting *Table 4.2*, one has of course to be aware that, even if our weighting procedure does imply some improvement, it certainly cannot satisfy all possible objections uttered by skeptics. As has already been indicated above, nobody who has ever been working in the field of innovations research, needs to be reminded of the numerous problems concerning topics such as the representativeness of sources, the randomness of selection principles, the distinction between "major" and "minor" events, an appropriate sample size, or the determination of innovation years.

If, in spite of all these problems, we want to arrive at a somewhat safe judgment about Schumpeter's above-sketched hypothesis, we should compare evidence from as many sources as possible. Fortunately, owing to the painstaking work by Baker (1976), there is still another long-run technology indicator that has been collected independently of the above basic innovation sources, and which will be considered in the following.

4.4. Testing the Baker Data

While the above data on basic innovations consist of years when the first successful commercialization of new products or processes, perceived to be of fundamental importance, occurred, Baker (1976) collected about 1,000 "breakthrough" patents which refer to 363 important items (these items range, in alphabetical sequence, from the addressograph up to the zip fastener). It should be mentioned that the basic innovation data are, in principle, *world* innovation data, whereas Baker's breakthrough patents are mainly patents registered at the British Patent Office. It can nonetheless be argued that they might be taken as a *world* innovation indicator, since "The United Kingdom's role in the international world of commerce has been of sufficient importance throughout the history of the patent system to ensure that most inventions of significance would have been subject of patent applications in this country" (Baker, 1976, p.21).

As compared with "direct" innovation data, the Baker patent data have three notable drawbacks. Firstly, the year of publication of a breakthrough patent on a new item is not necessarily identical with the year of the innovation (i.e., the first successful commercialization of the item), although it should come reasonably close to it. Secondly, the Baker sample covers a certain number of key patents that are related to radical *inventions* rather than *innovations*. Thirdly, a few cases are related to *improvement* rather than to "basic" innovations (see also the discussion in Kleinknecht, 1987).

These points are likely to constitute a bias in favor of a random-walk pattern. Consequently, we would expect fluctuations, as hypothesized in *Table 4.1*, to be less accentuated in the Baker data than in "pure" innovation data. A comparison between the Baker data in *Figure 4.3* and the supersample of basic innovations in *Figure 4.1* seems to confirm this.[4] Nonetheless, the results from application of the above-defined *t*-test to the Baker data, being documented in *Table 4.3*, confirm that the hypothesized fluctuations are still significant, even though significance levels are generally a bit lower than in *Table 4.2*.

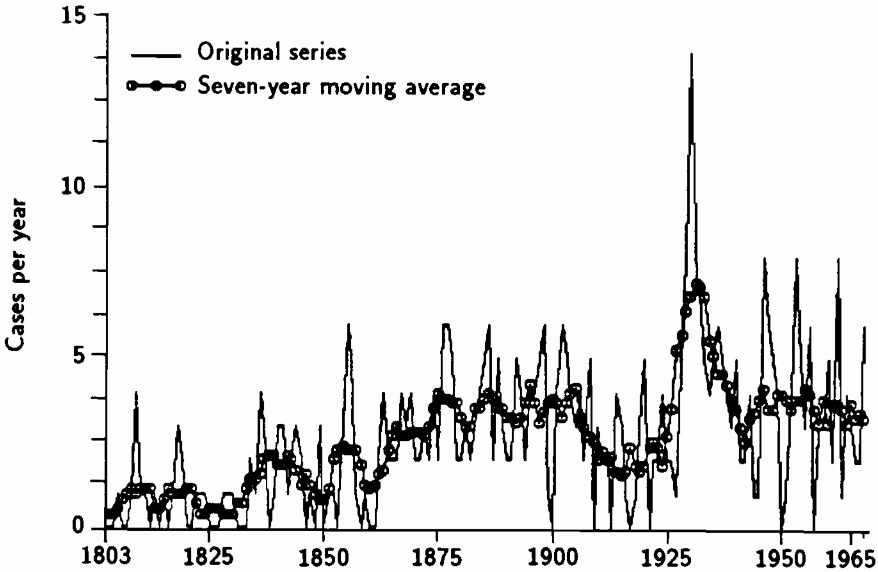


Figure 4.3. Product-related breakthrough patents from Baker (1976), according to classification in Kleinknecht (1987).

From *Tables 4.2* and *4.3*, as well as from the test on the individual sources in *Table 4.A1* (Appendix), it can be concluded that between 1881 and 1962 there is evidence of three successive periods of above- and below-average rates of innovations, the differences being distinguished from random variations at high levels of significance. Judging from the (somewhat more reliable) supersample in *Table 4.2*, this holds even for the period from 1861 to 1968.

As to the very last period (1962–1971), it is of course right that it is often only in retrospect with a certain time lag that one can decide what are "major" innovations or "minor" ones. Hence, the result from *Table 4.2* can be taken only as a very preliminary indication of a decline of innovation rates during the 1960s. Nonetheless, judging from conventional wisdom, it does not seem to be too bold

Table 4.3. Calculations of *t*-test for upswings and downswings of long waves: the Baker data.

Period	*Mean*	*SD*	*SE*	*t-value*	*d.f.*	*Prob.*
1861–1881	2.9523	1.4992				
			0.4832	0.985	20	0.160
1881–1901	3.4285	1.6300				
			0.4853	1.722	20	0.052
1901–1927	2.5925	1.7155				
			0.5761	3.262	26	0.002
1927–1962	4.4722	2.8333				
			0.8331	1.046	9	0.200
1962–1971	3.6000	2.1705				

a prognosis that the 1960s and 1970s will eventually turn out to have been a period of poor innovation performance, followed by a renewed upsurge of radical innovation in the 1980s and 1990s. This would also be a logical implication of my theoretical explanation of innovation waves that is beyond the scope of this paper (see Kleinknecht, 1987, for an extensive discussion).

Although a theoretical explanation is important for the issue of *cyclicity* of the observed waves, this paper is restricted to the statistical evidence that has been questioned by Solomou (1986a, 1986b). Summarizing the above considerations, we can say that Solomou, being right in his critique of Mensch's cluster hypothesis (and its statistical support) draws the wrong conclusions.[5] Innovation flows have *not* been constant. Besides a twentieth-century wave of radical innovations, there is evidence of a period of accelerated innovation activity in the 1880s and 1890s, followed by a deceleration up to the late 1920s, which Solomou will not be able to explain by whatever exogenous shock event.

Solomou's random-walk hypothesis may hold, however, for the period of early capitalism. Optical inspection of the various time series suggests that, up to the mid-nineteenth century, the flow of innovations in *aggregate* data experienced only a monotonous increase. This suggests that Schumpeter's innovation long-wave hypothesis as a macroeconomic phenomenon is valid only for developed capitalism.[6]

4.5. Suggestions for Further Research

This paper was restricted to empirical evidence of long waves in the incidence of major innovations, which is of course closely related to the issue of long waves in economic life. The explanation of innovation waves that has been put forward in Kleinknecht (1986, 1987) has been discussed controversially. However, "alternative" explanations that stress the importance of "science push" and institutional change (Clark *et al.*, 1983; Coombs, 1987), or which focus on the "social structure of accumulation" (Gordon *et al.*, 1982) are not necessarily inconsistent with my argument about a restructuring of the technological base of capital accumulation being triggered by a prolonged depression. However, an explanation that integrates the various views would be a task for another paper.

Another interesting issue would consist of linking innovation waves to long-run profit rates. The idea of long waves in aggregate profit rates has recently been advocated by several theorists, such as Fontvieille (1985), Menshikov and Klimenko (1985), Poletayev (1985), and Reati (1986). In a disaggregated analysis of West German manufacturing profit rates from 1950 to 1977, I have argued that sectors that can be closely related to the twentieth-century wave of major innovations did exercise during the 1950s (and in part during the 1960s) a counteracting influence against a rapid fall of the aggregate profit rate (Kleinknecht, 1987a). Doing such an analysis for other countries and periods (and, if possible, at a finer level of aggregation), might shed a new light on the discussion of the tendential fall of the profit rate.

Interpreting the innovation waves from the viewpoint of demand theory may be another research topic, being particularly attractive to Keynesian economists. In explaining why innovation waves may cause waves of expansion and contraction in the economy, one would have to consider that launching of an innovation involves considerable investment in R&D, in know-how, and eventually in the buildup of production facilities; the powerful multiplier effects that result from such investments may be conceived as a positive function of the degree of radicalness of an innovation, the amount and impact of subsequent (major and/or minor) innovations, and the degree of market success (diffusion).

Of course, the boom created by such innovation multipliers (which may end in an overshooting such as described in the above quotation by Joan Robinson) still needs to be adequately modeled before it can be integrated into long-run macro-models.

The relationship between demand and innovation still has another implication. To the extent that the "demand-pull" hypothesis (which is not necessarily inconsistent with my "depression-trigger" hypothesis) is valid in explaining innovation, it has an impact on government demand management that has been largely neglected even by Keynesian economists. Government demand, besides having the multiplier effects that are well known from the textbooks, may influence the flow of innovations (and in doing so create extra demand by means of the above-mentioned "innovation multiplier"). Of course, from a Schumpeterian viewpoint, one would not advocate macroeconomic demand impulses. The latter may be (in part) even counterproductive in that they (also) contribute to preserve existing product lines. Rather one would advocate *specific* demand impulses that are directed toward assisting the emergence of new industrial activities; i.e., government demand may systematically increase the chances of new technological options to survive in the process of Darwinian selection on the marketplace. Such a demand policy would have the advantage of not only increasing effective demand as such, but also of allowing political choices concerning socially desirable new technologies.

The above-sketched arguments may indicate that the hypothesis of innovation waves, if correct, calls for a lot of research still to be done.

Acknowledgment

A first draft of this paper has been prepared during my stay at the International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. I wish to thank the colleagues from the Technology, Economy, and Society Program, and, in particular, Robert U. Ayres, for providing a hospitable working environment and for helpful discussions.

Notes

[1] Discussing innovation data on the twentieth-century chemical industry, they concluded: "All of this supports the notion of bunches of basic inventions and innovations leading to the take-off of new industries. . . . It does not, however, demonstrate any direct connection between this process and the 'trigger' of depression" [Clark *et al.*, 1983, p. 74f.].

[2] A similar argument is likely to apply with respect to other long-run structural changes, such as the rise of the professional R&D lab during the twentieth century.

[3] For the twentieth century I took the Mensch data as revised by Clark *et al.*, (1983, p. 68f.).

[4] It should be mentioned that the Baker data in *Figure 4.3* and *Table 4.3* refer to product-related breakthrough patents. The process-related patents show a different pattern. A detailed discussion and documentation of the classification of the Baker cases by product versus process patents can be found in Kleinknecht (1987, ch. 4).

[5] The main difference between my results and those by Solomou does not seem to be due to the use of a different test formula, but rather to a different *periodization*. For example, Solomou applies the Mensch (cluster) periodization to the van Duijn data. Since the latter show a (broad) wave pattern rather than a narrow cluster pattern, it is not surprising that Solomou finds an almost perfect random-walk pattern (see Solomou, 1986a, p. 109).

[6] Recent work at the International Institute for Applied Systems Analysis at Laxenburg, Austria, suggests that the diffusion paths of specific technologies (e.g., in the energy and transportation sector or in the steel industry) seem to fit into the framework of Kondratieff long waves, even during those periods in the eighteenth and nineteenth century for which evidence of macroeconomic long waves is poor; see Marchetti (1986), Nakicenovic (1986), or Grübler and Nakicenovic (1987).

Appendix

Table 4.A1. Calculations of t-test for upswings and downswings of long waves.

Basic innovations according to van Duijn (1983)

Periods	*Mean*	*SD*	*SE*	*t-value*	*d.f.*	*Prob.*
1861–1881	0.8571	0.7928				
			0.3278	0.871	20	0.192
1881–1901	1.1428	1.2761				
			0.3137	2.225	20	0.019
1901–1927	0.4444	0.7510				
			0.2428	2.287	26	0.015
1927–1962	1.0000	1.1710				
			0.2478	2.824	9	0.015
1962–1971	0.3000	0.4830				

Basic innovations according to Haustein and Neuwirth (1982)

Periods	*Mean*	*SD*	*SE*	*t-value*	*d.f.*	*Prob.*
1861–1881	0.8571	0.9102				
			0.4208	2.149	20	0.022
1881–1901	1.7619	1.7001				
			0.4042	2.343	20	0.015
1901–1927	0.8148	0.8337				
			0.2873	2.191	26	0.019
1927–1962	1.4444	1.3404				
			0.4008	0.428	10	insignif.
1962–1972	1.2727	1.1037				

Mensch's 20th-century basic innovations as revised by Clark et al. (1983)

Periods	*Mean*	*SD*	*SE*	*t-value*	*d.f.*	*Prob.*
1901–1927	0.4814	0.8024				
			0.2411	3.302	26	0.002
1927–1962	1.2777	1.1112				
			0.3507	2.013	7	0.040
1962–1968	0.5714	0.7867				

References

Baker, R., 1976, *New and Improved ... Inventors and Inventions that Have Changed the Modern World*, British Museums Publications, London, UK.

Bieshaar, H. and Kleinknecht, A., 1984, Kondratieff Long Waves in Aggregate Output? An Econometric Test, *Konjunkturpolitik* **30**(5), October.

Bieshaar, H. and Kleinknecht, A., 1986, Kondratieff Long Waves? Reply to S. Solomou, *Konjunkturpolitik* **32**(5), October.

Clark, J., Freeman, C., and Soete, L., 1981, Long Waves and Technological Developments in the 20th Century, in D. Petzina and G. van Roon, eds., *Konjunktur, Krise, Gesellschaft*, Klett-Cotta, Stuttgart, Germany, F.R.

Clark, J., Freeman, C., and Soete, L., 1983, Long Waves, Inventions and Innovations, in C. Freeman, ed., *Long Waves in World Economic Development*, Butterworths, London, UK.

Coombs, R., 1987, Long Waves, Depression and Innovation, Comments on Kleinknecht, *De Economist* **135**, September.

Dosi, G., 1982, Technological Paradigms and Technological Trajectories, *Research Policy* **11**:147.

van Duijn, J.J., 1983, *The Long Wave in Economic Life*, George Allen & Unwin, London, UK.

Fontvieille, L., 1985, Contribution a l'Etude des Fluctuations Longues: Le Capital Fixe Productif a la Compagnie d'Anzin 1716–1897, *Cahiers de Rechercces et Travaux de l'Institut d'Histoire Economique et Sociale de l'Universitè de Paris I*, Paris, France.

Gordon, D.M., Edwards, R., and Reich, M., 1982, *Segmented Work, Divided Workers*, Cambridge University Press, Cambridge, UK.

Grübler, A. and Nakicenovic, N., 1987, *The Dynamic Evolution of Methane Technologies*, WP-87-2, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Haustein, H.-D. and Neuwirth, E., 1982, Long Waves in World Industrial Production, Energy Consumption, Innovations and Inventions, and Their Identification by Spectral Analysis, *Technological Forecasting and Social Change* **22**(1), September.

Kleinknecht, A., 1981, Observations on the Schumpeterian Swarming of Innovations, *Futures* **13**(4), August.

Kleinknecht, A., 1986, Long Waves, Depression and Innovation, *De Economist* **134**, March.

Kleinknecht, A., 1987, *Innovation Patterns in Crisis and Prosperity: Schumpeter's Long Cycle Reconsidered*, Macmillan, London, UK, and St. Martin's Press, New York, NY, USA.

Kleinknecht, A., 1987a, Rates of Profit in the Long Wave, in T. Vasko, ed., *The Long-Wave Debate*, Springer-Verlag, Berlin, Heidelberg, Germany, F.R., New York, NY, USA.

Kuznets, S., 1940, Schumpeter's Business Cycle, *American Economic Review* **30**(2):257–271.

Mahdavi, K.B., 1972, *Technological Innovation. An Efficiency Investigation*, Beckmans, Stockholm, Sweden.

Mansfield, E., 1983, Long Waves and Technological Development, *American Economic Association (Papers and Proceedings)* **73**(2), May.

Marchetti, C., 1986, A Fifty-Year Pulsation in Human Affairs, *Futures*, June.

Mensch, G., 1975, *Das technologische Patt*, Umschau, Frankfurt, Germany, F.R., [English translation, Cambridge, MA, USA, 1979].

Menshikov, S. and Klimenko, L., 1985, On Long Waves in the Economy, in G. Bianchi, *et al.*, eds., *Long Waves, Depression and Innovation*, CP-85-9, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Nakicenovic, N., 1986, The Automobile Road to Technological Change, *Technological Forecasting and Social Change* **29**:309–340.

Poletayev, A.V., 1985, *Pribyl Americanskikh corporatsy*, [Profits in American Enterprises, in Russian], Nauka, Moscow, USSR.

Reati, A., 1986, The Rate of Profit and the Organic Composition of Capital in the Post-1945 Long Wave, *Review* **9**, Spring.

Robinson, J., 1979, Solving the Stagflation Puzzle, *Challenge*, November–December.

Roobeek, A., 1987, *De Rol Van de Technologie in de Economische Theorievorming*, Scheltema, Amsterdam, Netherlands [in Dutch].

Rostow, W.W., 1975, Kondratieff, Schumpeter, Kuznets: Trend Periods Revisited, *Journal of Economic History* **35**.

Scholz, L., 1976, Kontroverse um das Technologische Patt, *IfO-Schnelldienst*, 29/30/1976 [in German].

Schumpeter, J.A., 1939, *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, 2 Vols., McGraw-Hill, New York, NY, USA.

Solomou, S., 1986a, Innovation Clusters and Kondratieff Long Waves in Economic Growth, *Cambridge Journal of Economics* **10**, June.

Solomou, S., 1986b, Kondratieff Long Waves in Aggregate Output? A Comment, *Konjunkturpolitik* **32**(5), October.

Wonnacott, T.H. and Wonnacott, R.J., 1977, *Introductory Statistics for Business and Economics*, 2nd ed., Wiley, New York, NY, USA.

CHAPTER 5

A Theory of Growth Rate Discontinuities

Robert U. Ayres

5.1. Introduction

The circumstances of accelerating use, and possible near- or medium-term exhaustion, of fossil energy resource – together with major uncertainties as to the feasibility, cost, and timing of downstream substitutes – constitute a challenge for economic analysis. Several frameworks are possible. The case where technology offers no substitution possibility was examined many years ago by Gray (1914) and Hotelling (1931). Decades later, Nordhaus (1973) considered a variant case in which the supply curve becomes infinite at some finite price, where the so-called "backstop" technology takes over and provides unlimited energy availability. Stiglitz (1974, 1979) assumes technological progress occurs at a constant rate, regardless of policy, indefinitely.

Dasgupta and Heal (1974, 1979) introduced a different twist. In their earlier models, the new technology eliminates the *need* for the resource, but it arrives exogenously and costlessly at some uncertain time in the future. In more recent work these authors, as well as Kamien and Schwartz (1978, 1982) and others, have examined variations in which the new development itself becomes endogenous and costly. In the context of energy analysis, these models largely retain the backstop concept, the focus being on a single, millennial, breakthrough technology and on optimal policy during the interim period. A useful recent summary of the status of this literature has been given by Huettner (1981).

A simple framework adopted in the present paper differs from some of those cited above in several ways.[1] First, no technological millennium, in the Nordhaus sense, is envisaged. Rather, technical progress is identified with a

Reprinted from *Journal of Environmental Economics and Management* 15 (1988):439–461.

continuously increasing function of technological knowledge, T, which is taken to be an explicit, endogenously determined factor of production. Second, all factors of production, including technological knowledge, are assumed to be forms of condensed or "embodied" forms of information. Information is used in the technical sense introduced by Hartley (1928) and elaborated by Shannon (1948), Brillouin (1953), and others. The substitutability (or interconvertibility) of factors of production follows naturally. Third, following from the above, both resource inputs and outputs of the production process can also be thought of as forms of condensed information and measured in "bits." The ratio of aggregate outputs (e.g., GNP) to input thus becomes a natural Generalized measure of the state of technology at a given time.

Much more can be said on the last point. It is one to which many economists raise objections, almost reflexively. However, many of the common objections are rooted in intuitive and rather imprecise uses of the concept of information. The following section is intended to provide some explanatory background material on this topic. It can be skipped by any reader who is either (a) already moderately comfortable with concepts of standard information theory (as used by engineers and physicists) or (b) willing to suspend disbelief and accept – for purposes of argument – that all economic quantities (labor, capital, resources, outputs) can be quantitatively measured in the same physical unit ("bits").

One caveat is essential at this point. The assertion that all economic quantities can be measured in bits does not preclude their also being measurable in value units (e.g., dollars). The two kinds of units need not be proportional, any more than the relative prices of two materials necessarily coincide with their relative masses. The model introduced later does not seek to maximize the absolute information content of economic output. It does seek to maximize the *utility* of that information output. Thus, a subtle and possibly controversial feature of this model is that it assumes the existence of such a utility function, i.e., a consistent relationship between the information embodied in final products and services produced by the economy and the utility thereof. If there is to be a debate, it should probably focus on whether such a utility function can consistently be determined.

5.2. Information

Technically speaking, information is a measure of *uncertainty* (Shannon, 1951), of negative entropy or *negentropy* (Brillouin, 1953), or of *distinguishability* or generalized distance (Tribus & McIrvine, 1971). The more distinguishable or nonrandom a subsystem is, the more information it embodies. This is true of telegraphic or telephonic messages, wireless transmissions, photographs, atomic or molecular assemblages, materials, shapes, and physical structures. It is also true of organizations and social systems.

Methods for numerical computation of information content are available for communications applications and for homogeneous physical–chemical systems. Computational schemes can be developed, in principle, for the more complex cases. In general, the information content of a manufactured thing corresponds roughly to the number of symbols or words that would be required to describe it efficiently (e.g., in a computer program).

Solar radiation is information-rich because it is highly distinguishable (in terms of equivalent black-body temperature) from the low-temperature background radiation. High-quality metal ores contain information because their composition is highly distinguishable from the surrounding earth's crust; purified metals contain even more information for the same reason; and so on.

Knowledge is a *useful subset* of information that can be regarded as a factor of production. Not all information is knowledge, but all knowledge is information. "Useful," in this context, merely means that it contributes to the production of useful goods and services. A more extended discussion of the relationship between information and knowledge has been included as Appendix A.

While knowledge can be assumed to increase, in principle, without physical limit (if one continues investing in R&D), its impact on productivity is assumed to be subject to diminishing returns. Both the assumption of concavity – or diminishing returns – and the assumption that technological knowledge is endogenous to the productive system are in contrast to views in some of the extant economic growth literature. However, one important notion underlying the approach described in this paper is that natural resources, labor, physical capital, and knowledge are all condensed forms of information and therefore mutually substitutable, within limits to be discussed later.

In fact, it requires no great leap of the imagination, at this point, to interpret physical capital stock as knowledge (i.e., useful information) embodied in material form. Similarly, various skill levels of labor can readily be interpreted as knowledge embodied in human workers. When capital equipment depreciates due to wear and tear, the (useful) information content embodied in its design (form and function) is gradually lost. As a cutting tool loses its physical edge, its distinguishability is obviously decreased, as is its economic productivity.

The interpretation of capital and labor as embodiments of knowledge does not alter the desirability of taking into account the fact that the economic system also depends on a continuing flow of available energy or essergy. Available energy (essergy) is the ultimate resource, in the same sense that all other material resources can be extracted from the earth's crust, in principle, if enough energy is available. Energy (essergy) flux from the sun is, of course, the ultimate source of all localized negentropic (information) accumulation on the earth. This being so, the solar energy flux is, in effect, a *flux* of information. Similarly, the earth's store of fossil fuels can be regarded as a stock of information. Some of the latter can be captured and embodied by biological and/or technological processes in other (even more condensed) forms, such as capital goods or products.

5.3. Technical Efficiency and Technological Knowledge

The essential equivalence of resources and energy is widely accepted (e.g., "energy is the ultimate resource"), and the equivalence of useful or available energy and information (negentropy) has already been discussed. Thus, in the final analysis both economic inputs (resources) and economic outputs (goods and services) can be viewed as forms of information. These forms differ primarily in terms of the extent to which information is embodied in composition, structure, shape or form, and knowledge content or quality.

The model discussed hereafter assumes that the modern economic system as a whole is a kind of information processor, which continuously converts massive amounts of crude information (negentropy) into a much smaller quantity of refined information. The latter takes the form of knowledge stocks and human services.[2] Both kinds of information flow are measurable in bits per sec. ***The processing efficiency of the economic system can be defined as the ratio of information output flux to information input flux.*** This statement is both trivial and truly profound, as will be seen.

It is convenient at this point to introduce a variable $E(T)$, where T is a measure of technological knowledge T, such that E is constrained to the range zero to unity. For reasons that will be clearer subsequently, it is convenient to think of E as a generalized efficiency measure. It is convenient to let

$$E = [1 + \exp(T_0 - T)]^{-1}, \tag{5.1}$$

where T_0 is a large number (by assumption) such that $E = 0.5$ when $T = T_0$. Evidently if T_0 is large, E is very small for small values of T ($T << T_0$) and asymptotically approaches unity for very large T ($T >> T_0$). Solving for T,

$$T = T_0 + \ln(E/1 - E). \tag{5.2}$$

The growth of the stock of technological knowledge T can be presumed, for purposes of the model, to follow a simple law, *viz.*,

$$\dot{T} = J, \tag{5.3}$$

where J is the annual creation (or destruction) of new knowledge. J is a function of time, of course. The rate of embodiment (or fixation) of knowledge in capital, labor, products, etc., is presumably proportional to the rate of acquisition of new knowledge owing to R&D over some prior period.

The productivity measure E satisfies a nonlinear differential equation, *viz.*,

$$\dot{E} = E(1 - E)J, \tag{5.4}$$

where J (previously defined) is the aggregate annual rate of addition to the stock of knowledge. It can be seen that E is an elongated more or less S-shaped curve. It is exponentially rising, at first, but after passing a point of inflection, it enters a concave region of saturation, asymptotically approaching unity. If J is a constant, it may be noted that the solution to equation (5.4) is the familiar logistic curve. This qualitative behavior is, incidentally, characteristic of most individual technology measures over time.

As technical efficiency E asymptotically approaches unity (i.e., progress continues for a very long time), the economic system generates the maximum possible output of final services, per capita, from a given resource (crude information) flux. Nothing whatever is implied about the need for physical materials, as such, since materials can always be recycled from the environment if enough energy is available.

5.4. An Optimal Economic Growth Model

I now introduce an explicit optimal growth model incorporating many of the concepts outlined in preceding paragraphs. In this model, it is assumed that labor force is an exogenous variable proportional to population and independent of other economic variables. For the sake of concreteness, let

$$L = bN, \tag{5.5}$$

where N is the total population.

It is conventional in the economic literature to make the usual Malthusian assumption, for convenience, that population N grows exponentially over time, at a constant rate g. This seems simplistic (on biological grounds) and unnecessary. A more reasonable assumption seems to be that humans can, and eventually will, regulate their population to the level that can be supported by the physical environment. In fact, the rate of world population growth has declined significantly in the last 20 years. A simple differential equation having roughly the desired asymptotic behavior is

$$\dot{N} = gN(1 - N/\bar{N}), \tag{5.6}$$

where $\bar{N}$ is the maximum population theoretically sustainable by conventional agriculture, given existing world soil characteristics, rainfall, insulation, and topographic conditions (Pearl, 1922). One need not be concerned at present with the numerical value of $\bar{N}$.[3] I will focus attention, subsequently, on aggregate production and consumption, with the understanding that per capita measures are derivable from them.

Next, consider the stock of fixed (constant vintage) invested capital K. The usual assumed accumulation law is

$$\dot{K} = I - dK, \tag{5.7}$$

where I is the current level of investment and d is the rate of physical depreciation, assumed to be constant, for convenience. The nonnegativity of investment $I \geq 0$ implies that fixed capital cannot be consumed although the stock can decline as a result of depreciation. For internal consistency, K measures the quantity of constant-vintage capital referred to in a given vintage year (e.g., 1985). It is, of course, true that successive technological improvements will tend to increase the capabilities of machines and/or structures built at later times. Thus, a given quantity of constant capital will be equivalent in productive capability to a smaller quantity of current capital, at any future time. This performance improvement reflects the continuous embodiment of new technological knowledge. The specific mathematical form describing the embodiment of technology in capital need not be considered at this stage.

For purposes of this model, it is necessary to define two distinct kinds of capital K_1 and K_2. By assumption, K_1 is used in the production of final goods and services while K_2 is used in the direct capture of solar energy (renewable resources). This is assumed for convenience to be a capital-intensive activity, though it could also be labor-intensive. Thus, we define

$$K = K_1 + K_2. \tag{5.8}$$

Similarly, capital investment has two components:

$$I = I_1 + I_2. \tag{5.9}$$

In the model, crude information (essergy) resources are required to drive economic activity. The quantity of essergy R needed is a function of the total output of goods and services by the economy, $\Pi(K, L, E)$, where $E = E(T)$. Given the view that economic output $\Pi(K, L, E)$ can be measured in terms of information (bits) and resource input R is also a measure of information input, it makes sense to define E as the dimensionless ratio of aggregate outputs Π to aggregate essergy inputs R (both measured in bits)

$$E = \frac{\Pi}{R}. \tag{5.10}$$

Note that this ratio is necessarily less than unity because energy becomes increasingly unavailable (i.e., entropy increases) at each stage of the production process from materials extraction to final assembly. As entropy increases, stage by stage, the total information (negentropy) contained in product-plus- environment necessarily decreases. Thus, equation (5.10) has physical content. In fact, the condition $\Pi/R < 1$ is required by the second law of thermodynamics. Evidently, the essergy resource requirement at any time is precisely

$$R = \frac{\Pi}{E}. \tag{5.11}$$

The supply of essergy R at any given time may come from either of two sources, *viz.*, fossil fuels or some renewable source (such as biomass) originating in the solar flux. In reality fossil fuels are not free by any means, since they must be extracted, processed, and distributed. However, for purposes of the model, it is interesting to assume the existence of an initial stock S_0 of essergy that can be extracted costlessly at any desired rate until it is exhausted.[4] Useful essergy can also be extracted from the sun, but only in proportion to the amount of capital K_2 invested for that purpose. To be consistent with the viewpoint adopted above, it is also convenient to divide aggregate production itself into two components

$$\Pi = \Pi_1\{K_1,\, L,\, E(T)\} + \Pi_2\,\{K_2\}, \tag{5.12}$$

where Π_2 is the output of the "renewable essergy" sector. The latter can be conceptualized as a set of unmanned solar satellites and ground stations, embodying capital K_2 (although it might equally well be some other kind of infrastructure). Since the solar-powered utility sector consumes no essergy, equations (5.10) and (5.11) can be simplified by substituting Π_1 for Π. The essergy resource supply at any time can be written

$$R = -\,\dot{S}_1 + C_2K_2, \tag{5.13}$$

where $\dot{S}_1$ (a negative number) is the rate of change of the stock S_1 of fossil essergy and C_2 is a parameter.[5] Equation (5.13) can thus be rewritten to eliminate R, *viz.*,

$$-\,\dot{S}_1 = \Pi_1/E - C_2K_2. \tag{5.14}$$

The total amount of exhaustible resources extracted over time is limited to the size of the original stockpile, *viz.*,

$$S_1(0) = -\int_0^\infty \dot{S}_1\,dt = \int_0^\infty (\Pi_1/E - C_2K_2)\,dt. \tag{5.15}$$

using equation (5.14).

An assumption adopted in some of the recent energy and economics literature is to treat the resource (essergy) flux R as a state variable (analogous to K) and thus as a factor of production.[6] This is compatible with the observed fact that the aggregate essergy flux is roughly proportional to the output Π of goods and services [equation (5.11)]. For recent empirical evidence in favor of this view, see Cleveland *et al.*, (1984).

However, notwithstanding the fact that essergy is essential for production – a point emphasized by Dasgupta and Heal (1974, 1979) – I believe that to include it as a factor of production on a par with capital and labor would involve some undesirable double counting of factors. Essergy is both an intermediate and a final good. It is embodied to a small extent in materials; but for the most part, intermediate essergy is used to operate capital equipment. To a large extent energy (essergy) is a complement not a substitute for other factors.[7] Hence, to increase the essergy supply without changing capital or labor would have little or no impact on total output. I assume, in effect, that essergy availability is not a limiting factor in the medium term, though it might be a constraint in the very short run (less than 10 years) or the very long run (millions of years).

It is intuitively obvious that investment in capital stock of the second type, K_2, is infeasible until a considerable conventional productive capacity exists. Thus, investments in the earliest period must be either in "ordinary capital" K_1 or in knowledge T, depending upon which is more productive at the time. It is not quite obvious which of these two comes first. (It is a problem not unlike "the chicken or the egg" conundrum) Quite possibly the optimal choice is to invest simultaneously, as will be seen later.

Starting with our already industrialized economy, it is clear that both K_1 and T are essential. There can be no production without a finite stock of both ordinary capital and knowledge. On the other hand, we have not yet begun to invest to a significant degree in building up the stock of extraordinary capital (K_2) needed to provide renewable (non-fossil) essergy in the future.[8]

5.5. Formulation as an Optimal Control Problem

It is appropriate now to introduce a utility function $U(Y)$ in which Y is aggregate consumption and what is consumed is information in some condensed form. This is the point in which many economists may choose to differ with the assumptions in this paper. It is not clear, *a priori*, that such a utility function can be consistently defined. I have already commented briefly on the equivalence of goods to embodied information. Goods, in turn, generate services, which contribute to the maintenance, extension, and enjoyment of life. The purpose of life itself is arguable, but human life – after early infancy – seems to be intimately concerned with awareness or consciousness. Awareness, in turn, is impossible without sensory stimulus and response. The fact that a TV set or book "delivers" information services to consumers is obvious. It is perhaps slightly less obvious that a house or car also delivers services (via the senses) and these services are also equivalent to information. In any case, I assume that services constitute a form of information flux, in the same sense that knowledge is a form of information stock.[9]

Having said this, one can make the usual assumption that $U(Y)$ is strictly concave and twice differentiable. It follows that $U'(Y)$ is a decreasing function of Y. To be consistent, I now define current consumption in terms of production and investment:

$$Y = \Pi_1(K_1, L, T) + I_1 - I_2 - J. \tag{5.16}$$

In any realistic case one can assume that $Y \geq 0$, where $I_1 + I_2 + J < \Pi_1$. It remains to ascertain the optimal path for consumption and the three types of investment.

An optimal consumption–investment policy requires that one maximizes an integral (representing welfare) over time, subject to a number of constraints. The expression to be maximized is

$$W = \int_0^z \exp(-\delta t) U(Y) dt + a_1 K_1(z) + a_2 K_2(z) \tag{5.17}$$

$$+ a_3 T(z) + a_4 S(z),$$

where δ is an assumed intertemporal discount rate or interest rate and z, fixed in advance, is the end of the planning period. In this case z is taken to be very large, but finite. The constants a_1, a_2, a_3, and a_4 are inserted to guarantee that the terminal conditions for an optimal solution will be satisfied. They are chosen to put a prohibitively high penalty on negative values of the state variables at the terminal point. Apart from this, the a_i need not be specified further (see Arrow, 1968).

The integral W in equation (5.17) must be maximized subject to a number of formal restrictions, including the first-order constraints on state variables, *viz.*, (5.1) or (5.4), (5.7), (5.8), (5.9), (5.11), (5.17), and (5.13); plus the nonnegative investment conditions $I_1 \geq 0$, $I_2 \geq 0$, $J \geq 0$; and the nonnegative rate of fossil resource extraction ($S \leq 0$). The latter can be expressed in integral form, as in equation (5.15).

It should be noted that the assumed population growth equation (5.6) is completely independent of the rest of the system and affects the optimal path of consumption only to the extent that the total of available output must be shared among the entire population at any given time. It is also noted that the current resource (essergy) flux is *not* a state variable inasmuch as it is absolutely dependent on the total output of goods and services, which defines the demand for resource inputs. It can therefore be eliminated from the equations.

As already pointed out, I have assumed four kinds of "stocks": productive capital (K_1), energy capital (K_2), knowledge (T), and fossil essergy (S_1). The technical efficiency variable E is defined by (5.1) in terms of knowledge T, and vice versa (5.2). The solutions to the optimization problem are derived in Appendix B. It is interesting that the equations are separable and the shadow price trajectories can be derived explicitly.

5.6. Implications

The implications of the model can best be seen by examining the behavior of the four shadow price variables P_{K_1}, P_{K_2}, P_T, and P_S over time. The important thing to observe is that both P_{K_1} and P_T are initially declining functions of time, while P_S and P_{K_2} are initially increasing. The shadow prices at the starting time $t = 0$ need not be identical, but the optimal investment policy is always to invest in that form of capital whose shadow price times marginal productivity (or price-productivity product) is highest. As a stock increases, its shadow price comes down, and conversely.

Now it is worthwhile to examine the behavior of the four shadow prices: P_S equation (5.41), P_{K_2} equation (5.42), P_{K_1} equation (5.44), and P_T equation (5.46). From the structure of equation (5.41), it can be seen that P_S is a monotonically increasing function (exp δt) with a monotonically *decreasing* function that will eventually become negative (after the end of the planning period Z). The product clearly increases to maximum, followed by a smooth decline.

From equation (5.42), which has the same structure, it is clear that the shadow price P_{K_2} of capital invested in alternative (or renewable) resource production has the same qualitative behavior as P_S. In brief, it rises monotonically to a maximum, then declines smoothly and monotonically toward the end of the planning horizon.

The expression (5.44) for P_{K_1} is more complicated, and it has a different behavior. The first term is a monotonically *decreasing* function whenever the exponent is negative, which is true *whenever the marginal productivity of capital K_1 is large enough for long enough.* During periods of investment in K_1 ($I_1 > 0$), Q_{K_1} must vanish identically and the integral in the second term (in braces) of equation (5.44) is necessarily positive. Thus, during periods of active investment, P_{K_1} is the product of a decreasing exponential function times a term (in brackets) that starts at a constant, rises rapidly at first (because of the integral over P_S, which is initially increasing), but approaches a maximum as the argument of the integral approaches zero. In short, P_{K_1} is, roughly, a declining exponential multiplied by an increasing "S curve." It is complex enough, however, to have "wiggles," corresponding to periods if (or when) the integral over marginal productivity of capital of type K_1 falls below a critical level, such that the exponent shifts from negative to positive.

It can be seen that the structure of equation (5.46) is similar to the structure of equation (5.44), and the behavior of P_T is qualitatively similar to that of P_{K_1}.

Thus, at the beginning of the planning period, two of the shadow prices (P_S, P_{K_2}) are *increasing* and two of them (P_{K_1}, P_T) are *decreasing.*

It is common sense to assume that at the beginning of the period P_{K_1} and P_T are large and P_S, P_{K_2} are zero or negligible. (If this were not the case, there could never be any investment in ordinary productive capital K_1 and/or knowledge t_1 without which there could be no economic output from which savings can be extracted for any subsequent investment in alternative resources.) Given the assumption that P_{K_1} and P_T are initially large but declining, while P_S and P_{K_2} are initially small but increasing, an intersection in trajectories is inevitable.

Whenever two price-productivity products (*PPPs*) intersect, the optimal policy is to shift investment from one to the other form of capital, until the curves cross again, and so on. In principle, such investment switches may occur arbitrarily often. The welfare loss that would result from a compromise policy of investing simultaneously in two (or three) types of capital is therefore negligible. Hence we can safely assume, hereafter, that the two declining *PPPs*, P_{K_1} and P_T, are identical, at least during the early period of unrestrained growth.

It can be seen, now, that the optimal sequence of events, in general terms, consists of four distinct phases:

Phase I $(0 \le t \le t_1)$ is characterized by declining P_{K_1} and P_T and investment, alternately or simultaneously, in two types of productive capital K_1 and T. During this phase either $I_1 > 0$ or $J > 0$, or both. But, during Phase I, P_{K_2} is increasing monotonically and $I_2 = 0$. Time t_1 is defined by the condition $P_{K_1} = P_T = P_{K_2}$. It can be shown without difficulty that this must occur before the final exhaustion of fossil resources (t_s).

Phase II $(t_1 \le t \le t_2)$ is a transitional period, during which investment is exclusively directed at building up the alternative energy capital K_2. Thus $I_2 > 0$ and $I_1 = 0$, $J = 0$. During this phase P_{K_2} continues to increase, but at a decreasing rate, until it reaches a maximum value, before beginning to decrease. Meanwhile, P_{K_1} and P_T also change slope. The two shadow prices P_{K_1} and P_T (and the corresponding P_{K_1} and P_T) do not coincide during this phase because the stock of productive capital K_1 *depreciates*, whereas the stock of technological knowledge T does not. Thus, on physical grounds one would expect P_{K_1} to increase and P_T to remain constant. Time t_2 is determined by the condition $P_{K_2} = P_{K_1}$.

Phase III $(t_2 \le t \le t_3)$ is a second transitional period, during which the optimal investment policy is a combination of K_1 and K_2, either simultaneously or in alternation. This continues until both P_{K_1} and P_{K_2} have declined to the point where they again equal P_T. This defines time t_3.

Phase IV $(t_3 \le t \le z)$ is the final phase during which $P_{K_1} = P_{K_2} = P_T$ all decline more or less simultaneously (i.e., in concert) to zero $(I_1 > 0, I_2 > 0, J > 0)$. It is convenient to equate this point with z, the end of the planning period.

The sequence of phases is shown schematically in *Figure 5.1.* The implications for economic growth are shown in *Figure 5.2.* It is important to observe that during Phase II, while investment is devoted exclusively to the buildup of K_2 (energy capital), the stock of ordinary capital K_1 is actually declining, whence total output of ordinary goods and services Π_1 must also decline. A feasible (but suboptimal) policy is to invest simultaneously in K_1 and K_2 (as in Phase III), so as to just compensate for depreciation of K_1. It might even be feasible to maintain a slow rate of increase in Π_1 by investing simultaneously in all three forms of capital (as in Phase IV). Obviously either policy would stretch out the transition resulting in a somewhat lower level of output in Phase IV and a lower final level (at z).

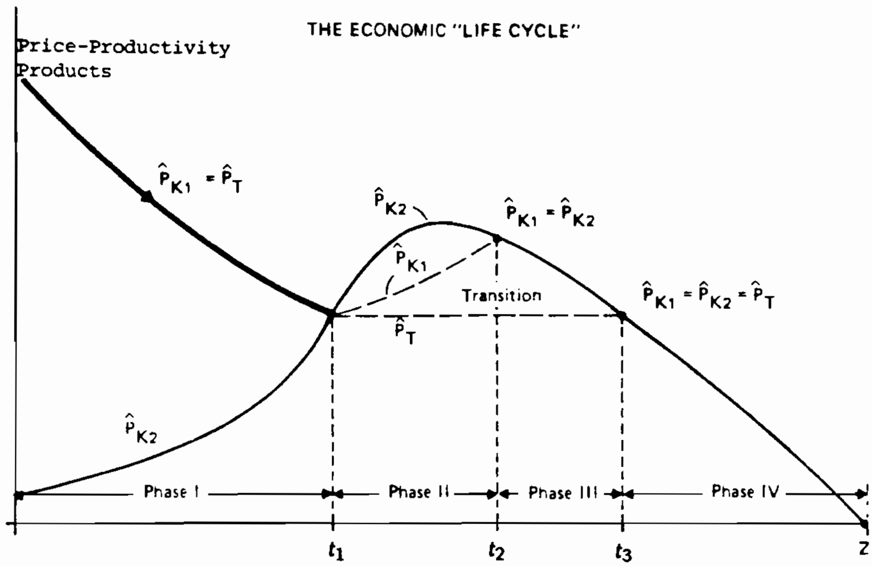


Figure 5.1. The economic "Life Cycle."

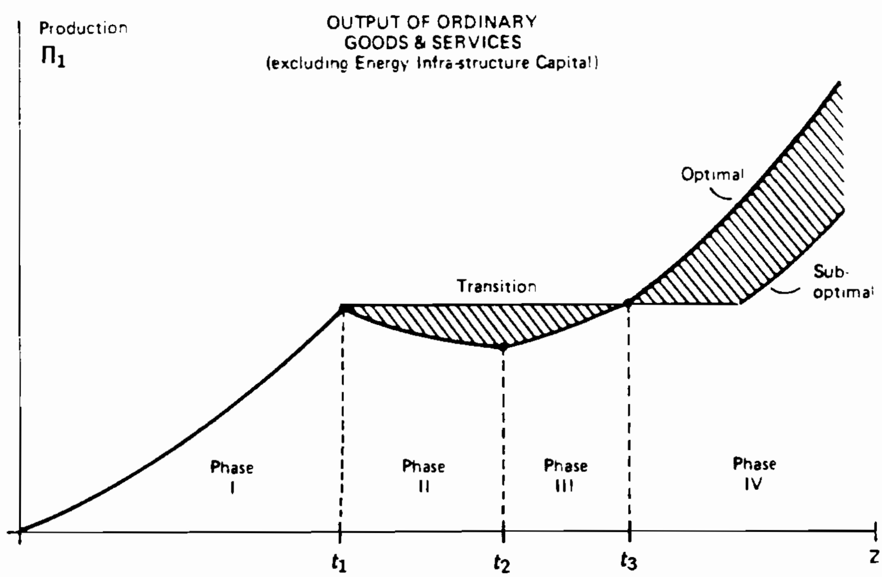


Figure 5.2. Output of ordinary goods and services.

There are two significant implications of this result. First, a long-run optimal policy [given the specification of welfare in equation (5.22)] is *inherently discontinuous*, at least as regards investment policy.[10] It is not optimal to invest in energy capital K_2 at early stages of the economic life cycle while the stock of fossil resources S is still large; and it is not optimal to invest in K_1 during the first part of the transition; finally, it is not optimal to invest in knowledge T after time t_1 until K_1 has been restored to its previous level. It follows from the shifts in optimal investment policy that economic growth will tend to follow an irregular path. In particular, sharp discontinuities in growth *rate* would be experienced, including a change from positive growth rate (slope) to negative growth rate at time t_1.

The reason for the discontinuities on the optimal path has been characterized by Arrow as "myopia." It must be remembered that the control model implicitly postulates an investment decision algorithm based on shadow prices and marginal productivities. In principle, these variables are continuously monitored in real time, and investment for the next period are shifted to whatever form of capital currently corresponds to the largest shadow price-productivity product (PPP).

Of course, growth rate discontinuities in the real world tend to be painful (and a more realistic utility function might attach higher utility to paths exhibiting less discontinuous behavior, *ceteris paribus*). An easier way out of the difficulty (suggested by Arrow) is to postulate a "central planner" with some foresight. The planner would be allowed to smooth over potential discontinuities by starting each investment shift somewhat early and extending it beyond the point of theoretical intersection of shadow prices. There would be a small welfare loss relative to the pure (myopic) optimum, but the planner could try to balance the welfare loss with the pain (loss) due to discontinuities.[11]

Second, the model inherently accounts for (i.e., predicts) structural changes in the economy. In the simple version described above, a *new sector* is created beginning at time t_1. In the generalized version, discussed later, it can be seen that this sector-creation process can be repeated many times. It may be noted that this seems to be a completely new feature of the present growth model. Earlier equilibrium and growth models of Harrod (1936), Domar (1956), or von Neumann (1945) are not compatible with structural change of the kind predicted here.

5.7. A Multiperiod Generalization

On reflection, the rather specialized model analyzed above can probably be generalized quite easily. The key feature of the model, as described, is the exhaustion of a stock of available "fossil" essergy and the buildup of a specialized stock of capital, K_2, whose only function is to permit the economic system to exploit renewable (solar) energy. However, the optimal path for economic growth would be unchanged if K_2 were interpreted, instead, as a stock of "infrastructure" capital required to enable the use of a *different* (less readily available) stock of exhaustible essergy. For analytic convenience, it was assumed that the building of this capital stock requires "ordinary" capital (and labor), but that, once built, each unit of such infrastructure generates a continuous but decreasing flow of essergy throughout its useful life without additional labor. This is a very reasonable description of a solar satellite (as noted earlier) or a hydroelectric plant. It is also a fairly realistic characterization of an oil or gas field, after the drilling is completed and the pipelines are in place.

Given this generalized interpretation, Phase IV of the one-period model would effectively become Phase I of a subsequent cycle. At some time, perhaps after t_3, but certainly before z, the "planner" would have to assess the magnitude of the second kind of resource stock (call it S_2), which need not be accurately known at the time of the initial plan, and identify the next specialized type of infrastructure capital K_3, and its annual essergy yield C_3. A new optimal plan would then be generated for the next period. *Figures 5.3* and *5.4* suggest, in schematic terms, how a multiperiod version of the model can be expected to behave.

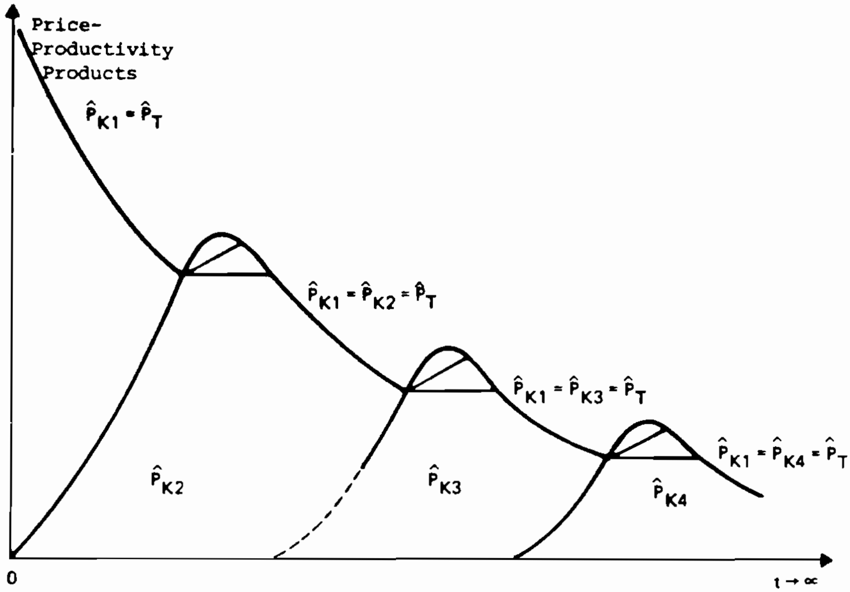


Figure 5.3. Generalized economic cycle.

It is undeniable that *Figure 5.4* bears some resemblance to the so-called Kondratieff long wave. Some economists still doubt that the cycle is "real." However, if the model described in this paper is at all realistic, a wavelike behavior should exist, though the periodicity need not be constant.

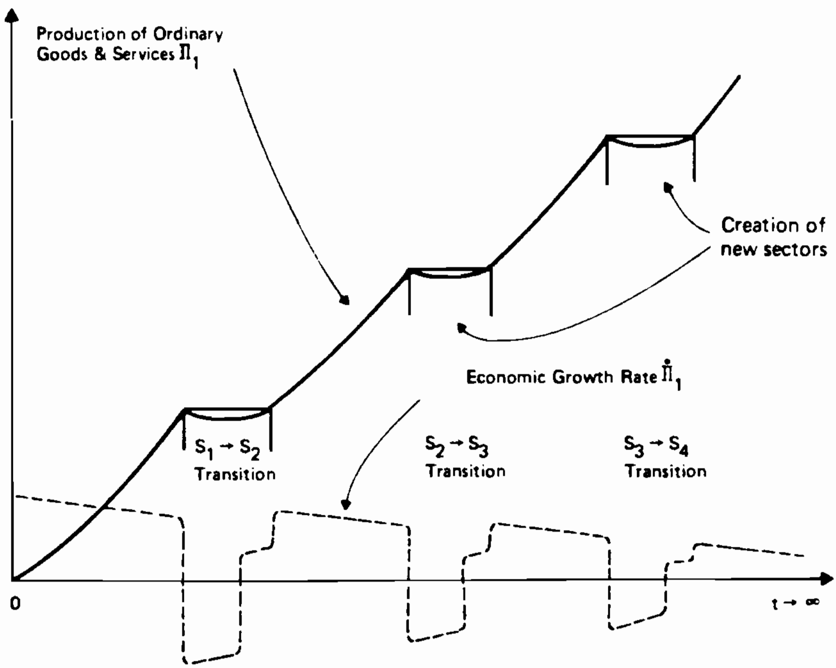


Figure 5.4. Generalized pattern of growth.

It is also undeniable that over the past 200 years of rapid industrialization there have been a series of fairly dramatic shifts in dominant energy (essergy) technology, from wood (charcoal) to coal, then to petroleum and electricity (derived primarily from fossil fuels), and currently to natural gas and/or nuclear power (Nakicenovic and Marchetti, 1979; Nakicenovic, 1986). The sequence of substitutions is shown schematically in *Figure 5.5*

According to the logic of the model, a period of slow growth in ordinary productivity should have occurred during the transition from wood to coal (1780s in the UK, 1880s in the US); again during the transition from coal dominance to oil dominance (the 1930s?); and finally during the transition from oil to gas (the 1980s?). This is a fascinating speculation, to be sure, but too heavy a burden to lay on such a simple model at this stage. Nevertheless, it is interesting to note that the behavior predicted by the model is, at least, qualitatively, consistent with some aspects of historical experience.

5.8. Conclusion

The picture is still too crude to reflect adequately what happens in the real world, of course. One obvious oversimplification is the implicit assumption that each essergy source is homogeneous in grade, with constant capital/output ratio (or yield factor C_2, C_3, . . .) over its lifetime. This is unrealistic, of course, and

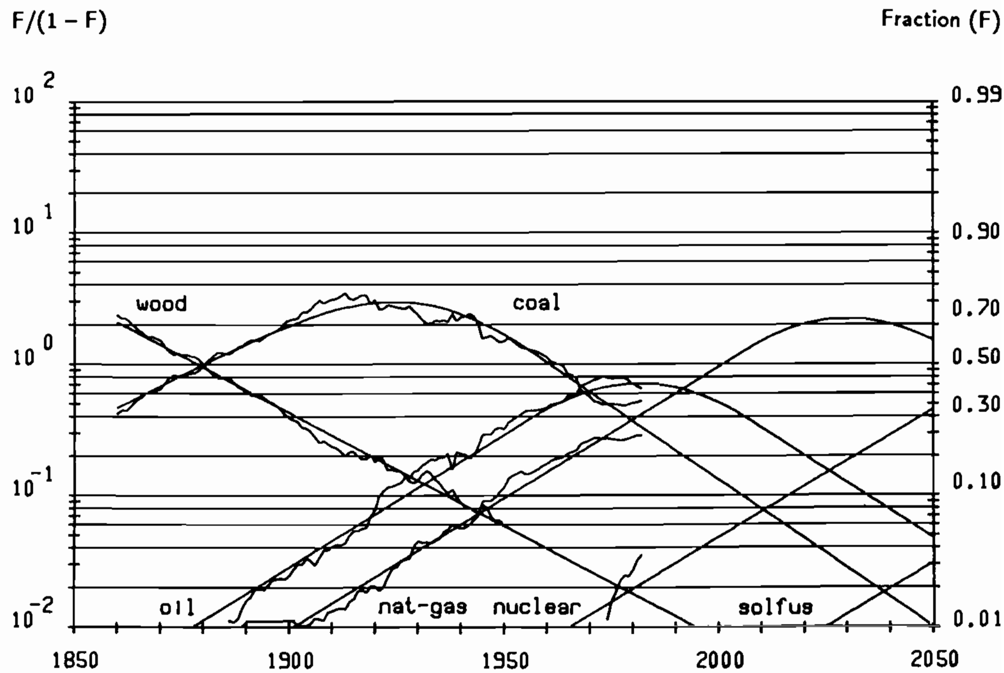


Figure 5.5. World primary energy substitution.

real resources are quite inhomogeneous. Moreover, it has been shown by Herfindahl (1967), among others, that it is optimal to utilize the highest grades of ore first. As a consequence, the quality or grade of the remaining stock of any fossil resource tends to decline over time, which implies that more and more economic effort must be devoted to extraction and refining activities over time. This means that the surplus for consumption or reinvestment lags increasingly over time, in comparison with what it would be in the idealized case illustrated by *Figure 5.4*.

Heterogeneity of actual resource stocks, together with heterogeneity of uses, explains why it can be optimal to exploit more than one different type of resource at the same time, as occurs in the real world.[12] Inhomogeneity and heterogeneity in the system undoubtedly help to smooth out, to some extent, the sharp discontinuities in economic growth rate shown in *Figure 5.4*. However, it is not likely that smoothing due to heterogeneity could totally eliminate the occurrence of changes in growth rate from time to time owing to periodic shifts in investment from one major resource infrastructure to another.

A more penetrating criticism of the present model might be that it is based on the assumption of a single utility function for society as a whole. It is certain that humans and organizations do not in general attempt to specify a utility function for decision making. In fact, most do not utilize any optimizing methodology in decision making. Even if firms or individuals can be assumed to behave like rational "utility maximizers," in the sense of von Neumann and Morgenstern (1944), it is unclear that the combined behavior of many independent individual decision makers would result in overall economic behavior equivalent to that of a single utility-maximizing entity. Thus, the validity of any such model as this is open to question on several points. Nevertheless, the model seems to capture two important but hitherto elusive aspects of macroeconomic behavior. This would appear to justify further investigations, at both the theoretical and empirical levels.

Acknowledgments

Some of the introductory material in this paper is substantially similar to that in an earlier paper by the author (with S. Miller), Ayres and Miller, 1980. The basic model described herein was also, in part, anticipated by that earlier paper. Unfortunately, the original model formulation was flawed, preventing a straightforward solution or interpretation. I am very grateful to Karl-Göran Mäler, Adam Rose, and Erno Zalai for helping me to find and eliminate some mathematical bugs. I am also grateful to Thomas Crocker, Ronald Cummings, Allen Kneese, Wilhelm Krelle, Pieter de Wolff, and an anonymous reviewer for carefully reading the manuscript in an earlier draft and drawing my attention to various deficiencies. None of above are responsible for any remaining errors, and several have even expressed serious reservations about the fundamental approach.

Notes

[1] Some features of the present framework were first suggested in *Resources, Environment & Economics* (Ayres, 1978).

[2] There is an obvious analogy between crude information and crude oil. Refined petroleum products have less energy content but much greater utility than crude oil.

[3] However, see Buringh *et al.*, 1975, for example. Obviously, if humans were able to colonize other planets or grow food in orbiting space colonies, terrestrial limitations would not apply.

[4] Calculating the optimal consumption of such a stock has been called the "cake-eating problem" for obvious reasons.

[5] The subscripts 1 and 2 are used to facilitate a later generalization regarding several kinds of alternative essergy stocks, S_1, S_2, S_3, and types of infrastructure $K_2, K_3, \ldots$.

[6] See, for instance, Hudson and Jorgenson (1974); Allen *et al.* (1976); Manne (1977); Hogan and Manne (1977).

[7] See Berndt and Jorgenson (1973); Berndt and Wood (1977); Griffin and Gregory (1976).

[8] Reasonable people may disagree on whether nuclear fuel is, in any sense, a substitute for fossil essergy or merely another kind of fossil fuel. It does not matter in the context of our model.

[9] I am indebted to T. Vasko for pointing out that the Russian scientist Academician V.A. Trapeznikov made essentially the same argument in 1966.

[10] It must be pointed out that the phase transition is *not* discontinuous in the sense that the consumable resource runs out suddenly. Nor is there any discontinuity in resource price at the point where the initial substitution of the alternative of essergy resource (renewable or not) for the depletable resource. This transition begins at a point where the shadow prices are equal. An extended discussion of this issue can be found in Tietenberg (1984, Chapter 6). I am indebted to a reviewer for calling my attention to this point.

[11] Probably the planner would use an optimal control model in a simulation mode.

[12] The "grade" of a resource can only be defined in relation to a specific use. Thus coal is a very low-grade resource in terms of providing liquid fuels for automotive vehicles or aircraft, and it is scarcely better in terms of providing gas for household heating. On the other hand, coking coal is a very high-grade resource for purposes of smelting iron ore.

Appendix A: Information and Knowledge

In fact, one can clearly identify and distinguish at least three distinct forms of information. There is an obvious analogy with the distinct forms of matter (solid, liquid, and gas), although I do not pursue it further here.

(1) *Disembodied* information is associated with the temperature on spectral characteristics of incoherent electromagnetic or thermal radiation (energy). It is quantitatively proportional to the available useful work (or essergy) content of the energy flux.

(2) Information is *embodied* in the (average) state and chemical composition of unstructured matter, whether gaseous, liquid, or solid, or in the physical microstructure of a crystal or glass.

(3) Information is embodied in the form or shape of a solid medium, (two or three dimensions) or in the structure of a macromolecule (such as DNA).

The first two categories are essentially thermodynamic. Explicit rules for computing each type of information content in quantitative terms have been formulated. Note that the third category includes information as we normally use the term, *viz.*, a photograph, symbols on a printed sheet of paper, a magnetized tape, a precision gear, or a pattern of impurities in a silicon chip. Information of the third kind can be (and often is) transmitted via telecommunications channels, converted from one form (e.g., analog data) to another equivalent form (e.g., digital data) and "processed" by computers.

Note that the third kind of information can only be stored and processed (i.e., utilized) by living organisms and/or material devices that also embody information of the second kind. Moreover, all such organisms and devices require a flux of available useful work (essergy) for their maintenance. Thus, information of the third kind is, in some sense, the *essence* or *condensate* of a much larger quantity of information of the first and second kinds. It can be termed "morphological."

Knowledge can perhaps be thought of as a fourth kind of information or as the "useful" component of information of the third kind. It has been suggested that knowledge is the minimum information required to decode a message or to reproduce forms or patterns. If this is true, knowledge is a form of information embodied in a decoder or copying machine, or possibly in a living reproductive cell or a brain. Knowledge is therefore literally undefinable in the absence of a supporting material system. The more knowledge is embodied in the decoder, the less information needs to be transmitted to reproduce the original message, or object.

There is no general means of computing the minimum information requirement to reproduce an object except for objects themselves defined in terms of computer languages. In this context, it is noteworthy that there is a computer science literature on algorithmic information theory, for example, Chaitin (1978). Although quantitative formulas are lacking in general, it is safe to assume that the knowledge component of stored or transmitted information of the third kind is normally quite small, compared with the total amount of information of all kinds that must be mobilized to store or transmit it. In other words, much form and structure information is actually redundant. It follows, incidentally, that while the amount of thermodynamic information (of the second kind) that can be extracted each year from all sources (fossil fuels plus solar flux) is indeed limited, this in itself imposes no practical limitation on the rate of accumulation of human knowledge relevant to the production of goods or services.

Appendix B: Solution to the Optimization Problem

To solve the optimization problem stated above (following Takayama, 1974) we define a Hamiltonian system with three "controls" I_1, I_2, and J:

$$\begin{aligned} H = {} & \lambda_0[U(\Pi_1 - I_1 - I_2 - J) + a_1\dot{K}_1 + a_2\dot{K}_2 + a_3\dot{T} + a_4\dot{S}_1] \\ & + \hat{P}_{K_1}(I_1 - d_1K_1) + \hat{P}_{K_2}(I_2 - d_2K_2) + \hat{P}_TJ \\ & - \hat{P}_S(\frac{\Pi_1}{E} - C_2K_2) + Q_{K_1}I_1 + Q_K\, I_2 + Q_TJ + Q_SS_1. \end{aligned} \tag{5.18}$$

It can be shown without difficulty that λ_0 can be set equal to unity without loss of generality. Moreover, the three terminal conditions are automatically satisfied by defining $P_{K_1} = \hat{P}_{K_1} + a_1$, $P_{K_2} = \hat{P}_{K_2} + a_2$, $P_T = \hat{P}_T + a_3$, and $P_S = \hat{P}_S - a_4$. This yields the simpler equivalent Hamiltonian:

$$\begin{aligned} H = {} & U(\Pi_1 - I_1 - I_2 - J) + P_{K_1}(I_1 - d_1K_1) \\ & + P_{K_2}(I_2 - d_2K_2) + P_TJ - P_S(\frac{\Pi_1}{E} - C_2K_2) \\ & + Q_{K_1}I_1 + Q_{K_2}I_2 + Q_TJ + Q_SS_1. \end{aligned} \tag{5.19}$$

The co-state variables P_{K_1}, P_{K_2}, P_T, and P_S are canonical conjugates of K_1, K_2, T, and S_1, respectively. They are usually interpreted as shadow prices of the corresponding stocks K_1, K_2, S_1, and T. The Lagrange multipliers Q_K, Q_T, and Q_S are zero or positive, but the products Q_{K_1}, I_1, Q_{K_2}, I_2, Q_TJ, and Q_SS_1 are all identically zero. Thus, introducing the nonnegativity constraints:

$Q_{K_1} = 0$ whenever $I_1 > 0$; otherwise $Q_{K_2} \geq 0$
$Q_{K_2} = 0$ whenever $I_2 > 0$; otherwise $Q_{K_2} \geq 0$
$Q_T = 0$ whenever $J > 0$; otherwise $Q_T \geq 0$
$Q_S = 0$ whenever $S_1 > 0$; otherwise $Q_S \geq 0$.

Two other nonnegativity constraints could be included, *viz.*,– $\dot{S} > 0$ (resources are never put back into the ground) and $I_1 + I_2 + J < \Pi_1$ (investment never exceeds current production). However, the constrained and unconstrained solutions are essentially identical.

The first three Euler–Lagrange equations for an optimal path are obtained by partially differentiating the Hamiltonian (5.19) with respect to I_1, I_2, and J, respectively:

$$\frac{\partial H}{\partial I_1} = 0 = -U'(Y) + P_{K_1} + Q_{K_1} \tag{5.20}$$

$$\frac{\partial H}{\partial I_2} = 0 = -U'(Y) + P_{K_2} + Q_{K_2} \tag{5.21}$$

$$\frac{\partial H}{\partial J} = 0 = -U'(Y) + P_T + Q_T. \tag{5.22}$$

It follows from equations (5.20), (5.21), and (5.22) that

$$P_{K_1} + Q_{K_1} = P_{K_2} + Q_{K_2} = P_T + Q_T = U'(Y), \tag{5.23}$$

where $U'(Y)$ is the maginal utility of aggregate consumption. It follows from (5.23) and the nonnegativity conditions that the optimal investment rule is to invest in that type of capital with the largest shadow price. This can be demonstrated by assuming the contrary. For instance, let $P_{K_1} < P_{K_2}$ but assume simultaneous investment in both; i.e., $I_1 > 0$ and $I_2 > 0$ at the same time. Then from the nonnegativity conditions $Q_{K_1} = Q_{K_2} = 0$. From (5.23) it would follow that $P_{K_1} = U'(Y)$ and that $P_{K_2} = U'(Y)$. But this is not possible, by assumption that $P_{K_1} < P_{K_2}$.

Taking this line of reasoning further, one can now derive the following expressions for the Qs:

$$Q_{K_1} = \begin{cases} 0 & \text{if } P_{K_1} > P_{K_2}, P_T \\ P_{K_2} - P_{K_1} & \text{if } P_{K_2} > P_{K_1}, P_T \\ P_T - P_{K_1} & \text{if } P_T > P_{K_1}, P_{K_2} \end{cases} \tag{5.24}$$

$$Q_{K_2} = \begin{cases} 0 & \text{if } P_{K_2} > P_{K_1}, P_T \\ P_{K_1} - P_{K_2} & \text{if } P_{K_1} > P_{K_2}, P_T \\ P_T - P_{K_2} & \text{if } P_T > P_{K_1}, P_{K_2} \end{cases} \tag{5.25}$$

$$Q_T = \begin{cases} 0 & \text{if } P_T > P_{K_1}, P_{K_2} \\ P_{K_1} - P_T & \text{if } P_{K_1} > P_{K_2}, P_T \\ P_{K_2} - P_T & \text{if } P_{K_2} > P_{K_1}, P_T. \end{cases} \tag{5.26}$$

The co-state variables P_{K_1}, P_{K_2}, P_T, and P_S, together with the corresponding state variables K_1, K_2, T, and S, satisfy the following canonical system of differential equations, which are conditions for a solution:

$$\frac{\partial H}{\partial P_{K_1}} = \dot{K}_1 \tag{5.27}$$

$$\frac{\partial H}{\partial P_{K_2}} = \dot{K}_2 \tag{5.28}$$

$$\frac{\partial H}{\partial P_T} = \dot{T} \tag{5.29}$$

$$\frac{\partial H}{\partial P_S} = \dot{S} \tag{5.30}$$

$$\frac{\partial H}{\partial K_1} = -\left(\dot{P}_{K_1} - \delta P_{K_1}\right) \tag{5.31}$$

$$\frac{\partial H}{\partial K_2} = -\left(\dot{P}_{K_2} - \delta P_{K_2}\right) \tag{5.32}$$

$$\frac{\partial H}{\partial T} = -\left(\dot{P}_T - \delta P_T\right) \tag{5.33}$$

$$\frac{\partial H}{\partial S} = -\left(\dot{P}_S - \delta P_S\right). \tag{5.34}$$

To solve the set of eight differential equations (5.27)–(5.34) we need eight constants of integration. These are determined by so-called transversality conditions. For the four state variables K_1, K_2, T, and S, it is reasonable and sufficient to fix initial values at time $t = 0$. The initial values can be zero or finite. Except for P_S, the corresponding co-state variables must be fixed at the terminal point $t = \mathbf{z}$. Here it is reasonable (though not necessary) to assume

$$P_{K_1}(\mathbf{z}) = P_{K_2}(\mathbf{z}) = P_T(\mathbf{z}) = 0. \tag{5.35}$$

However, if we specify $S(\mathbf{z}) = 0$, then $P_S(\mathbf{z}) > 0$ and conversely. The case of a declining, but still nonzero, resource stock would imply $P_S(\mathbf{z}) > 0$, with the useful simplification that $Q_S = 0$ for all $t \leq t_s$. The following derivation is based on the simplification that $\mathbf{z} = t_s$, whence $Q_S = 0$ at all times $t \leq \mathbf{z}$. In the more general case where $t_s < \mathbf{z}$, it can be shown that P_S is discontinuous at t_s and declines for $t_s < t < \mathbf{z}$. In this period Q_s is nonzero and must be determined by using the condition $\dot{S} = 0$, which implies [from (5.14)] that $C_2 K_2 = \Pi_1 / E$.

It is interesting to note, by the way, that the condition of vanishing shadow prices (5.35) implies that the marginal utility of consumption $U'(Y)$ also declines to zero [by (5.23)], which means that a consumption plateau is finally reached at $t = s$. When the indicated differentials of the Hamiltonian are carried out, the results are a set of four first-order differential equations for the shadow prices P_{K_1}, P_{K_2}, P_T, and P_S as follows:

$$0 = \dot{P}_{K_1} - (\delta + d_1)P_{K_1} + U'(Y)\left(\frac{\partial \Pi_1}{\partial K_1}\right) - P_S\left(\frac{1}{E}\frac{\partial \Pi_1}{\partial K_1}\right) \tag{5.36}$$

$$0 = \dot{P}_{K_2} - (\delta + d_2)P_{K_2} + C_2 P_S \tag{5.37}$$

$$0 = \dot{P}_T - \delta P_T + U'(Y)\left(\frac{\partial \Pi_1}{\partial T}\right) - P_S\frac{\partial}{\partial T}\left(\frac{\Pi_1}{E}\right) \tag{5.38}$$

$$0 = \dot{P}_S - \delta P_S + Q_S. \tag{5.39}$$

It should be pointed out that (5.38) assumes that output Π_1 is explicitely dependent on T, but that there is no implicit dependence through K_1 or K_2. In other words,

$$\partial K_1/\partial T = \partial K_2/\partial T = 0. \tag{5.40}$$

This reflects the fact that K_1 and K_2 are pure measures of the quantity of capital. Improvements in the quality of capital and labor are reflected by increases in T alone.

The most general solution of (5.39) is

$$P_S(t) = e^{\delta t}\left[P_S(0) - \int_{ts}^{t} Q_S(t')dt'\right] = P_S(0)e^{\delta t}, \tag{5.41}$$

since $Q_S = 0$ holds for all $t < t_s$. The next step is to substitute equation (5.41) for (5.37) and solve. The result is

$$P_{K_2} = \exp(\delta + d_2)t\left[P_K(0) - C_2\int_0^t P_S(t')\exp(d_2 t')dt'\right], \tag{5.42}$$

which rises to a maximum value (when $\dot{P}_{K_2} = 0$) and then falls, becoming negative when the term in square brackets becomes negative. The initial value of $P_K(0)$ must be chosen large enough such that P_{K_2} vanishes at $t = s$, as required by (5.35).

To solve (5.36) and (5.38) we can again substitute (5.41) and also use (5.23). Equation (5.36) becomes

$$0 = \dot{P}_{K_1} - \left(\delta + d_1 - \frac{\partial \Pi_1}{\partial K_1}\right)P_{K_1} - \left(\frac{\partial \Pi_1}{\partial K_1}\right)\left(\frac{P_S}{E} - Q_{K_1}\right) \tag{5.43}$$

with the general solution

$$P_{K_1} = \exp\left[\int_0^t (\delta + d_1 - \frac{\partial \Pi_1}{\partial K_1}) dt'\right]$$
$$\times \left\{P_{K_1}(0) + \int_0^t (\frac{P_S}{E} - Q_{K_1})\frac{\partial \Pi_1}{\partial K_1} \exp\left[-\int_0^{t'} (\delta + d_1 - \frac{\partial \Pi_1}{\partial K_1})\, dt''\right] dt'\right\}, \tag{5.44}$$

where Q_{K_1} is given by (5.24). Similarly (5.38) becomes

$$0 = \dot{P}_T - \left[\delta - \frac{\partial \Pi_1}{\partial T}\right] P_T - \frac{\partial}{\partial T}(\frac{\Pi_1}{E})\ P_S + (\frac{\partial \Pi_1}{\partial T})\ Q_T \tag{5.45}$$

and has the solution

$$P_T = \exp\left[\int_0^t (\delta - \frac{\partial \Pi_1}{\partial T}) dt'\right]$$
$$\times \left\{P_T(0) + \int_0^t \left[P_S \frac{\partial}{\partial T}\left[\frac{\Pi_1}{E}\right] - Q_T \frac{\partial \Pi_1}{\partial T}\right] \exp\left[\int_0^{t'} (\delta - \frac{\partial \Pi_1}{\partial T})\, dt''\right] dt'\right\} \tag{5.46}$$

and Q_T is given by (5.26). When (5.24) is inserted to (5.44) to eliminate Q_{K_1}, the result is an integral equation. The same is true when (5.26) is substituted into (5.46). It can be verified without much difficulty that (5.44) and (5.46) are well behaved for reasonable values of parameters. In particular, the shadow price of productive capital, P_{K_1}, is a (generally) decreasing function of time – as it should be – as long as the marginal productivity of capital $\partial \Pi_1/\partial K_1$ is greater than the sum of pure utility discount rate (if any) plus the depreciation rate; i.e.,

$$\frac{\partial \Pi_1}{\partial K_1} > \delta + d_1.$$

Similarly, the shadow price of technological knowledge P_T is a generally decreasing function of time, provided the marginal productivity of knowledge exceeds the discount rate

$$\frac{\partial \Pi_1}{\partial T} > \delta.$$

Note that Q_{K_1} and Q_T are, respectively, nonzero when, and only when, the corresponding investment terms (I_1 and J) vanish. The effect of nonzero values of Q_{K_1} and Q_T is to decrease (or even reverse) the decline P_{K_1} and P_T, again as one would expect. In fact, Q_{K_1} and Q_T, appearing in equations (5.44) and (5.46), act as negative feedback stabilizers, in effect. They vanish at points where the shadow price trajectories intersect and increase as they diverge.

Actual solutions of equations (5.44) and (5.46) require forward integration with assumed starting values of P_{K_1}, P_T to $t = z$, followed by a set of successive corrections until the terminal conditions are satisfied.

References

Allen, E.L., *et al.*, 1976, *U.S. Energy And Economic Growth, 1975–2010*, Publication ORAU/IEA-76-7, Institute for Energy Analysis, Oak Ridge Associated Universities, Oak Ridge, TN, USA.

Arrow, K.J., 1968, Applications of Control Theory to Economic Growth, in G. Dantzig and A. Veinott, eds., *Mathematics Of The Decision Sciences*, American Mathematical Society, Providence, RI, USA.

Ayres, R.U., 1978, *Resources, Environment and Economics: Applications Of The Materials/Energy Balance Principle*, Wiley, New York, NY, USA.

Ayres, R.U. and Miller, S.M., 1980, The Role of Technological Change and Resource Constaints on an Optimal Growth Path, *Journal of Environmental Economics and Management* 7: 353–371.

Berndt, E.R. and Jorgenson, D.W., 1973, Production Structure, in *US Energy Resources and Economic Growth*, Ford Foundation Energy Policy Project, New York, NY, USA.

Berndt, E.R. and Wood, D.O., 1977, *Engineering and Econometric Approaches to Industrial Energy Conservation and Capital Formation: A Reconciliation*, WP MIT-EL-77-040WP, Energy Laboratory, Massachusetts Institute of Technolgy, Cambridge, MA, USA.

Brillouin, L., 1953, Negentropy Principle of Information, *Journal of Applied Physics* **24**(9):1152–1163.

Buringh, P., van Heemst, H., and Staring, G., 1975, *Computation of the Absolute Maximum Food Production of the World*, Agricultural University, Wageningen, Netherlands.

Chaitin, G., 1978, Toward a Mathematical Definition of Life, in R. Levine and M. Tribus, eds., *The Maximum Entropy Formalism*, MIT Press, Cambridge, MA, USA.

Cleveland, C.J., Costanza, R., Hall, C.A.S., and Kaufmann, R., 1984, Energy and the US Economy: A Biophysical Perspective, *Science* **225**(Aug. 31):890–897.

Dasgupta, P. and Heal, G., 1974, The Optimal Depletion of Exhaustible Resources, *Review of Economic Studies*.

Dasgupta, P. and Heal, G., 1979, *Economic Theory and Exhaustible Resources*, Series: Cambridge Economic Handbooks, Cambridge University Press, Cambridge, UK.

Dasgupta, P., Heal, G., and Majundar, M., 1977, Resource Depletion and Development, in M. Intriligator, ed., *Frontiers of Quantitative Economics*, North Holland, Amsterdam, Netherlands.

Denison, E.F., 1967, *Why Growth Rates Differ*, Brookings Institution, Washington, DC, USA.

Domar, E.D., 1956, Capital Expansion, Rate of Growth and Employment, *Econometrica* **14**:137–147.

Gray, L.C., 1914, Rent under the Assumption of Exhaustibility, *Quarterly Journal of Economics* **28**:466–489.

Griffin, J.M. and Gregory, P.R., 1976, An Intercountry Translog Model of Energy Substitution Responses, *American Economic Review* **66**(December):845–857.

Harrod, R.F., 1936, *The Trade Cycle*, Oxford University Press, Oxford, UK.

Hartley, R.V.L., 1928, Transmission of Information, *Bell System Technical Journal* **7**.

Herfindahl, O., 1967, Depletion and Economic Theory, in M. Goffney, ed., *Extractive Resources and Taxation*, University of Wisconsin Press, Madison, WI, USA.

Hogan, W.W. and Manne, A.S., 1977, Energy-Economy Interactions: The Fable of the Elephant and the Rabbit, in C.J. Hitch, ed., *Modeling Energy-Economy Interactions: Five Approaches*, Series: Research Paper R5, Resources for the Future, Washington, DC, USA.

Hotelling, H., 1931, The Economics of Exhaustible Resource, *Journal of Political Economy* **39**:137–175.

Hudson, E. and Jorgenson, D., 1974, US Energy Policy and Economic Growth, 1975–2000, *Bell Journal of Economics & Management Science* **5**:461–514.

Huettner, D.A., 1981, Energy, Entropy and Economic Analysis: Some New Directions, *Energy Journal* **2**(2):123–130.

Kamien, M. and Schwartz, N., 1978, Optimal Exhaustible Resource Depletion with Endogenous Technical Change *Review of Economic Studies* **45**.

Kamien, M. and Schwartz, N., 1982, *Market Structure And Innovation*, Series: Cambridge Surveys of Economic Literature, Cambridge University Press, Cambridge, UK.

Lotka, A.J., 1956, *Elements of Mathematical Biology*, Dover Publications, (Original title: *Elements of Physical Biology*, 1924).

Manne, A.S., 1977, ETA-Macro, in C.J. Hitch, ed., *Modeling Energy-Economy Interactions: Five Approaches*, Series: Research Paper R5, Resources for the Future, Washington, DC, USA.

Nakicenovic, N. and Marchetti, C., 1979, *The Dynamics of Energy Systems and the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Nakicenovic, N. 1986, *Patterns of Change: Technological Substitution and Long Waves in the United States*, WP-86-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

von Neumann, J., 1945, A Model of General Economic Equilibrium, *Review of Economic Studies* **13**:1–9.

von Neumann, J. and Morgenstern, O., 1944, *Theory Of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA.

Nordhaus, W.P., 1973, The Allocation of Energy Resources, *Brookings Papers on Economic Activity*, Vol. 3, Brookings Institution, Washington, DC, USA.

Pearl, R., 1922, The Population Problem, *Geographical Review* **638** (October).

Shannon, C.E., 1948, A Mathematical Theory of Communication, *Bell System Technical Journal* **27**.

Shannon, C.E., 1951, Prediction and Entropy of Printed English, *Bell System Technical Journal* **30**:50–64.

Stiglitz, J., 1974, Growth with Exhaustible Natural Resources: Efficient and Optimal Growth Paths, *Review of Economic Studies*.

Stiglitz, J., 1979, A Neoclassical Analysis of the Economics of Natural Resources, in V.K. Smith, ed., *Scarcity and Growth Reconsidered*, Resources for the Future, Washington, DC, USA.

Takayama, A., 1974, *Mathematical Economics*, Dryden Press, Hinsdale, IL, USA.
Tietenberg, T., 1984, *Environmental and Natural Resource Economics*, Scott, Foresman and Company, Glenview, IL, USA.
Tribus, M. and McIrvine, E.C., 1971, Energy and Information, *Scientific American* **225**(3):179–188 (September).

CHAPTER 6

From Life Cycles to Long Waves to Catastrophes

S. Menshikov and L. Klimenko

6.1. Introduction

In recent years substantial evidence has been accumulated to show that growth in many products and technologies passes through a life cycle, experiencing, in consecutive order, takeoff, acceleration, deceleration, and stagnation. In statistical and econometric terms such processes can be approximated by the logistic curve.

The purpose of this paper is to show that the life-cycle approach has a more general application in economic analysis, helping to explain not only individual instances of growth in products or technologies, but also overall macroeconomic trends, fluctuations, and sudden drastic changes in performance.

6.2. The Life Cycle as a General Rule

In economic life any product or factor that is not immediately consumed, but utilized gradually, due to either its material substance or its economic necessity, undergoes some sort of life cycle. Among these factors are natural resources, labor, technical know-how, and those parts of national product that are saved, accumulated, and subsequently used in production or consumption. In this

Modified version of this paper was published in *Collection of Works of All Union System Research Institute*, No. 20, 1988, pp. 54–69, Moscow.

paper we shall leave aside natural resources, labor, and demographic variables, and deal exclusively with renewable material resources, which play an important role in determining life cycles of products, technologies, plant, equipment, and inventories.

All these factors operate as capital stock participating in various production processes, and are themselves created by accumulating the results of such processes. This applies also to stocks of materialized technical know-how, consisting of plant and equipment specifically designed to produce new products and new technologies. The nature and role of the economic life cycle is explained by the interaction between stocks and processes.

The economic duration of a particular life cycle is determined largely by the time that is necessary to return the full value of the stock plus an increment intended to maintain a given rate of growth. The maximum duration is, of course, set by the physical durability of the particular kind of stock. However, the resource usually becomes economically obsolete long before it ceases to function in physical terms.

The return of the stock per unit of time varies at different stages of its operation. The overall rule is that it gradually reaches a maximum somewhere near the middle of its lifetime ("maturity"), and then tends to fall, so that the overall picture is reminiscent of the normal probability distribution. Smaller returns at the start are caused by natural limitations of any new resource, which need time and expense to be overcome, while at the finish smaller returns are due to aging and to competition of newer products and technologies.

This process can be described in terms of function

$$E(t) = \frac{are^{-rt}}{(1+e^{-rt})^2}, \tag{6.1}$$

where $E(t)$ is return per unit of time and r is speed of adjustment to optimal conditions. Starting at 0, $E(t)$ gradually accelerates and then levels off reaching a maximum ($ME = ar/2$) and then declines back to zero (*Figure 6.1*).

A higher speed of adjustment can in some cases be achieved by using newly produced equipment of a well-established type so as to avoid high costs of initial introduction. It is possible in this way to gain higher initial returns, but eventually the user may find himself lagging behind in technical progress.

Given the described behavior of return per unit of time, total accumulated return will have a maximum limit a. At any given moment it will be approximated by the logistic curve

$$F(t) = \frac{a}{1+e^{-rt}}, \tag{6.2}$$

where $F(t)$ is total accumulated return of the stock and a is total demand over time for product of the stock or as the maximum limit of total return.

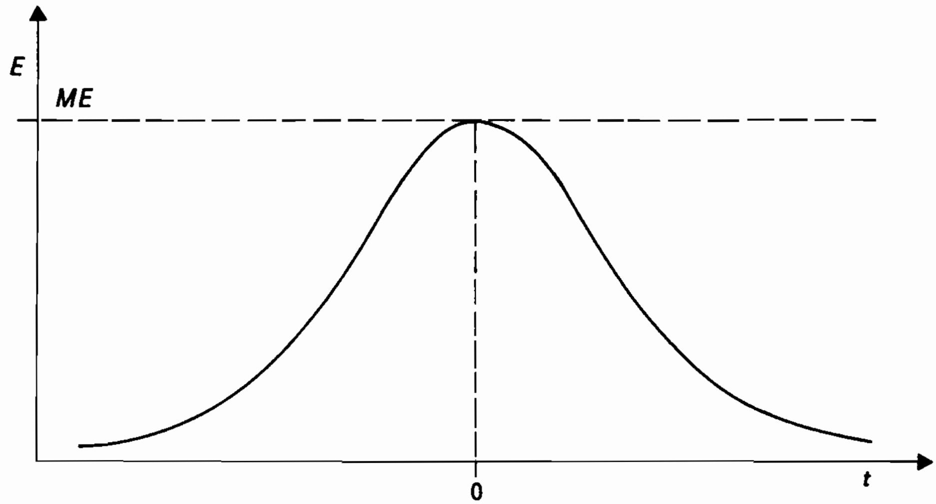


Figure 6.1. Wave of return of stock function.

The duration of a life cycle in terms of this formula is directly correlated with total demand, and inversely correlated with the speed of adjustment (both return and demand are measured in the same units). Total demand depends on a number of factors, including aggregate purchasing power and relative competitiveness of the product or process.

Let us assume that from the macroeconomic perspective r represents the average speed of the utilization of stocks in the economy, or the speed of diffusion of new products and new technologies. It follows that the ratio r/a is extremely important in determining whether an economy is keeping abreast of technical progress, or is lagging behind other economies.

Any renewable material stock is created by capital investment seeking adequate return. Assume that expected accumulated return equals

$$EF(t) = K(O)\ (1+q)^t, \tag{6.3}$$

where $EF(t)$ is expected accumulated return at moment t, $K(O)$ is initial value of the capital stock, and q is expected average annual growth rate of K equal to average rate of return.

Compare expected and actual total return (*Figure 6.2*). There are a few possibilities. Assuming that $EF(t) = F(t)$, both curves intercept at t a point that is close to total demand a. The economically determined lifetime of the stock will terminate at this point, since its further use would fall below the expected rate of return. By this time the total actual return would be adequate to replace the old stock by a new one which could be larger and, presumably, more efficient.

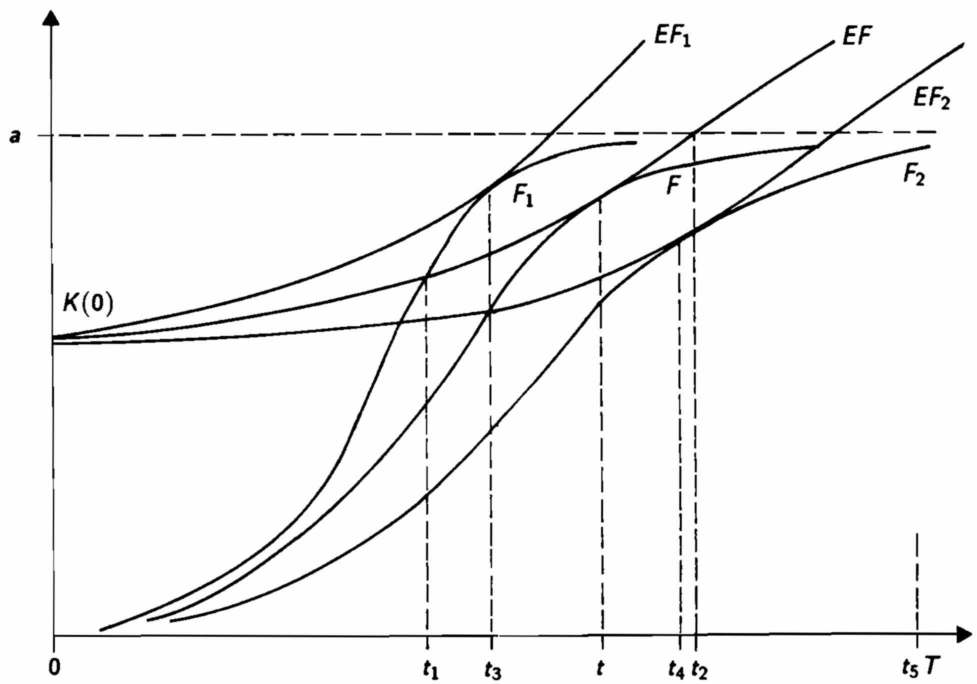


Figure 6.2. Comparison of expected and actual total return.

If the speed of utilization (diffusion), r, is sufficiently high, actual total return may intercept EF at an earlier point in time (t_1). The replacement of old stock would then be possible earlier than expected, or old stock could be kept until any moment before t_2, since this would bring additional total return. However, if the rate of profit, rather than total return were being maximized, the lifetime of the stock would terminate at t_3. The outcome would depend on the actual criterion, which would be chosen by entrepreneurs or planners at this time.

This duality of outcome is important. It indicates that higher speeds of utilization (diffusion) may change the duration of the life cycle in both directions. Whether it becomes shorter or longer, depends on the economic rationale that prevails at this point. As shown below, this may be tantamount to the existence of more than one equilibrium position, and thus to the possibility of sudden leaps from one to another (bifurcation).

It is also possible that r turns out to be rather slow, so that total actual return fails to intercept the expected return curve. This again may lead to several eventualities. One would be to take a loss at expected time t, and replace the old stock with a more efficient one. Another would be to wait until maximum average return (or actual profit rate) was achieved at t_4. In both cases the effective rate of return would be lower than expected, but the duration of the life cycle would not change appreciably. However, if profit maximization was replaced by output maximization, then the life cycle could be lengthened appreciably from t to t_5. Again, this opens the door to bifurcation.

The concept, as described above, applies directly to life cycles in products and technologies, though it is true that the notion of "stock," as related to particular products or technologies, is not easy to define in precise terms. Output of a family of products or technologies needs so many different kinds of capital stock with so many different durabilities that special research would be needed to single out the particularly long-lasting components of plant, equipment, and technology, which are associated with the particular family of products or technologies constituting one common life cycle.

When the life-cycle concept is applied to equipment that is later replaced by modernized versions of the same kind, one may argue that a represents not total demand, but rather potential accumulated total return. The latter is reached when incremental return becomes zero or negative. This is possible if equipment is used beyond its normal retirement age owing to expensive capital repairs, patching up, and other costly methods. This may be practical in extreme deficit conditions, but normally could result in technical regress, inefficiency, and physical breakdowns.

At the beginning of this section inventories were mentioned as an example of stocks with a life cycle. This too needs explanation. Capital invested in inventories is similar to any other in that it expects a certain rate of return. But does actual return follow a logistic curve? It is obvious that maximum momentary return from inventories is reached only at a certain optimal ratio between inventories and output. Investment into inventories is made in order to achieve savings in transportation and other costs. It follows that the maximum momentary return is reached only somewhere in the middle of its lifetime, since inventories are either higher (at the beginning) or lower (at the end) than their optimal level.

The length of the life cycle in inventories depends in part on the physical durability of the product that is being saved. But apart from this consideration it is determined by the ratio between intermediate costs and value added. Klimenko (1972) has shown that the average duration of the inventory cycle is practically equal to the ratio of inventories to value added in various industries. The larger the share of intermediate costs in total value, the higher, on average, is the optimal ratio of inventories to product. And the lower the share of value added in total value, the smaller the average relative return obtained from the use of inventories and the longer the life cycle in inventories.

6.3. The Life Cycle and the Long Wave

The particular combination of life cycles in an economy serves as a material basis for the existence of a spectrum of fluctuations with different frequencies. The simplest explanation of how one translates into another is that the intermittency in stock creation leads to an echo effect.

However, if one assumes an ideally planned economic system, the time when stocks are created can be distributed so evenly that the echo effect would practically disappear. But even such an ideal system is not free from external shocks, which may impose an echo, depending on its speed of adjustment and correction. As experience of the 1970s and 1980s has shown, the adjustment of centrally planned economies to drastic increases and falls in the price of oil and other phenomena of the structural crisis in market economies has been relatively slow, which served to increase foreign indebtedness and internal economic imbalances.

In a predominantly market economy the echo effect may be either dampened or sustained owing to the existence of other sources that generate economic cycles of different duration. Intuitively, it seems logical to explain such cycles by predominant average duration of life cycles of different kinds of stocks.

The mechanism that translates life cycles into macroeconomic fluctuations has yet to be fully explored. Two approaches are feasible. One approach is to study the speeds of reaction and adjustment in an economic system that transforms echo effects and life cycles into close to periodic fluctuations. This approach is reviewed in more detail later in this paper. Another approach could be based on analyzing the connection between the cost structure of the various industries comprising the economy and the various life cycles.

As mentioned, the duration of the inventory cycle is directly related to the share of intermediate costs in total value of the product of various industries. On the same basis it would be logical to assume that the frequency of cycles connected with fixed capital is related to the share of depreciation and the cost of capital in total value. However, these issues should be explored in conjunction with the mechanism of investment determination.

An attempt to combine these approaches was made in Menshikov (1972), which discusses dynamic properties of input–output models supplemented by econometric difference equations that determine components of final demand (equipment investment by industry, construction investment by type, personal consumption by group) by variables derived from the cost structure of industries (intermediate inputs, depreciation, labor income, other income). The general structure of such models is (in matrix form)

$$X = (E-A)^{-1}\ Y \tag{6.4}$$

$$Y = S + I + G \tag{6.5}$$

$$sS = f_1\,[W(t-k),\, P(t-k)] \tag{6.6}$$

$$sI = f_2\,[P(t-k),\, K(t-k)] \tag{6.7}$$

$$W = f_3\,[V(t-k)] \tag{6.8}$$

$$P = V - W \tag{6.9}$$

$$V = (E-\alpha) X \tag{6.10}$$

$$K = (E-d) K (t-k) + I \tag{6.11}$$

$$S = \sum_{jj=1}^{3} s\bar{S} (i, jj) \tag{6.12}$$

$$I = \sum_{j=1}^{n} s\bar{I} (i, j) \tag{6.13}$$

$$s\bar{S} (i, jj) = c(i, jj)\ sS\ (jj)\quad i=1, \ldots, n;\ jj=1, 2, 3 \tag{6.14}$$

$$s\bar{I} (i, j) = b(i,j)\ sI(j)\quad i=1, \ldots, n;\ j=1, \ldots, n \tag{6.15}$$

$$\sum_{i=1}^{n} c(i, jj) = 1 \tag{6.16}$$

$$\sum_{i=1}^{n} b(i, j) = 1, \tag{6.17}$$

where:

X	is the vector or total value of industry output.
Y	is the vector of total final demand.
S	is the vector of personal consumption.
sS	is the vector of personal consumption by three types.
$s\bar{S}$	is the matrix of personal consumption by type and producing industry.
G	is the vector of exogenous final demand.
I	is the vector of total investment demand by producing industry.
sI	is the vector of industry investment.
$s\bar{I}$	is the matrix of investment by investing and producing industries.
W	is the vector of labor income.
P	is the vector of nonlabor income.

k	is time lag.
V	is the vector of value added.
K	is the vector of capital stock.
b	is the coefficient of investment structure.
E	is the unity matrix.
c	is the coefficient of consumption structure.
A	is the matrix of input–output coefficients.
$E-\alpha$	is the diagonal matrix with industry shares of value added in total value.
$E-d$	is the matrix of depreciation rates.

Estimated on the basis of US time series and using actual US input–output coefficients, the model generated a spectrum of periodic oscillations, including 3.5-, 8-, 20-, and 40-year cycles. However, the analysis was not completed and the precise mechanism, which chooses a specific group of dominant macroeconomic cycles among a long list of individual frequencies that depend on the various cost structures, has yet to be explored.

Another approach, as indicated above, is to study the interaction of different speeds of adjustment in systems of differential equations that describe the operation of an economic system. These equations represent the modes of adjustment of stocks to processes and vice versa, as well as between various stocks, processes, and their relation to macroeconomic ratios (i.e., profit rate, capital–output ratio). Each differential equation contains an exponential lag of one variable relative to another:

$$dx/dt = -a(x - by), \tag{6.18}$$

where a is the speed of adjustment of x to y and b is the coefficient of average proportionality between x and y. Coefficient b, besides assuring compatibility of variables measured in different units, also shows the direction in which x changes relative to bY, i.e., to its equilibrium position.

In market economies a is usually >0 and <1. When $a=0$, which is typical of a command economy, the initial deviation of x from its equilibrium remains constant, meaning that x ceases to react to changes in y, as it should. In command economies a is often <0. Then the initial deviation tends to grow exponentially. Not only is the equilibrium never reached, but the normal average relation between x and y is distorted, which may lead to the destruction of the whole mechanism, or to the creation of another one with substantially different characteristics (i.e., a shadow economy interlocked with a bureaucratic command system).

When a is close to or larger than unity, any deviation from equilibrium is eliminated practically immediately. This does not provide much flexibility, and may lead to sudden sharp reactions to large external shocks. Economic stability is better served by somewhat slower, but smoother, adjustment.

Though individual relationships between variables do not generate cycles *per se*, coefficient a plays a leading role in determining the frequency of the cycles, when generated by a system of equations of type (6.18). On the other hand, a is inversely dependent on the duration of the life cycle that underlies the relation between the variables.

Indeed, total return of a momentary investment is obtained only gradually. Equilibrium is reached only when total actual return equals expected return. It follows that speeds of adjustment should gravitate toward relative incremental speeds in actual life cycles. In other words, a in equation (6.18) is directly correlated with r in equation (6.1).

As mentioned, the sign of coefficient b shows the direction in which one variable changes relative to another. In economic systems variables are correlated both positively and negatively. The negative correlation is likewise a result of the life cycle. Though flows of investment finance are continuous, actual investment expenditure and the placement of capital stock are concentrated in relatively short periods of time. The speed of increasing capital stock is inversely correlated with accumulated capital stock. After adequate capital stock has been created, the necessity to replace or increase it emerges only after some time has passed.

Thus, life cycles generate both gradualness in the use of stocks (positive correlation) and discontinuity in their creation (negative correlation). A combination of both is a necessary condition for the emergence of periodic fluctuations. This is possible in both market and nonmarket economies, but their probability in market economies is much larger.

In a market economy negative correlations are also caused by the general use of certain criteria that control its dynamics. For instance, in the short run acceleration of fixed capital stock is positively correlated with the rate of profit, but negatively with accumulated capital stock.

Systems of differential equations with positive and negative relationships are analogous to simple mechanical systems, i.e., pendula, springs, flexible constructions with dampers. However, there is an important difference. In mechanics flexibility and dampening are usually determined independently, while in economic systems they are products of different combinations of mutually dependent forces. To take a simple example, in the well-known Samuelson model of the business cycle both flexibility and dampening are determined by different formulas, in each case containing different combinations of the accelerator and propensity to consume. In more complex models such combinations include many common structural coefficients.

Consider a simplified version of our Siena model of the long wave, presented in Menshikov and Klimenko (1985):

$$dy/dt = -a\ (y - bk) \tag{6.19}$$

$$dk/dt = -c\ (k - gp) \tag{6.20}$$

$$p = y - k, \tag{6.21}$$

where y is the growth rate of labor productivity, k is the growth rate of capital intensity (capital stock per man-hour), p is the growth rate of the profit rate, and a, b, c, and g are the structural coefficients. Note that in this model left-hand variables are actually second derivatives with respect to time and that y, k, and p are growth rates rather than absolute values. Labor productivity adjusts to capital intensity, and the latter to the profit rate. To simplify matters it is assumed that the share of profit in national product is constant, so that p is approximately equal to $y-k$.

The dynamic properties of this model are determined by characteristic equation

$$x^2 + [a+c(1+g)]x + ac(1+g-bg) = 0, \tag{6.22}$$

or, if $b=1$ (which is approximately true in the long run),

$$x^2 + [a+c(1+g)]x + ac = 0. \tag{6.23}$$

Regular cycles in growth rates are generated when $g=-2$ and $a=c$. Given these conditions, long waves of 50 to 60 years appear if speeds of adjustment equal 0.11–0.12; 20-year cycles, when $a=c=.34$; a 7-year cycle when they are both = 1; and a 3.5-year cycle, when $a=c=2$.

The unabridged model (6.19)–(6.21) with coefficients estimated from US time series for 1889 to 1982 generated cycles of 53.7 years. In this estimation a and c were far from equal ($a=.048$; $c=.25$). This is in agreement with expectations, since investment, represented by dk/dt tend to adjust to profit rates (the negative correlation) much faster than output adjusts to capital stock (the positive correlation).

However, the difference in coefficients a and c is important in determining the dampening force. The larger the ratio c/a, the smaller the critical absolute value of g necessary to produce explosive cycles. This follows from

$$a + c + cg < 0 \tag{6.24}$$

and

$$g < -(1+a/c)\,. \tag{6.25}$$

With $a=.048$ and $c=.25$, g has only to be smaller than −1.192 to generate explosive cycles. The estimated value for the whole period was −.61, or twice as large, meaning that there was a dampening factor of .273.

An enlarged version of the Siena model was presented in Menshikov and Klimenko (1986). Here fixed capital investment was disaggregated (using the latent variable technique) into three structural components:

(1) Extensive investment, representing increase in stock of existing technologies.
(2) Productivity investment, introducing new technologies that increase factor productivity.
(3) New products investment, used to create new products or spheres of production.

The speeds of adjustment $SA(i)$ of the ith type of investment were estimated to be $SA(1) = .094$, $SA(2) = .059$, and $SA(3) = .022$. The fastest adjustment was made by extensive investment, which follows changes in output. Nearly twice as slow was the adjustment of investment into new technologies, which followed past rises in productivity. The slowest was new products investment, which adjusted to changes in the profit rate.

Computer simulations of this model produced a superposition of fluctuations of approximately 20, 30 to 40, and 60 years. Particularly important was the dominance of the longer-term wave in total capital stock. Total net investment tended to gravitate toward 20-year fluctuations, apparently tied to extensive investment, while both new technology and new products investment determined the 30-to-40-year pattern.

6.4. Dynamic and Static Bifurcation

Linear differential models with constant coefficients, such as presented above, may turn out to be dynamically unstable when they generate explosive cycles. Instability of a different kind appears in them when the system has "negative flexibility," associated either with negative speeds of adjustment or with the absence of negative feedback relationships between variables. This is typical of command economies. In this case even smaller shocks to the system leading to deviations from the equilibrium path will tend to increase exponentially, making special and, perhaps, continuous exogenous efforts necessary to restore stability.

However, if a differential system is nonlinear, then the "galloping" effect may occur for both exogenous and endogenous reasons. Consider

$$y'' + b_1 y' + c_1 y + a_1 m V y' = 0, \tag{6.26}$$

where an external force, having velocity V, influences the system, whose receptivity equals a_1. Particularly important is coefficient m, which transforms the external influence in such a way that it may be either positive or negative.

If $m<0$, then the overall dampening effect $(b_1 + a_1 m V)$ may become negative, if V is larger than critical speed

$$Vc = b_1/[a_1 abs(m)]. \tag{6.27}$$

This generates explosive oscillations. In linear systems the explosion is infinite, leading to total destruction. In nonlinear systems, such as

$$y'' + by' + cy - aV[m_1 y' + m_2(y')^2] = 0, \tag{6.28}$$

the explosion continues only until it reaches a cycle of certain finite amplitude, which may, however, be sufficiently large so as to make impossible the normal operation of the system.

What is the economic meaning of these formulas? In real economic systems the gradual accumulation of a positive external effect may reach a certain critical point when the net effect suddenly becomes negative. For instance, when the Keynesian multiplier was first introduced, its accompanying negative effects were so small as to be largely ignored. However, over later decades these negative effects, caused by inflation and monopoly, weakened the supply response of the market to such an extent that the value of long-term, and in some cases of short-term, multipliers became negative.

Figure 6.3 shows the dependence of $m = d\text{GNP}/dG$ on $G/P\text{GNP}$, where G represents government purchases and PGNP is the potential output of the economy. If the share of government expenditure in total potential output is relatively low, the effect of an increment in G has an increasingly positive effect on actual output. However, as the share of government expenditure and GNP approaches capacity, the inflationary and bottleneck effects tend to lower the value of the multiplier. When both $G/P\text{GNP}$ and $\text{GNP}/P\text{GNP}$ pass critical levels, the multiplier becomes increasingly negative. This is particularly true of military expenditures, as shown in a number of recent studies.

Assume that the aggregated Siena model is modified by introducing restraints that add to equation (6.23) component $(+\ a_1\ m\ V\ y')$ and $(c_2\ y')$ where $c_1 = c + c_2$. In this modified form $(m\ y')$ represents the change in the growth rate of labor productivity (GNP/employment), produced by a change in the share of government purchases in GNP, whereas coefficient $a_1 = G/\text{GNP}$ measures the receptivity or susceptibility of the system. Then, using actual coefficients of the model and arbitrary assumptions of m, we obtain the following values of critical velocity $Vc = .1455/[a_1\ abs(m)]$:

a_1	m	Vc
.1	+.2	–
.2	+.1	–
.3	−.1	4.85
.4	−.2	1.82

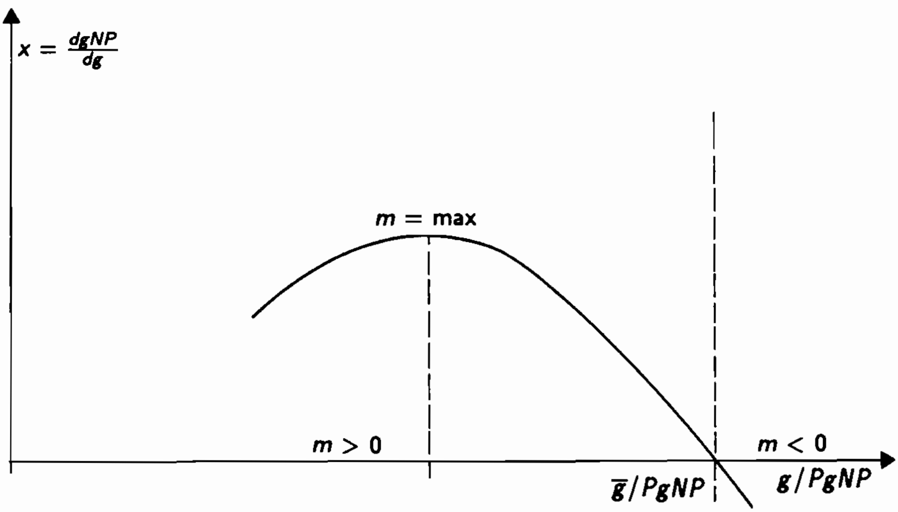


Figure 6.3. Dependence of multiplier on the share of government purchases in GNP.

Similar effects may be associated with other forces that may be totally or partially external. For instance, large government deficits have had adverse influence on the economy not only through the inflationary mechanism, but also through the "crowding out" effect in the money markets. The rise of transnationals and the globalization of capital markets have distorted the traditional mutual responses between money exchange rates and trade flows, leading to unprecedented imbalances in international payments and, presumably, to major stock market crashes.

To sum up, in cases when a previously positive influence suddenly becomes negative, there is danger of dynamic and, for that matter, static instability.

External, or partially external, factors may act to change structural coefficients of a system, which may themselves oscillate as a function of time. For instance, in Menshikov and Klimenko (1984) the Siena model was fitted not only for the full 90-year period, but also for various shorter periods of 20 to 24 years. In some cases the coefficients showed relative stability over time, while in others there were significant changes.

Thus, coefficient a, i.e., the speed of adjustment of labor productivity, changed relatively little from .25 to .5, but was much higher than the long-term value of .048. Coefficient b, i.e., proportionality of labor productivity to capital intensity, changed from a high of 1.4 in 1904–1924 to a low of .3 in 1924–1928, but remained stable in both 1894–1914 and 1946–1977 at .99 and .98. Coefficient c, i.e., speed of adjustment of capital intensity, was extremely volatile, reaching a high of .8 in 1900–1924 and 1953–1977 (long-term average was .25). Coefficient g, i.e., proportionality between capital intensity and the profit rate, was relatively stable at (–.3; –.4) from 1894 to 1948, but very volatile thereafter.

When fitted with these coefficients, the model generated cycles of 15 to 19 years duration (in one case – 9 years) that were mostly explosive. In real economic life structural coefficients tend to oscillate in a nonlinear manner even within cycles of medium (7 to 10 years) or short (3 to 4 years) duration. This may also, at times, produce explosive cycles. When an economy finds itself in dynamic instability, it cannot maintain an equilibrium on its own merits, and has to undergo substantial structural change to avoid complete destruction.

In differential systems, like those described above, most economic variables are assumed to be interdependent. Such a system has only one equilibrium for one set of coefficients. Any deviation from this equilibrium tends either to be eliminated with time (autocorrection) or to increase infinitely or to a certain point which is far from equilibrium.

But in real life, economies may at any given moment have two or more equilibrium positions, with the possibility of sudden change from one to another. In mathematical terms such changes are called static bifurcations and are explored in the Catastrophe theory.

In dynamic bifurcations the catastrophe (or "galloping") is associated with the possible destruction of the system. A new system with new structural coefficients and another regulatory mechanism may emerge instead.

In static bifurcations changes from one equilibrium to another may occur within the same system of structural coefficients, and may not necessarily lead to its destruction. Such a catastrophe may be followed by an anti-catastrophe, which restores the former equilibrium.

Static bifurcation in an economy usually occurs when there is a significant and lasting deviation of two or more variables from their optimal mutual relationships as reflected by the structural coefficients. For instance, as shown in part one, this happens when the actual return curve either intercepts the expected return curve earlier than expected, or not at all. In both cases the impossibility of completing the life cycle and passing to a new one, as originally planned, creates additional pressure or friction in the system. This is particularly true in the latter case, when the necessity of foregoing previous profit maximization targets is forced upon the economy. In the former case the situation is more favorable, but the possibility of a sudden drastic change is still possible. Both situations may happen within market and planned economies.

The existence of a second (third, etc.) equilibrium may serve as a certain guarantee that the underlying system may survive. The seeming absence of alternative equilibria, on the other hand, does not prove the durability of the system that may well break down under pressure. Of course, potential equilibria may exist that are tantamount to nonexistence of the current system. This is so in physics (consider the change of cold starlike mass from small dwarves to neutron stars) and in economics (like the sudden transformation of a poorly planned economy into a mixed command/shadow market system).

Even in an economy with damped cycles external shocks may generate such large amplitudes that transcend critical pressures or, in the nonlinear case, lead to chaotic movement (cf. Rasmussen *et al.*, 1985). The critical pressures are defined as frontiers beyond which a bifurcation becomes possible, and in some cases inevitable.

The mathematical Catastrophe theory is intended to establish such critical frontiers, or zones. It is possible to single out from any n dimensional system one or more variables, as control parameters, whose small and gradual changes in the vicinity of certain points bring about bifurcations, or catastrophes inherent to the system. The concrete form of catastrophe depends on the number of control parameters.

Assume that in market systems output is determined by factors of production, however, not directly but through the profit rate. Simple example: if we choose the capital–output ratio as the only control parameter then there is only one type of catastrophe possible, namely, "the fold." The production function in this case is described as a third-order surface:

$$X = P^3 + (K_x)P, \tag{6.29}$$

where X is output, P is profit rate, and K_x is capital–output ratio. Their relation is shown in *Figure 6.4*. For relatively small K_x there is only one stable equilibrium for X relative to P (as demonstrated by the minimum points of curves X_1 and X_2. As K_x increases, so does the pressure on the system. After it reaches a certain critical value (represented by curve X_3) surface $X(P, K_x)$ sags in such a manner as to produce three alternative optimal points (illustrated by X_4 and X_5). The area between them defines the critical zone of the profit-rate values that make bifurcation possible and inevitable. The higher K_x the wider the critical zone bordered by curve P_{CR}.

The source of bifurcation in this example can also be traced to the life cycle. A high capital–output ratio may mean that actual return, as represented by output, is lower than expected return, as represented by capital. This creates alternative equilibria: the local minima are points where the profit rate is maximized relative to output, whereas the local maximum indicates maximization of output.

A mechanical analog to this system is pressure brought upon an arch where ends are connected by a spring. When pressure brought upon the arch reaches a critical level, the arch will suddenly jump from its upper to lower equilibrium position.

In Klimenko and Menshikov (1987) labor productivity (Y) and capital intensity (K) were chosen as the two control parameters, assumed to serve as substitutes for, respectively, intensive and extensive technical progress. Note that the ratio of K to Y is identical to the capital–output ratio, which in this case is expressed by two, rather than one control parameter. Output is assumed to be a nonlinear function of the profit rate and of the two parameters. Employment (L) influences output in a linear manner, thus not directly affecting bifurcational characteristics of the system.

With two control parameters any process $X(q)$, where q is the vector of independent variables, is described by surface

$$X = q^4 + \lambda_1 q^2 + \lambda_2 q \tag{6.30}$$

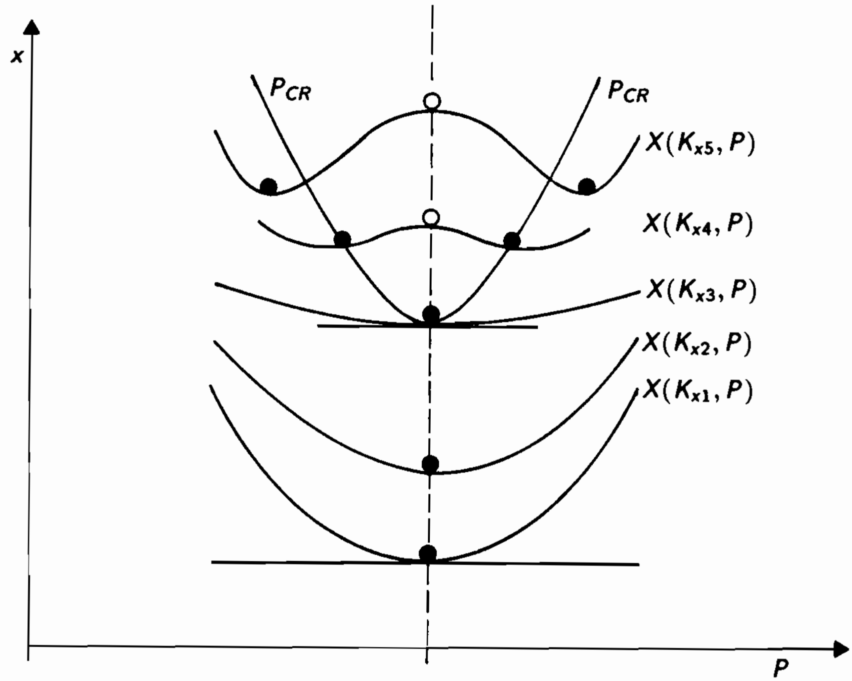


Figure 6.4. Static bifurcation depending on capital–output ratio.

and the possibility of catastrophes by set

$$dX/dg = 0\ . \tag{6.31}$$

Thus it is logical to express production functions described above as

$$X = a_1\, P^4 + a_2\, P^3 + a_3\, P^2\, Y + a_4\, PK + a_5 L \tag{6.32a}$$

or

$$X = a_1\, P^4 + a_2\, P^3 + a_3\, P^2 K + a_4\, PY + a_5\, L\ , \tag{6.32b}$$

where P is profit rate, Y is labor productivity, K is capital intensity, and L is employment.

When a_1 and a_2 are small, equations (6.32a) and (6.32b) are similar to usual production functions weighted by the profit rate. The two variants reflect conditions prevailing under, respectively, intensive and extensive growth. On those parts of the surface that are far from the critical zones both surfaces are well approximated by linear production functions.

We then test the hypothesis that actual output moves along paths that are close to optimal points of X relative to P, i.e., close to trajectory:

$$dX/dP = 4a_1\ P^3 + 3a_2P^2 + 2a_3P\ Y + a_4\ K + 0\ . \tag{6.33}$$

One would assume a positive correlation of output relative to productivity and, in most cases, to capital intensity (and thus $a_3>0$; $a_4>0$). This proved to be true in most empirical estimations. However, the sign and significance of a_1 and a_2, representing the highly volatile influence of the profit rate are subject to more speculation. As shown below the signs of these coefficients are dependent on the prevailing direction of technical progress and thus on the movement of the capital–output ratio within the period under estimation.

For any combination of Y and K, there may be three different values of P corresponding to this criterion. Then there are three equilibrium positions of X relative to P, which – depending on the sign of a_1 – are either two maxima and one minimum ($a_1>0$), or one maximum and two minima ($a_1<0$). If there are complex roots, only one equilibrium exists. Negative equilibrium values of X do not exist in real economies and are excluded from the analysis.

Equations of the type (6.32a) and (6.32b) were fitted to US time series for 1889–1982. It was found that actual values of GNP, indeed, were close to one of the equilibria, i.e., for each pair of actual Y and K the rate of profit calculated from equation (6.33) was close to actual P. As in the case of the Siena model the whole period was subdivided into sub-periods of 24 to 27 years corresponding to upturns or downswings in the long wave. In two periods (namely, from 1889 to 1939) there were two positive real roots, and in two other periods (from 1933 to 1982) one positive real root. In periods of economic crises there were leaps from one equilibrium to another, typically from Xmin in non-crisis years to Xmax in crisis years.

This was explained by the coexistence within the market economy of forces of monopoly and competition. Catastrophes occurred when the predominance of monopoly (and thus of profit maximization) gave way to prevailing competitive conditions (and thus to output maximization). It was also noted that bifurcation took place in periods when the ratio K/Y reached its local maxima, i.e., when the dominance of extensive technical progress was at a peak, opening the door for a new technical revolution. At such points forces of competition became particularly active helping to storm and destroy some of the traditional monopoly bastions. In the 1930s such changes were particularly noticeable due to the overall fall of commodity prices and exchange rates. In the 1970s and 1980s there was also evidence of drastic leaps, particularly in the prices of fuels and raw materials, as well as in exchange rates. Extraordinarily large imbalances in international payments and government budgets and enormous accumulated unrepaid external debts are symptoms of large inconsistencies and superpressures that may provoke a maxi-catastrophe. These peculiarities did not show up in the corresponding equation owing perhaps to the fact that the large imbalances appeared after 1980 and did not show in our time series.

Nonlinear output surfaces described above were further explored to establish the extent of their relative stability. For every nine years centered around every consecutive year in US time series from 1893 to 1977, coefficients a_1, a_2, a_3, and a_4 in equation (6.32) were estimated by a two-stage procedure. Assuming that they are functions of time and that output moved along paths close to equilibrium points of X relative to P, i.e., $dX/dP = 0$, we obtained

$$dX/dt = (dX/dY)\,(dY/dt) + (dX/dK)\,(dK/dt)\ . \tag{6.34}$$

Note that with $dX/dY{=}a_3\,P^2$ and $dX/dK = a_4\,P$ we get

$$dX/dt = a_3\,P^2\,(dY/dt) + a_4\,P\,(dK/dt) + e(t) \tag{6.35a}$$

or

$$dX/dt = a_4\,P\,(dY/dt) + a_3\,P^2\,(dK/dt) + e_t. \tag{6.35b}$$

Having estimated a_3 and a_4 from equation (6.35) we then estimate a_1 and a_2 by substituting a_3 and a_4 into equation (6.32). In this way equilibrium points of X relative to P were calculated for every combination of K and Y in every year from 1893 to 1977. The use of the nine-year centered estimation makes it possible to eliminate most shorter-term influences.

The results of this exercise have yet to be fully analyzed. Let us point to a few preliminary conclusions:

(1) The X-surfaces changed from year to year. In most instances the sign of a_1, which largely determines the quantity and sequence of maximum and minimum equilibria, remained unchanged from year to year. In 15 cases, however, it did change, indicating a more substantial transformation of the surface. This last case is important for the following reason. If two positive equilibrium points exist, then, if $a_1{>}0$ the left-hand one is Xmax and the right-hand one is Xmin. When $a_1{<}0$ the left positive equilibrium is Xmin and the right one is Xmax.

(2) In mechanical systems the Xmax equilibrium point is usually unstable, and Xmin is stable. The leaps, when they occur, are either from Xmax to Xmin, or from one Xmin to another. In economic systems the leaps, as evidenced by this exercise, are always between an Xmax and an Xmin, or vice versa. The reason is that the economy is driven to one of the equilibria by either profit or output maximization. When $a_1{<}0$ both forces work in the same direction: the larger the output, the higher the profit rate, and vice versa. If in these conditions the economy finds itself at a minimum, it is under the overwhelming influence of inadequate aggregate demand. If $a_1{>}0$ then the profit and output maximization work in opposite directions,

and a leap to Xmin may be caused by profit maximization being stronger than demand-pull and output maximization.

(3) It was found that in most cases the $a_1>0$ situation occurred when K was increasing faster than Y, i.e., when the capital–output ratio was rising, and vice versa. Thus a change in the sign of a_1 was generally correlated with the change in the direction of movement of K/Y over time. This is shown in *Figure 6.5*. An increase in K/Y usually corresponds to the downward movement in the long wave, and it is at this point that the forces of output and profit maximization act in the most mutually contradictory manner.

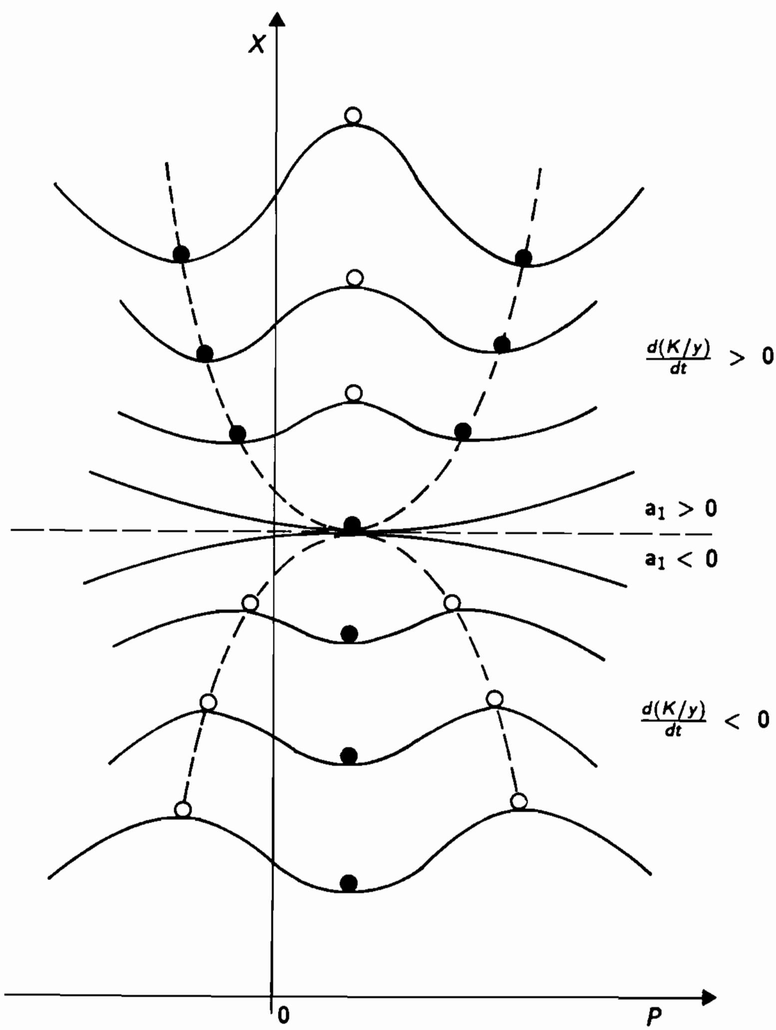


Figure 6.5. Static bifurcation depending on the relation of two control parameters K/Y.

(4) Contrary to this general rule a change to $a_1<0$ also occurred in 1930–1931 and in 1973–1975, in both cases at a time when K/Y was rising. In the first case there was a severe fall from Xmax to Xmin. These were conditions of acute crisis of overproduction, when demand inadequacy forced the economy down. In the second case output remained closer to Xmax. Apparently other forces, i.e., stagflation and the oil shock, were stronger than the more traditional demand inadequacy.

(5) In this exercise every estimated X-surface largely reflected the nine-year medium-term cyclical characteristics. Therefore its results could not be identical to the earlier exercise that concentrated on the longer-term cyclical features. From a longer-term perspective the economy chooses either Xmax or Xmin as its predominant and, therefore, stable equilibrium point, changing to the unstable equilibrium only in the time of crises. From a shorter-term perspective, the X-surface itself changes during critical times, making Xmin as stable during crises as Xmax is during times of cyclical upturns.

References

Klimenko, L., 1972, The Application of Spectral Analysis to Business Cycles Research, Doctoral Dissertation, Moscow [in Russian].

Klimenko, L. and Menshikov, S., 1987, Catastrophe Theory Applied to the Analysis of Long Waves, in T. Vasko, ed., *The Long-Wave Debate*, Springer-Verlag, Berlin, Germany, F.R.

Menshikov, S., ed., 1972, *Dynamic Models of the Economy*, Novosibirsk [in Russian].

Menshikov, S. and Klimenko, L., 1984, *Long Waves in the Economy (in Processes of Global Development: Models and Analysis*, Vol. 3, Moscow [in Russian].

Menshikov, S. and Klimenko, L., 1985, On Long Waves in the Economy, in G. Bianchi *et al.*, eds., Collaborative Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Menshikov, S. and Klimenko, L., 1986, Structural Change and the Long Wave, in *Technological and Social Factors in Long Term Fluctuations*, International Workshop, Siena, Italy.

Rasmussen, S., Mosekilde, E., and Sterman, J.D., 1985, Bifurcation and Chaotic Behavior in a Simple Model of the Long Economic Wave, in *Systems Dynamics Review 1*, Summer.

Thompson, J.M.T., 1982, *Instabilities and Catastrophes in Science and Engineering*, Chichester, New York, NY, USA.

CHAPTER 7

Economic Structural Changes and Waves in Technological Progress

Iouri Tchijov and Irene Sytchova

7.1. Introduction

One of the explanations of long-term waves in economic development is based on technological progress or the interpretation of technological waves as a main source of long-term oscillations in economic systems. On the other hand, technological progress can be measured at a macroeconomic level as a process of relative input saving. This is why it might be interesting to analyze long-term dynamics of total labor requirements as well as their components.

If the cost decrease is a target of technological progress, one of the progress assessments is a change in the cost structure. It is due to the uneven impact of new technologies (NT) on different elements of product costs. Hence it is possible to determine three types of technological progress: labor saving, fixed capital saving, and material saving. In reality, any new technology or technological progress as a whole changes all the three elements of cost, but historically it is possible to determine the periods when one of them dominated. The reasons for such an uneven influence are the dynamics of relative prices of each resource (on the demand side) and the objective dynamic possibilities of scientific and technological development.

7.2. Methodological Approaches

In our investigations we use seven input–output tables of Japan for 1951, 1955, 1960, 1965, 1970, 1975, and 1980.[1–3] All the tables were aggregated to 18 industries and re-estimated into the 1970 prices.[4] Thus we have the time series of completely comparable input–output tables:

$$b_j = \sum_i a_{ij}\, b_i + l_j, \tag{7.1}$$

where b_j and b_i are the total labor requirement coefficients, l_j is the direct labor requirement coefficient, and a_{ij} is the input–output coefficients. If we divide a_{ij} into two parts,

$$a_{ij} = \alpha_{ij} + \beta_{ij}\ , \tag{7.2}$$

where α_{ij} reflects the use of materials produced by a ith industry and β_{ij} reflects ith capital consumption for unit production in a jth industry.

To estimate β_{ij} it is necessary to disaggregate capital consumption allowances for each industry into its elements – fixed capital goods produced by certain industries. We have chosen five such industries (construction, c; general machinery, gm; electrical machinery, em; transportation equipment, te; and agriculture, a) as well as a technological structure of fixed assets (houses, 1; constructions, 2; equipment and machines, 3; ships, 4; other transportation equipment, 5; instruments and fixtures, 6; land improvement, 7; plants and animals, 8; and incomplete construction, 9). The algorithm of industrial classification transformation is shown in *Table 7.1*.

Then we got the experts' estimates of lifetime for different types of fixed capital assets as 33 years for houses and constructions, 11 years for industrial equipment, 5.7 years for transportation equipment, and 8 years for plants and animals in agriculture.[5] As a result we estimated the distribution structure for capital consumption allowances in each industry (Φ_j) (see *Table 7.2*) based on the use of the real structure for 1965 and assumptions about average lifetime mentioned above.[6]

The data in *Table 7.2* mean that for industry 8, for instance, capital allowances (Φ_8) will be distributed as flows of capital goods (x'') in the following way:

$$x''\ 3.8 = 0.30\ \Phi_8$$
$$x''\ 10.8 = 0.43\ \Phi_8$$
$$x''\ 11.8 = 0.07\ \Phi_8$$
$$x''\ 12.8 = 0.20\ \Phi_8\ .$$

This method permits us to develop the matrix of fixed capital allowances as

$$\beta_{ij} = x''_{ij}/x_j \tag{7.3}$$

estimated by the above-mentioned methods,

Table 7.1. Industrial classification transformation.

Industries	*Technological structure of assets*
c	1, 2, 9
gm	3
em	6
te	4, 5
a	7, 8

Table 7.2. Distribution structure for capital consumption allowances (%).

	Material production (No. 2-13, 17)[a]	*Agriculture (No. 1)*	*Transportation and communication (No. 16)*	*Nonmaterial production (No. 14, 15, 18)*
3,c	30	27	27	50
10,gm	43	40	40	–
11,em	7	7	6	12
12,te	20	20	27	37
1,a	–	6	–	–
Total	100	100	100	100

[a]Numbers of industries correspond to the list in *Table 7.3.*

$$x_{ij} = x'_{ij} + x''_{ij}, \tag{7.4}$$

$$\alpha_{ij} = x'_{ij}/x_j \tag{7.5}$$

from I–O tables, and

$$a_{ij} = \alpha_{ij} + \beta_{ij}$$

new I–O coefficients. Thus we have transformed the capital consumption row of the third quadrant of the I–O table into additional elements of input–output coefficients.

In the postwar economy of Japan, import traditionally covered a certain part of input. To exclude the influence of imports we need the same purification of I–O tables. The most reasonable approach is based on import subtraction from the I–O matrix, element by element. But there are only two import matrices in the Japanese statistics (for 1970 and 1980), and we had to develop an approximate algorithm of such an elimination.

If $M_{i,t}$ is a volume of imports of products of the ith industry, $x^d_{i,t}$ is a volume of this industry's domestic production (all in year t), the share of domestic production in the total ith product consumed in the economy will be

$$d_{i,t} = x^d_{i,t}/(x^d_{i,t} + M_{i,t}). \tag{7.6}$$

Of course, we used a rather strong assumption that for a given year t the import share is the same for different ways of product use. But it will change in dynamics, from one year to another.

The modified flows, purified from the import part, will be for each I–O table:

$$x^d_{i,j} = x_{i,j} \cdot d_i \tag{7.7}$$

where x_{ij} is taken from equation (7.4).

To prove the applicability of this method we compared two matrices of domestic flows – the first one was estimated from the official import statistics for 1970 and the second one was estimated for the same year by using the proposed method.

The correlation coefficients for these two vectors estimated for each industry (from 1 to 18) were more than 0.99 except for one case – transportation and communication in which the coefficient was 0.92. These results can be treated as evidence of the method acceptability.

Thus we reconstructed the input–output tables where the flows of products were purified from import and divided into material and fixed capital consumption. It is possible to disaggregate the coefficients of total labor requirements (b) into three coefficients: direct labor requirements (l), material requirements (b^m), and final capital requirements (b^c). If A is a matrix of direct material and capital requirements and A^m is a matrix of direct material requirements, b and its components will be defined as follows:

$$b = l(E - A)^{-1}, \tag{7.8}$$

$$b^c = l(E - A)^{-1} - l(E - A^m)^{-1}, \tag{7.9}$$

$$b^m = l(E - A^m)^{-1} - l, \tag{7.10}$$

$$b = l + b^m + b^c. \tag{7.11}$$

The shares of labor, material, and capital consumption in the total cost of product will be defined, respectively:

$$S_l = \sum_j l_j \, x_j / \sum_j b_j \, x_j, \tag{7.12}$$

$$S_m = \sum_j b_j^m \, x_j / \sum_j b_j \, x_j, \tag{7.13}$$

$$Sc = \sum_j b_j^c \, x_j / \sum_j b_j \, x_j. \tag{7.14}$$

Consequently, to define the type of technological progress for a certain period, it is necessary to estimate the impacts of these three components on the total labor requirement reduction and to find the main one.

7.3. Technological Progress, Economic Growth, and Structural Changes

The data obtained for the output growth as well as for the direct and total labor requirements reduction are shown in *Table 7.3*. It is obvious from both viewpoints that uneven growth took place in the Japanese economy – in dynamics and in the industrial structure. The accelerated output growth of the 1950s and 1960s was substituted for the economic growth deceleration in the 1970s. A similar tendency was observed in labor requirements (direct and total).

Let us analyze two hypotheses. The first one is that the postwar technological progress promoted a decrease of big industrial differences (in total labor requirements), which took place at the beginning of the 1950s. The second hypothesis claims that industries with higher labor requirements at the beginning tended to decrease their requirements to an average level, or industries with a higher starting point experienced a more rapid decrease later.

To test the first hypothesis we estimated the dynamics of the relation between a standard deviation and a sample average ($S/\bar{Y}$) for the total labor requirements. The results (see *Figure 7.1*) show that technological progress in the Japanese economy led to a reduction of the total labor requirement in all industries, and the industrial differences moved down during 1951–1965 and then up again.

We tested the second hypothesis for the total labor requirements (TLR), estimating the correlation between decrease rates (TLR 80/TLR 51) and the starting levels of the variables in 1951 (kTLR 51). These estimates were made for 17 industries (No. 2–18 in *Table 7.3*). The results, displayed in *Figure 7.2*, show that this hypothesis is acceptable for the total labor requirement case.

If we compare the growth rates for industrial outputs with the decrease rates for the total industrial labor requirement coefficients we can find a certain relationship between them. Higher reductions in the cost of production usually correspond to higher production growth rates. The total rank correlation coefficient of these two variables for 17 industries was (0.65) from this point of

Table 7.3. Dynamics of direct and total labor requirement coefficients (number of employees per 1 million yen).

	Direct								*Total*								*Output growth*
Industries	*1951*	*1955*	*1960*	*1965*	*1970*	*1975*	*1980*	*1980–1951*	*1951*	*1955*	*1960*	*1965*	*1970*	*1975*	*1980*	*1980–1951*	*1980–1951*
1. Agriculture	3.9	2.9	2.3	1.8	1.4	1.0	0.8	0.207	4.5	3.8	3.0	2.3	1.8	1.2	1.0	0.230	1.73
2. Mining	1.8	1.6	1.2	0.6	0.2	0.1	0.1	0.056	3.7	2.7	1.9	1.0	0.5	0.4	0.3	0.072	3.65
3. Construction	0.5	0.6	0.6	0.4	0.2	0.2	0.2	0.460	1.1	1.4	1.6	1.0	0.6	0.6	0.5	0.395	7.91
4. Food	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.258	2.0	2.8	1.9	1.3	0.9	0.7	0.5	0.236	5.40
5. Textile	1.1	0.8	0.6	0.5	0.4	0.3	0.3	0.234	6.1	4.0	2.0	1.5	1.0	0.7	0.5	0.085	6.55
6. Paper	0.8	0.6	0.4	0.3	0.2	0.1	0.1	0.125	5.1	3.4	1.8	1.2	0.7	0.4	0.4	0.073	16.53
7. Chemicals	0.9	0.5	0.3	0.2	0.1	0.1	0.1	0.069	4.9	2.8	1.4	0.8	0.4	0.4	0.2	0.045	21.92
8. Primary metals	0.4	0.3	0.3	0.2	0.1	0.1	0.1	0.143	4.9	1.6	1.3	0.8	0.5	0.3	0.2	0.045	15.62
9. Fabricated metal products	1.2	1.3	1.0	0.6	0.3	0.3	0.2	0.198	3.0	2.3	1.8	1.0	0.7	0.6	0.4	0.131	22.81
10. Nonelectrical machinery	0.7	0.6	0.4	0.3	0.1	0.1	0.1	0.085	2.3	1.7	1.3	0.9	0.5	0.4	0.2	0.091	30.43
11. Electrical machinery	1.6	1.2	0.5	0.4	0.2	0.2	0.1	0.050	4.7	2.8	1.6	1.1	0.6	0.4	0.2	0.041	150.24
12. Transportation equipment	1.1	0.8	0.4	0.3	0.2	0.1	0.1	0.072	3.8	2.0	1.4	0.9	0.5	0.4	0.3	0.066	47.40
13. Other manufacturing	0.9	0.8	0.7	0.5	0.3	0.3	0.2	0.268	2.6	3.6	2.4	1.5	0.9	0.6	0.5	0.182	9.32
14. Trade	2.3	2.0	2.0	1.2	0.7	0.6	0.5	0.208	3.3	2.6	2.4	1.5	0.9	0.8	0.6	0.174	14.89
15. Finance, real estate	0.4	0.3	0.2	0.2	0.1	0.1	0.1	0.237	0.7	0.7	0.4	0.4	0.3	0.3	0.2	0.297	20.20
16. Transportation, communication	1.5	1.2	0.9	0.7	0.4	0.2	0.2	0.147	2.9	2.2	1.4	1.0	0.7	0.5	0.4	0.140	15.01
17. Public utilities	0.5	0.4	0.2	0.2	0.1	0.1	0.1	0.130	1.6	1.4	0.8	0.6	0.4	0.3	0.2	0.135	11.82
18. Services	0.8	0.6	0.7	0.6	0.5	0.5	0.4	0.443	2.5	1.1	1.1	1.0	0.8	0.8	0.5	0.218	5.70
Change in total, %	–	–27.0	–23.0	–30.0	–40.0	–19.0	–19.0	–	–	–26.0	–27.0	–33.0	–38.0	–21.0	–29.0	–	

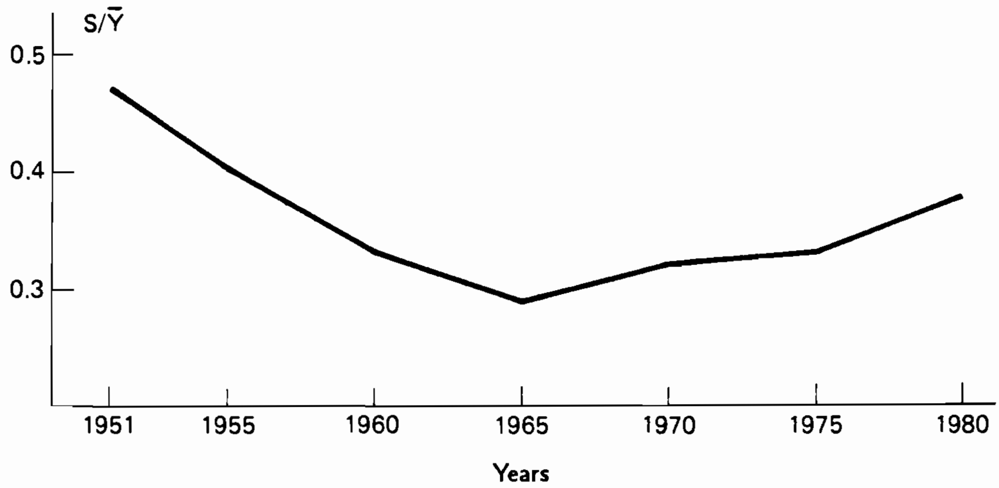


Figure 7.1. The dynamics of industrial differences (measured as relation of standard deviation to sample average) in total labor requirements.

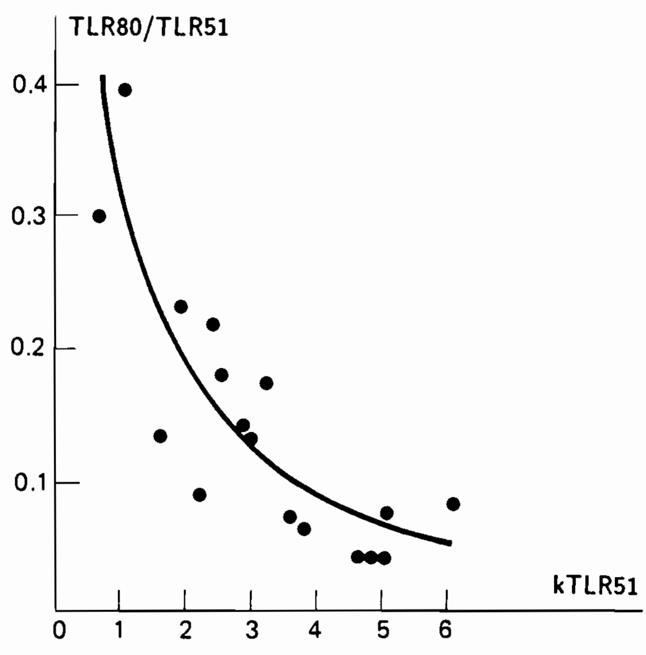


Figure 7.2. The dependencies of labor requirement coefficients changes on their starting values.

view it is possible to define "progressive" industries (like electrical and nonelectrical machineries, transportation equipment, and chemicals) and less technologically intensive industries (like agriculture, services, construction, and food).

For example, output in electrical machinery increased by a factor of 150 in 1951–1980, in transportation equipment by 47, and in nonelectrical machinery by 30. During the same period total labor requirements in these industries dereased by a factor of 24, 15, and 11, respectively. On the other hand, output in agriculture, services, and the food industry increased only by a factor of 2, 6, and 7, respectively, and total labor requirements declined by 4–5 times. One can find that there is a certain interdependence between the rate of output growth and the rate of cost (total labor requirements) reduction (see *Figure 7.3*).

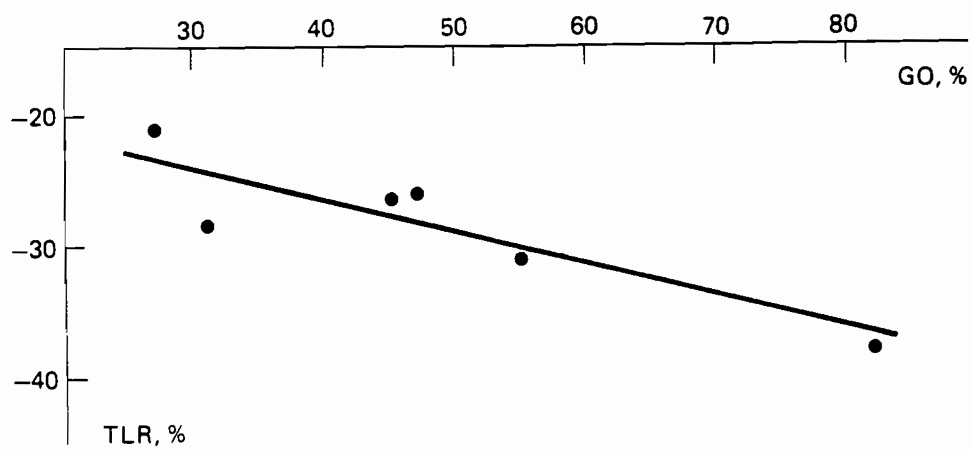


Figure 7.3. Five-year changes in total labor requirement (TLR) versus five-year changes in gross output, constant prices (GO).

But during the postwar period we found big differences in this relationship. Let us compare the interpolated regression between the rate of total labor requirement reduction (TLR) and changes in industrial outputs (IO) (see *Figure 7.4*). The dynamics of the slope coefficient reflects a tendency to decrease from 0.37 for the first period up to 0.22 for 1965–1970 and a lack of such a relationship in 1970–1975. Then the coefficient increased again in 1975–1980. T values shown in parentheses reflect the decrease of the coefficient estimates probability.

Thus we can observe certain tendencies in long-term dynamics of variables reflecting waves in technological progress. The dynamics of total and direct labor requirements, the changes in industrial differentiation (see *Figure 7.1*), and the changes in the relationships between economic growth and cost reduction (see *Figure 7.4*) show that the first part of the 1970s was a turning point in long-term waves, connected with technological progress.

7.4. Three Types of Technological Progress and Their Dynamics

If the total labor requirements are disaggregated into three components – labor, material, and fixed capital inputs, see equations (7.12)–(7.14) – the technological progress materialized in total cost reduction can be divided into three types: labor, material, and capital savings. By using the method described earlier and the seven input–output tables for the postwar economy of Japan we derived the data reflecting these three types of technological progress in dynamics (see *Table 7.4*).

It is obvious that at any time during the investigated postwar period the technological progress combined all three types. But in each period one type usually dominated. The material-saving type took place in all periods playing a main role in 1955–1965. The labor-saving type of technological progress played a growing role and reached the biggest share in the total cost reduction in 1965–1970. The important impact of capital saving on total cost (or labor requirements) reduction took place in 1975–1980.

The first period (1951–1955) of technological progress belonged to the labor- and material-saving types. In 1955–1965 material saving dominated, but the role of labor saving was growing and in 1965–1970 labor saving was the leading type of cost reduction. In 1970–1975, when the first postwar energy crisis occurred, material saving became more important again. Finally, 1975–1980 was the only period when fixed capital saving dominated.

As a result (see *Table 7.5*), the labor share in total cost grew in the 1950s, was stable in the 1960s up to 1975, and then grew again. The material share decreased during the whole period but with different rates, and the capital share in total cost increased up to 1975 and decreased afterward.

There were certain correlations between the three types of technological progress in the industries. The comparison of the change rates (1951–1980) in labor (LR), material (MR), and capital (CR) savings is shown in *Table 7.6*. For instance, the value of a direct labor requirement coefficient in agriculture in 1980 equals 21% of its value in 1951, the value of a material input coefficient in construction in 1980 equals 32% of its initial value, etc.

From the viewpoint of labor saving, the three best industries were electrical machinery, mining, and chemical, but the three worst were food, services, and construction. From the material-saving viewpoint the best ones were primary metals, electrical machinery, and chemicals, and the worst ones were finance (plus real estate), construction, and agriculture. And finally, from the capital-saving viewpoint the best industries were electrical machinery, primary metals, and chemicals. Finance, agriculture, and construction can be regarded as the worst three. On the whole the most progressive industries (minimum rank sum) in the Japanese economy were electrical machinery and chemicals; construction and agriculture can be regarded as smokestack industries.

The theoretical discussion about the relationships between labor–material–capital savings (either they compete or coincide in dynamics) might be added with our estimates based on the rank data shown in *Table 7.6*. The coefficients of the rank correlation are the following: 0.83, between material

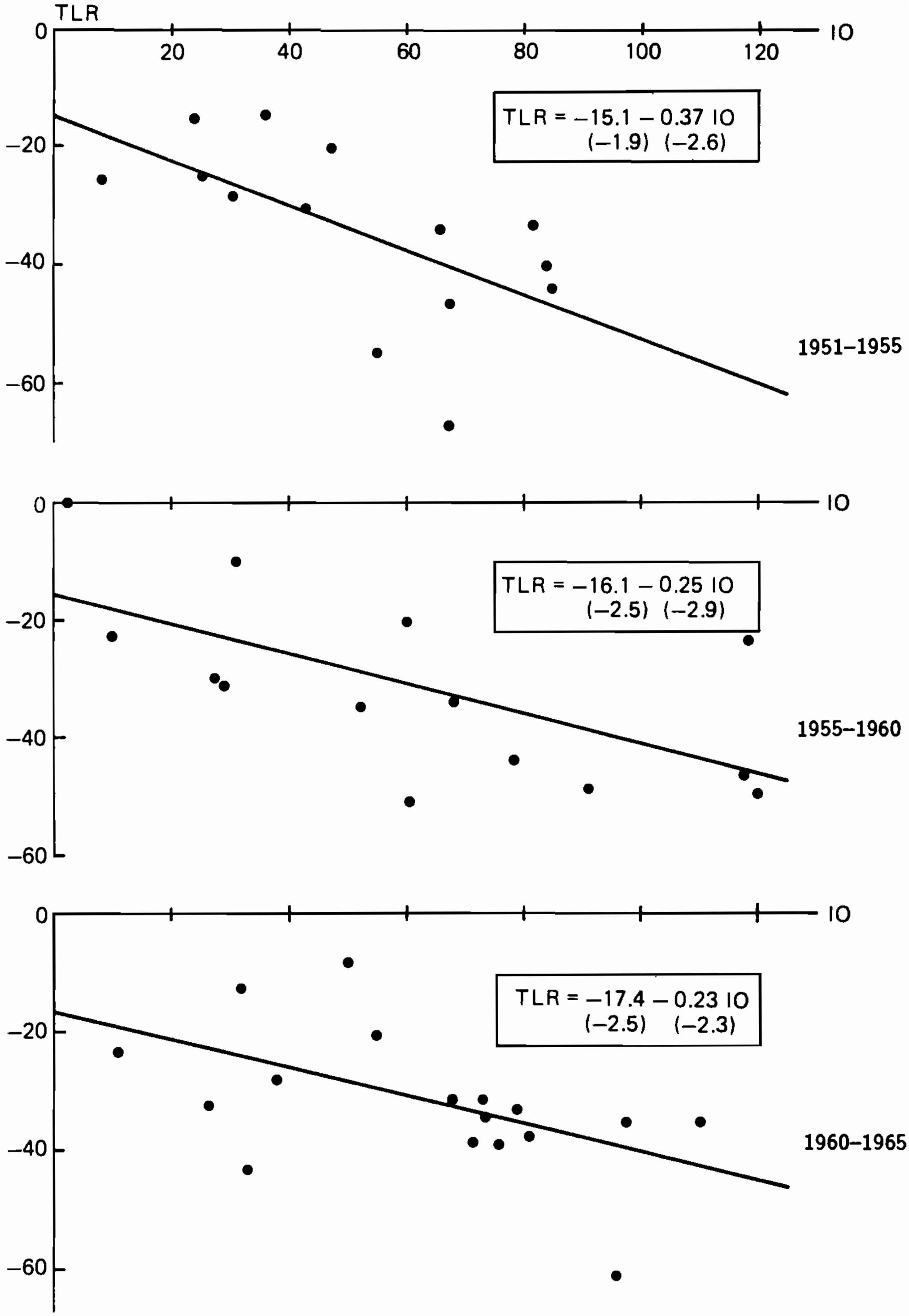


Figure 7.4. Corelation between five-year changes in industrial output (IO) and in total labor requirements (TLR), %.

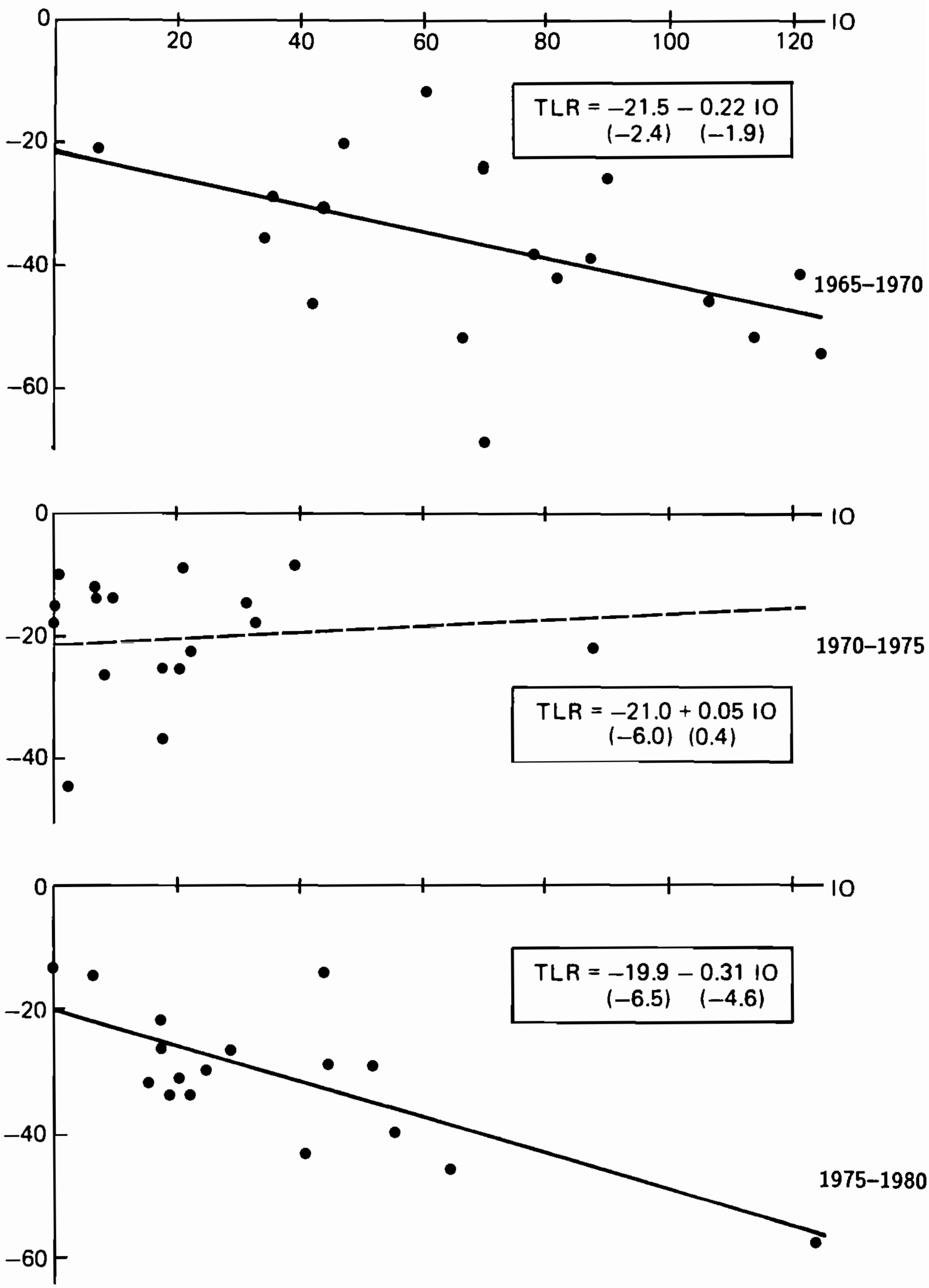


Figure 7.4. Continued.

Table 7.4. Rate of cost elements changes (%).

	1951–1955			1955–1960			1960–1965			1965–1970			1970–1975			1975–1980		
Industry	*L*	*M*	*C*	*L*	*M*	*C*	*L*	*M*	*C*	*L*	*M*	*C*	*L*	*M*	*C*	*L*	*M*	*C*
1	−24*	+44	+39	−20	−29	−36*	−25*	−24	+ 7	−19	−21*	− 4	−32	−35*	−18	−17	−15	−20*
2	−12	−41*	−30	−22	−48*	+ 5	−52*	−49	−11	−61*	−42	−21	−39*	+13	−22	−29	−33	−39*
3	+24	+21	+57	−10*	+24	+43	−30	−40*	−12	−38*	−38	−12	0	−18*	− 8	− 4	−31	−48*
4	−35*	+58	+75	−20	−32*	−31	− 6	−34*	+ 2	−27	−31*	−17	− 9	−31*	− 5	−20	−35	−45*
5	−26	−37*	−10	−28	−57*	−45	−17	−33*	+3	−27	−34*	−24	−19	−39*	−18	−10	−34	−46*
6	−28	−36*	− 8	−34	−50*	−22	−32	−40*	+ 5	−42	−44*	−40	−13	−60*	−45	−23*	+26	0
7	−41	−44*	−27	−41	−54*	− 8	−43	−48*	−11	−47	−51*	−45	−11	−24*	+ 4	−25	−37	−51*
8	−33	−71*	−60	− 4	−28*	+20	−37	−40*	−13	−47*	−47	−33	−22	−35*	−13	−14	−31	−52*
9	+12	−44	−54*	−24	−25*	+11	−43*	−42	−12	−39*	−37	−26	− 6	−25*	− 6	−28	−34	−38*
10	−20	−34*	−11	−39*	−17	+ 8	−17	−37*	−17	−55*	−43	−34	−15	−21*	−10	−45	−47	−53*
11	−24	−48	−52*	−63*	−29	+45	−22	−43*	−23	−46*	−46	−40	−21	−37*	−10	−47	−55	−61*
12	−24	−55*	−54	−48*	−25	+66	−41*	−38	−13	−38	−41*	−32	−13	−22*	−18	−43	−45*	−31
13	−10*	+68	+68	−15	−41*	−33	−35	−42*	− 3	−42*	−42	−31	0	−40*	−21	−27*	−24	−26
14	−11	−43*	+ 5	− 3	−41*	+17	−37*	−25	−22	−46*	−45	−27	−13*	− 3	−11	−23	−40	−38*
15	−34*	+25	+46	−32	−63*	−50	0	− 2*	+21	−24*	−16	−17	−38*	+31	−6	+13	−50*	−21
16	−22	−36*	+ 2	−27	−49*	−18	−24	−38*	−12	−34	−42*	−35	−44*	+36	−32	− 8	−26	−42*
17	−24*	−12	+ 8	−41	−59*	+ 3	−26	−30*	− 4	−39	−42*	−41	−18	+ 6	−24*	−22	−24	−54*
18	−25	−71*	−51	+14	−13*	+ 2	−10	−22*	+17	−25*	−22	−11	+ 2	+ 1	+ 6	−24	−33	−50*
No. asterisks	5	10	2	4	13	1	5	13	0	10	8	0	4	12	1	2	2	14
Average	−27	−27	−9	−23	−32	−10	−30	−36	0	−40	−38	−22	−19	−26	−14	−19	−35	−50

L = labor, M = material, C = fixed capital inputs, *maximum reduction.

Table 7.5. Structure of total cost of production (%).

Cost element	*1951*	*1955*	*1960*	*1965*	*1970*	*1975*	*1980*	*1980–1951*
Labor	44.7	44.5	46.5	47.3	46.2	47.4	53.3	119
Material	51.7	50.9	47.8	44.6	44.1	41.4	37.8	73
Fixed capital	3.6	4.6	5.7	8.1	9.7	11.2	8.9	245
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	

Table 7.6. 1951–1980 reductions in LR, MR, and CR (%) for 18 industries.

	LR		*MR*		*CR*		
Industries	*Value*	*Rank*	*Value*	*Rank*	*Value*	*Rank*	*Sum of ranks*
1. Agriculture	21	11	35	18	61	17	46
2. Mining	6	2	7	7	25	10	19
3. Construction	46	18	32	17	84	18	53
4. Food	26	16	22	15	54	15	46
5. Textile	23	14	5	4	17	4	22
6. Paper	13	6	5	5	25	9	20
7. Chemicals	7	3	3	3	17	3	9
8. Primary metals	14	8	3	1	11	2	11
9. Fabricated metal products	20	10	8	8	20	6	24
10. Nonelectric machinery	9	5	8	10	22	8	23
11. Electrical machinery	5	1	3	2	11	1	4
12. Transportation equipment	7	4	5	6	25	11	21
13. Other manufacturing	21	13	15	14	44	14	41
14. Trade	21	12	8	9	39	13	34
15. Finance, real estate	24	15	25	16	55	16	47
16. Transportation, communication	15	9	12	13	19	5	27
17. Public utilities	13	7	12	12	22	7	26
18. Services	44	17	10	11	28	12	40

and capital savings; 0.64, between labor and material savings; and 0.62, between labor and capital savings. This means that a growing saving of one element usually leads to the same tendency in other cost elements and coincidence in this process is much stronger than competition.

What are the reasons or determinants of the process when one type of technological progress is going ahead? Of course, from the long-term viewpoint the relative prices of the three factors (labor, materials, and fixed capital) as well as national availability of the resources are the important determinants of one or another type of technological progress.

In 1960–1980 labor prices doubled in Japan, prices of materials increased 2.5 times, and prices for capital goods increased 1.6 times. This led to higher savings of material input and lower savings of capital input (see *Table 7.5*). The first real capital-saving period took place in the second half of the 1970s. One reason was that prices of capital goods began to grow practically only in the 1970s and had been very stable before.

The supply of resources influenced these processes too. For example, material saving dominated during the periods of the postwar reconstruction of the Japanese economy when it lost the traditional sources of raw materials abroad and in 1970–1975 – the period of a first severe energy crisis. The 1965–1970 period was of a labor-saving type. At that time the rural labor resources were exhausted, the wage rate growth accelerated by 2 times, and the labor-saving type of technological progress became mostly preferable.

Coming back to the subject of our discussion, long waves, it is possible to draw the following conclusions from the analysis of different types of technological progress. Besides long-term oscillations of the technological progress as a whole, discussed above, the elements of cost reduction – labor/material/capital savings – have their own long-term harmonics dependent on both the common reasons and the specific ones.

Notes

[1] Showa 35 (40,45,50) nen sangyo renkanhyo. - Gyosei kanricho, Tokyo 1964 (1969, 1974, 1979).

[2] Input–Output Tables. Data Report (1) February 1984, Government of Japan.

[3] Uno, K. Input–Output Tables in Japan 1951–1980. March 1983, University of Tsukuba.

[4] Annual Report of National Accounts, 1982. Government of Japan.

[5] *Japan: Modern Monopolistic Capitalism.* Moscow, 1981 [in Russian].

[6] *Japan Statistical Yearbook*, 1967.

Reference

Uno, K. Long-Term Change in Input-Output Framework. Paper presented at the 8th International Conference on Input–Output Techniques, Sapporo, Japan, July/Augst 1986.

CHAPTER 8

Technology Diffusion in a Long-Wave Context: The Case of the Steel and Coal Industries

Arnulf Grübler

"La société, c'est l'imitation, et l'imitation c'est une espèce de somnambulisme."
G. Tarde, Les lois de l'imitation, 1895

8.1. Introduction

This paper is intended to illustrate, with the aid of selected examples from the steel and coal industries, the intimate relationship (or rather the *causality*) between technological change and *Wechsellagen* in long-term economic development. For this purpose periods of relatively stable homogeneous growth (characterized by regular diffusion patterns of new technologies, industries, and products into the market) and periods of turbulence and stagnation in economic activity (during which a number of leading technologies and industries, which were the main driving force of the previous upswing period, reach saturation in their market niches) will be differentiated. During these saturation phases new important technological innovations are often (although not always) introduced into the market.

It is my contention that during periods of depression leading technologies and products (in terms of their contribution to previous economic growth) are saturating. This saturation can be observed in either one or a combination of the following indicators: total output, market share, further cost reduction potential along the learning curve of a technology, and productivity indicators.

It is necessary at this point to emphasize that the argument presented here does not rely on a perfect *synchronization* of the saturation of existing and/or the emergence of new technologies as causal phenomena to explain long-term economic fluctuations, although empirical evidence suggests a certain degree of clustering in the appearance of innovations (see Chapter 4 by A. Kleinknecht). Rather, the argument is based on the high degree of overlap in the main growth periods (i.e., the period needed to increase from 10 percent to 90 percent of the ultimate market potential) of a number of key sectors and technologies, which would be sufficient to result in *Wechsellagen* of economic development.

In upswing periods one can observe regular growth in output, productivity, etc., as well as regular technological diffusion patterns, all of which convert to oscillatory behavior when reaching the saturation phase of market growth and technology diffusion. However, to describe these periods of saturation and turbulence (depression) not only in dimensions of changes in output and technological diffusion, economic and social indicators (such as prices and strike rates) will also be considered to characterize periods of stable growth and turbulence in long-term economic development.

In Chapter 9 N. Nakicenovic deals with the term technology in a wider context, in that primarily the evolution of transport infrastructures and of the energy system is analyzed. Infrastructures and energy constitute to some extent higher-level aggregates of a large number of underlying technologies. In this paper this view will be complemented (arriving at similar conclusions) by zooming from the macro level down to two specific sectors and discussing in detail changes in specific technologies, in factor inputs (energy and labor), and in major market outlets in individual countries. It will be shown that periods of turbulence and depression constitute important turning points in the *structural change* of the system with regard to all indicators mentioned above. Thus, the concept of the long wave constitutes a powerful paradigm through which long-term technological change can be systematized and better understood.

8.2. Multidimensional Indicators of Technological Change

Technological change is described along two dimensions: time and a number of indicators of technological change *per se*. With regard to the time dimension, a number of different indicators will be used to identify secular periods of stable economic growth and time periods of turbulence and stagnation. The identification of these time periods is especially important as there is evidence that all indicators of technological change *per se* show that the market saturation of major technologies and products, as well as significant turning points in the structure of the technological system (appearance of new technologies, changes in factor inputs, changes in market outlets), tend to cluster during these periods of turbulence and stagnation in economic growth.

8.2.1. Indicators to identify secular periods of growth and periods of stagnation

A large number of possible methods to describe long-term fluctuation in economic development have been advanced in analyzing empirical evidence in arguing pro or con the long-wave hypothesis.[1] For the purpose of this paper a number of different indicators will be used to increase the confidence about the exact timing of the various phases of the long wave. This approach is taken, because it is considered here that the complex phenomenon of economic fluctuations can be described only by a number of different (to an extent independent) indicators. Thus, the following indicators will be considered to describe secular movements of economic activity:

Deviations from long-term trends in output, prices, social indicators, etc. Typically, long time series are analyzed with various filters. In this paper deviations from a 50-year moving average are analyzed, these deviations are then smoothed by applying a 9-year moving average to them.

Discontinuities in long-term trends. Economic upswing periods are characterized by continuous growth in output, while downswing and depression periods are characterized by stagnation and oscillatory behavior of output figures. In this context it is interesting to note that an analysis of long-term production trends often reveals an overlap of consecutive growth pulses, with production figures oscillating between the boundaries defined by the two consecutive growth pulses.

Introduction and/or saturation of technologies.

Evolution of economic and social indicators. Under this hypothesis the periods of depression and structural change are characterized by *flares* in economic (e.g., prices) and social (e.g., strike rates) indicators. One theoretical explanation of these flares would consider the saturation of existing technologies, implying that further (real-term) cost reductions along the learning curve become infeasible (law of diminishing returns of technology improvement) and thus marginal production costs of resources, commodities, and products increase substantially. This increase in costs (and prices), along with the phenomenon of market saturation, could in turn have repercussions on social relations, making the settlement of working disputes a more difficult task and flares in working disputes more common.

8.2.2. Indicators of technological change

A number of different indicators are proposed in this paper to describe technological change:

Output. Introduction of new technologies allows an intensification of production and results in significant growth in total output. A typical example of this intensification is provided in the case of the introduction of the Bessemer process, allowing mass production of steel.

(Real) production costs. A drastic reduction in (real-term) production costs is a typical indicator for the introduction and growing importance of a new technology. The introduction of the puddling process and later of the Bessemer process, for instance, resulted in a significant decrease of the nominal and real-term steel production costs and prices.

Technological substitution. Here the share of various competing technologies in the production of total output is analyzed, using a logistic substitution model (Fisher and Pry, 1971, and Marchetti and Nakicenovic, 1979).[2] This model describes the emergence of new and the replacement of old technologies in terms of their fractional market share by a set of logistic substitution functions, where the emergence date is defined as the date a new technology captures one percent market share and the saturation phase of a technology is defined by a (non-logistic) transition function linking the logistic functions describing the introduction of a technology and its replacement by a more recent one.

Changes in factor inputs of production (in particular energy and labor). This analysis considers the long-term evolution of the intensity of factor inputs (or its inverse, in the form of energy efficiency or labor productivity) and the changes in the composition of factor inputs.

Changes in market outlets and products.

The above-listed indicators of technological change represent a considerable expansion of the descriptive framework for understanding technological change, which in the past has concentrated on indicators of output growth and factor inputs. The various indicators are, of course, closely interrelated; however, if one is able to discern their synchronous development (as we attempt to show in this paper), additional evidence supporting the hypothesis that technological change is a main driving force of long swings in economic activity can be gained.

8.3. Technological Change in the Steel Industry

The importance of steel in the industrial development in the second half of the nineteenth century as well as in the post–World War II period is widely recognized. In analyzing total world crude steel output in the period 1860 to 1985 (*Figure 8.1*) one can easily discern periods of regular growth from periods of stagnation and fluctuations in world crude steel output.

The regular growth in total output can be described by two consecutive growth pulses in the form of logistic functions with regular growth periods from 1880 to 1910 and 1950 to 1970. *Figure 8.2* shows the same two consecutive

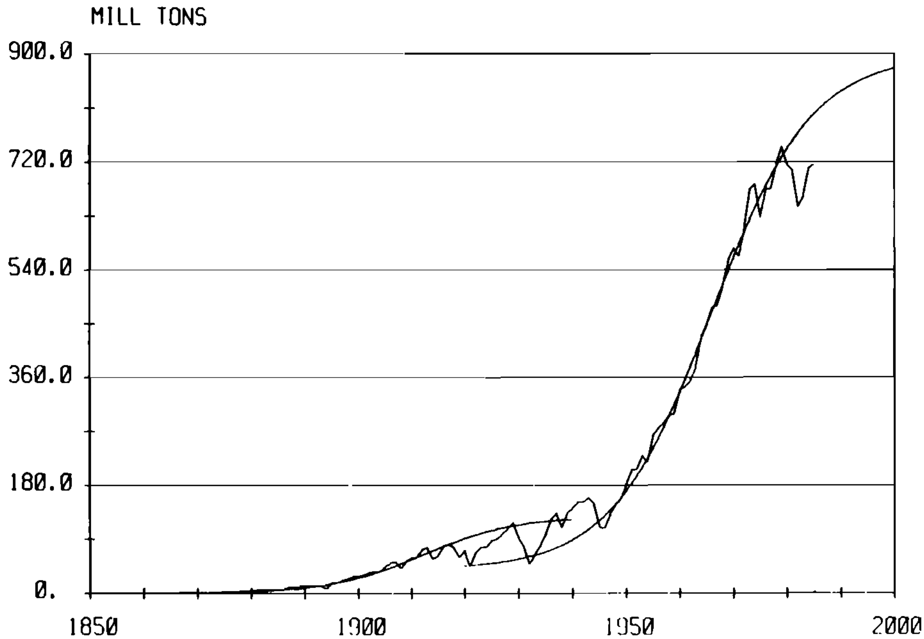


Figure 8.1. **World: growth pulses in raw steel production.**

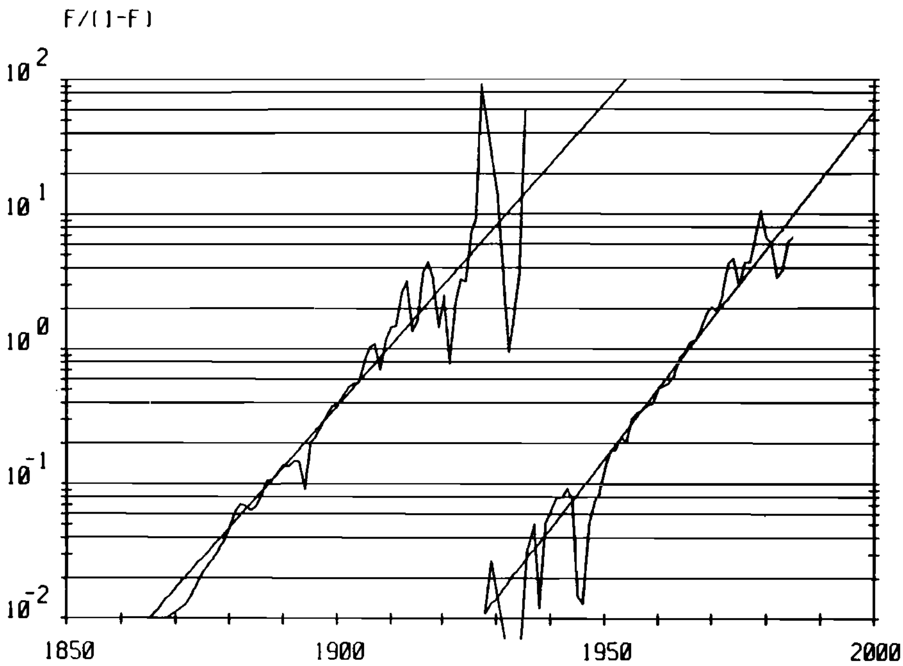


Figure 8.2. **World: growth pulses in raw steel production in $F/(1-F)$ transformation.**

pulses in the transformation $F/(1-F)$, where F represents the ratio of the world steel production in a particular year divided by the calculated saturation level of the logistic growth pulse (around 100 and 900 million tons, respectively) on a logarithmic scale. As a result, the deviations from the logistic growth trends (appearing as straight lines in *Figure 8.2*) can be more easily discerned. The two growth pulses overlap in the period 1920 to 1945, which is characterized by an oscillatory phase of turbulence, before a regular growth pattern is resumed after World War II. Since 1970, a similar deviation of total world steel production from the logistic growth pulse appears, indicating that the world steel industry is possibly facing a similar period of transition, as in the period 1920 to 1950. In *Figure 8.3* world per capita steel production is analyzed, and one can conclude that the picture is consistent with the situation illustrated in *Figures 8.1* and *8.2*.

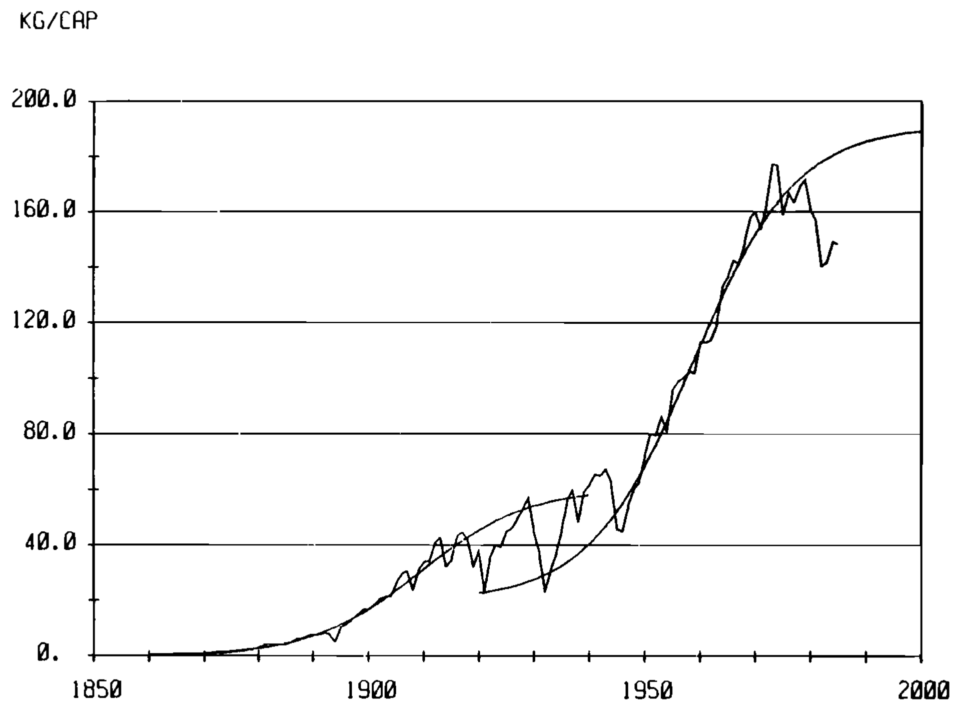


Figure 8.3. World: growth pulses in per capita raw steel production.

The reason for the rapid increase in total world steel production after 1870 can be explained by the introduction of the Bessemer steel production process in all steel-producing countries following its introduction in England in 1854. Through the Bessemer process a dramatic increase in mass production of steel became possible, compared with the production possibilities of the medieval bloomery process or the crucible and puddling processes introduced in the late eighteenth and early nineteenth centuries. The share of Bessemer steel in total steel production increased very rapidly in the period from 1860 to 1875, accounting for nearly 90 percent of total world steel production, before it was gradually replaced by the open hearth (Siemens-Martin) process.

There are two reasons for the rapid expansion of the steel production after World War II. First, in traditional steel-producing countries and regions (the UK, Europe, and the USA) new market outlets for steel products evolved, especially in the rapidly expanding automobile industry. Second, new principal producers, in particular Japan and "Newly Industrialized Countries" (NICs), appeared on the market. During this phase of expansion, we observed the rapid adoption of the basic oxygen blown furnace (BOF or LD) process and the replacement of the Bessemer and open hearth processes at the global level (see Roesch, 1979, and Chapter 9 by N. Nakicenovic).

During the same time period, one can observe a regular pattern of alternating periods of stable growth with periods of stagnation and fluctuations in output, as well as in the diffusion of new and the replacement of old steel production processes; one also observes the emergence of several new steel-producing countries and a change in the leadership of world steel production. *Figure 8.4* shows the share of principal steel-producing countries and regions (the UK, Europe including Eastern European countries, the USA and Canada, the USSR, Japan, and NICs) in total world steel production. *Figure 8.5* presents the same data transformed. The transform $F/(1\text{-}F)$ represents the share in total steel production of one country or region divided by the share of all remaining countries and is plotted on a logarithmic scale. This was done to make the data transformation underlying the logistic substitution model clearly visible.

Despite the turbulences induced by World War I, the 1930s, and World War II, one can nevertheless observe periods of significant structural change in the geographic location of steel production. These structural change periods are characterized by the emergence of new producing countries and the saturation and consequent loss of leadership (in terms of market share) of traditional producers. To better illustrate the dynamic shifts in the location of steel production, *Figure 8.5* is reconsidered by applying the multiple logistic substitution model to it (*Figure 8.6*). The transformation of the data from *Figure 8.5* compared with *Figure 8.4* makes the earlier (turbulent) phases of a particular country's market share more visible. Straight lines in *Figure 8.6* indicate logistic substitution paths of an increase and a decline in the market share. These model calculations are not intended as a method of forecast, but rather provide an instrument for data structuring that could better illustrate the dynamic geographic shifts in world steel production.

Again, periods of structural change can be distinguished from periods of regular growth and smooth transition. In periods of structural change (around 1870, 1920, and 1970 onward) new producing countries appear on the market and the long-term market share of the dominant producer(s) is reaching its saturation level in order to decline thereafter.

The period 1860–1880 is characterized by the loss of leadership in world steel production by the UK to Europe, which in turn saturated its share in total world steel production around 1870. During this time period two new producers emerged: North America (expanding very rapidly and overtaking the UK and Europe by 1890 in terms of market share) and the USSR.

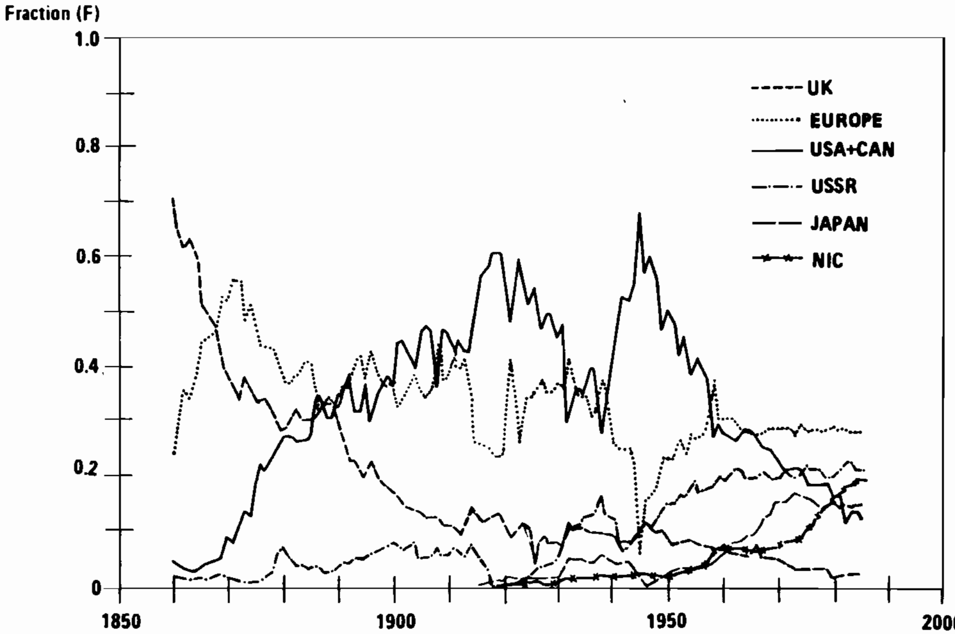


Figure 8.4. World: share of principal steel-producing countries in total crude steel production.

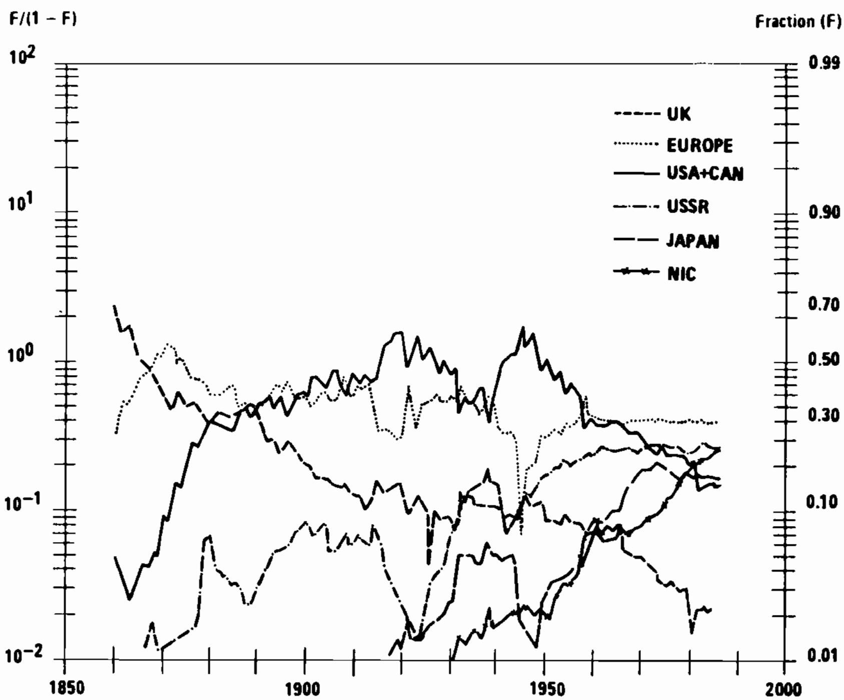


Figure 8.5. World: share of principal steel-producing countries in total crude steel production in transformation $F/(1-F)$. (Source: Nakicenovic, 1987.)

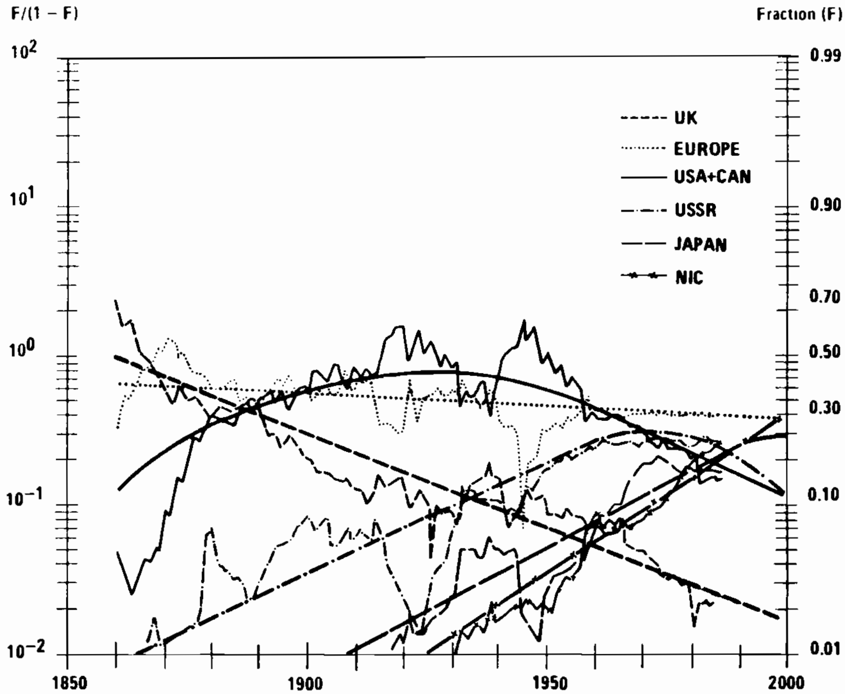


Figure 8.6. World: life cycle of principal steel-producing countries in terms of market share in total steel production $[F/(1-F)$ transformation]. (Source: Nakicenovic, 1987.)

In the period 1910 to 1930 a pattern similar to that of 50 years ago can be observed. The market share of the leading region in steel production (North America) reaches its saturation phase and new producers are entering the market (Japan and NICs). Currently the steel industry again appears to be in a phase of transition: the market share of the USSR in total world steel production is saturating (around 1975) and, based on the model estimates, one would expect a similar saturation for Japan by 1990. The rapid expansion of NICs is particularly at the expense of the North American market share, with Europe, the USSR, and ultimately Japan gradually loosing market share to these countries.

Despite the turbulences induced by wars and the crudeness of representation related to such a simple model, it is interesting to note that the logistic substitution model captures quite accurately the dynamic behavior, especially of the oldest and latest competitors. The UK has been loosing its market share since 1870 along the logistic path, which appears as a straight line in *Figure 8.6*. In a similar way the emergence of newly industrialized countries can quite accurately be captured by a logistic market share growth function, suggesting that the historically observed trend will continue into the future and will leave NICs as the only winners in the game.

After having established a framework of periods of change, turbulence, and saturation (around 1870, 1930, and since 1970) and of periods of smooth growth and transition at the macro level, the analysis will now zoom down to the national level to consider economic indicators and to analyze whether technological change coincides with the periods of turbulence and saturation identified above. The examples of a number of different countries point out, that the periods of fluctuation and depression coincide with the saturation of major technologies, in terms of both their market share and their performance (energy and labor efficiency). This indicates a close causal relationship between technological change and *Wechsellagen* of economic development.

Figure 8.7 shows the evolution of the per capita steel production in Germany (Federal Republic of Germany after 1945) and of the steel intensity per unit of GNP. One can observe a regular growth pattern of the two indicators up to 1910, with a stagnation period between 1870 and 1875. After 1910 a drastic slump in the steel intensity per capita and per unit of GNP can be observed as a result of the two world wars. Particularly visible is the effect of the depression of the 1930s. After 1945 the two indicators increase in line with the economic reconstruction of the Federal Republic of Germany and reach their saturation (for steel intensity around 1960 and for per capita steel production around 1970) and consequently started to decline, pointing to the fact that the material intensity of the German economy is decreasing.

Figure 8.8 presents the evolution of nominal and real-term (in 1913 Deutsche marks) steel prices since 1850. Although the data series available to date are still incomplete in order to analyze with regard to cyclic fluctuations (in particular, the time period of the 1880s, the period after World War I, and the time period of hyperinflation in the 1920s are lacking), they nevertheless provide a first-order picture on the evolution of real-term prices. Of particular interest is the dramatic decline of nominal and real steel prices between 1860 and 1875 as a result of the rapid diffusion of the Bessemer process. It is also interesting to note that the real-term steel prices have remained rather stable during the periods of stable economic growth (1890–1910 and 1950–1975).

On the basis of *Figures 8.7* and *8.8* one can conclude that the time periods with important changes and fluctuations to be analyzed in a technological change context are the periods from 1860 to 1875 (rapid fall of nominal and real steel prices), 1910 to 1935 (decrease of per capita steel production and steel intensity), and the period since 1970, where the shift away from a material (steel) intensive economy for the Federal Republic of Germany is particularly apparent. Let us now consider how technological change in terms of the principal steel production processes integrates into this picture.

Figure 8.9 shows the evolution of the share of different steel production processes in Germany. The evolution of the technological system can be described by a regular succession of growth, saturation, and replacement of different technologies in the production of steel. This regular pattern is quite remarkable in view of the territorial changes of Germany throughout this period and in particular between the German Reich before World War II and the Federal Republic of Germany thereafter.

The most important technological shifts in German steel production technologies start with the substitution of the medieval bloomery process by the puddling process, which saturates its market share in the period from 1855 to 1865

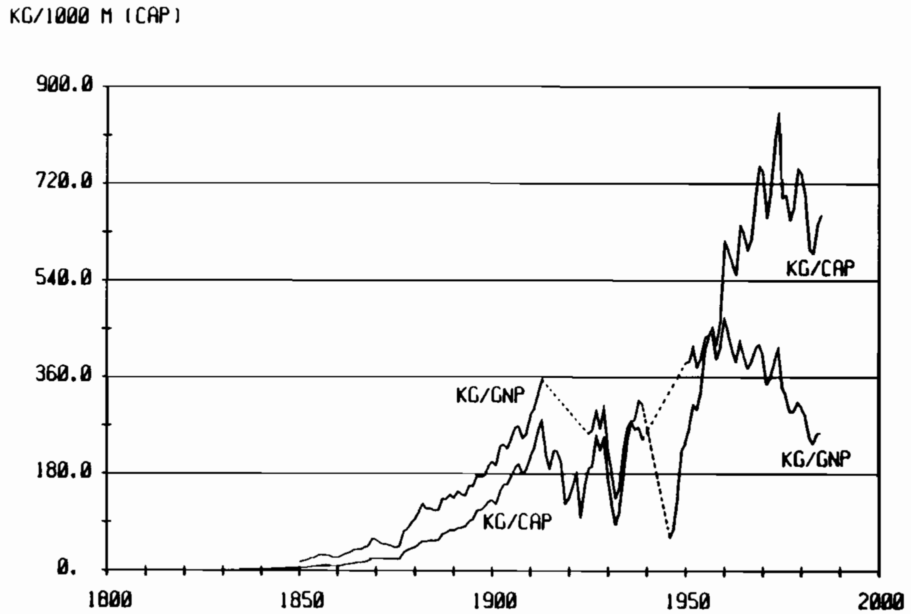


Figure 8.7. Germany (Federal Republic of Germany after 1945): steel production per capita and per unit of GNP.

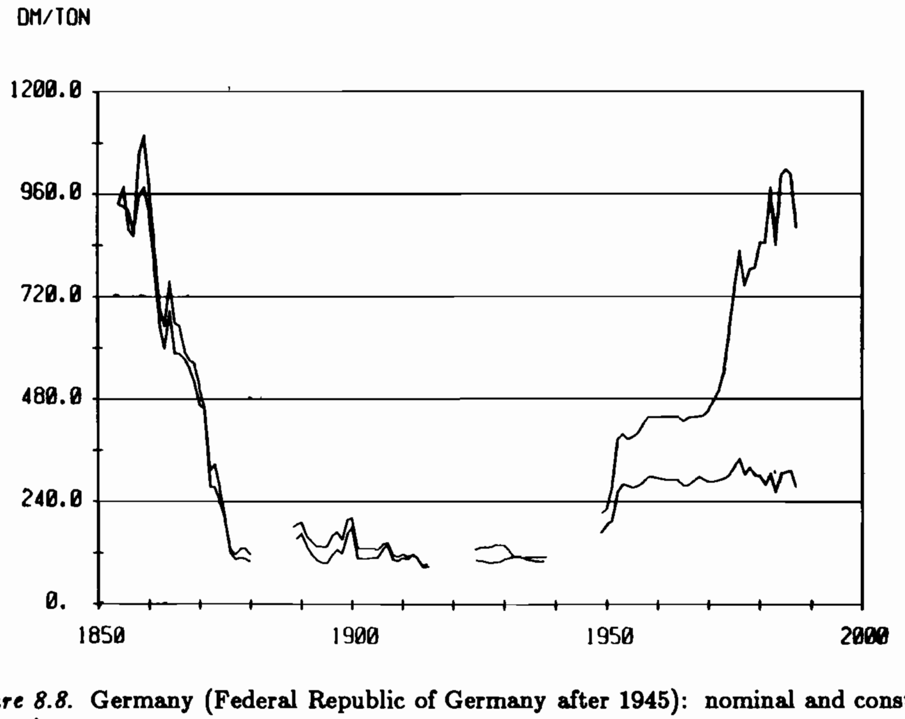


Figure 8.8. Germany (Federal Republic of Germany after 1945): nominal and constant steel prices.

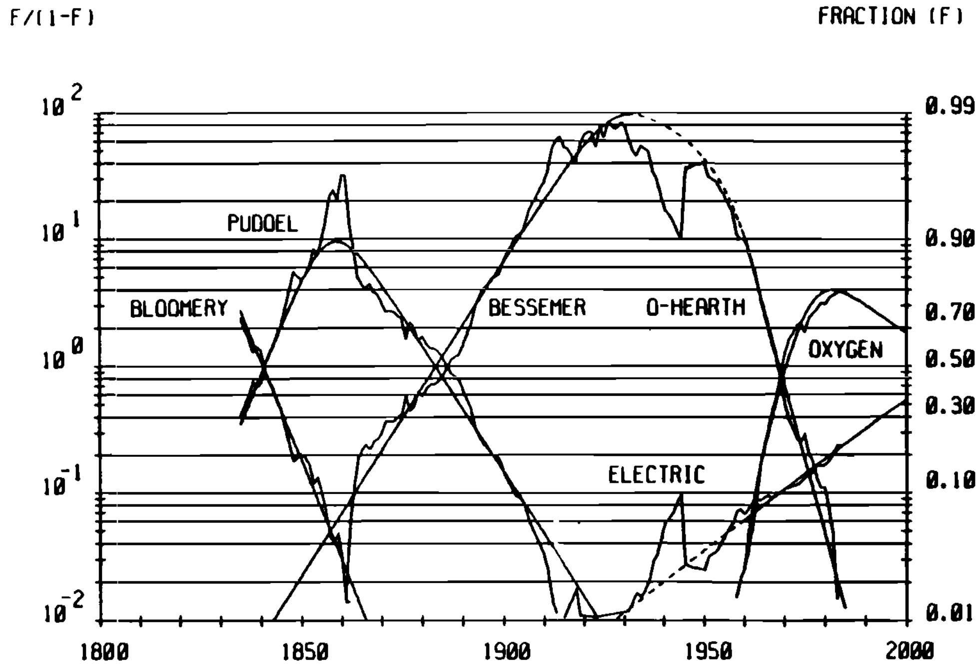


Figure 8.9. Germany (Federal Republic of Germany after 1945): substitution of steel production processes.

(period of introduction of the Bessemer process) and then is consequently replaced by the Bessemer and later by the open hearth processes.[3] The diffusion of the last two technologies follows a regular logistic path in the period from 1875 to around 1910, when the two technologies enter their phase of saturation. After 1910 the electric arc process emerges, which after some turbulence (the share of recycled scrap in total steel production increased dramatically during the wars) resumes a regular diffusion pattern in the second half of the twentieth century.

At the beginning of the 1960s one observes the appearance and fast adoption of another technology, i.e., the basic oxygen blown furnace (BOF or LD process). The emergence of a major technological innovation during a period of steady growth is in contrast to what we have observed in the past, when new technologies appeared during the phase of saturation of the predominant technology (coinciding to some extent with periods of economic fluctuation and depression). Still, it is interesting to note that, despite the appearance of the BOF technology falls outside the observed historical pattern, the diffusion of the BOF process enters its phase of saturation around 1973 and appears in the long term to be replaced by the electric arc process.

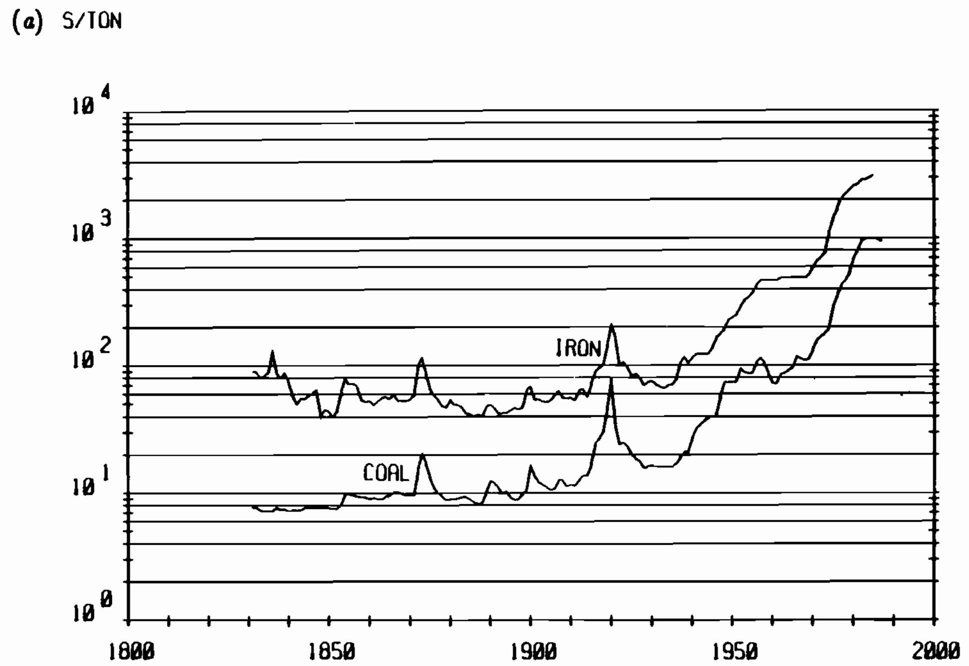
The case of steel production technologies in Germany supports thus the earlier statements in this paper that new technologies follow a regular diffusion pattern during phases of stable economic growth and approach their saturation stage during those periods defined as downswing and depression phases in the long-wave literature.

Figure 8.10 presents a time series for nominal and real prices of coal and pig iron (raw steel since 1965) for the United Kingdom. This example serves primarily to point at distinct price flares in the long-term evolution of coal and iron prices, whereas the cyclic nature of price fluctuations will not be further discussed on the basis of this particular example. *Figure 8.10* draws our attention in particular to the three distinct flares in nominal and real prices that occurred in 1835 (pig iron), after 1870 (coal and pig iron), and around 1920 (coal and pig iron). After 1940 nominal prices increase dramatically in line with inflation, and two flarings of real prices can be observed around 1950 (coal and pig iron) and around 1975 (raw steel) and 1980 (coal).

Under the assumption that price flares are indicators of periods of structural change in the system, we will now analyze the technology changes in crude steel production in the UK (*Figure 8.11*). Contrary to the example presented of Germany, the available data (Mitchell and Deane, 1971) for steel production according to processes go back only to 1871 and refer only to ingots and castings production. Thus, detailed statistics on a very important market segment of the UK steel industry in the nineteenth century – namely, the production of high-quality tool steel (Sheffield steel) by the crucible process – are missing, but it is clear that most of the steel production prior to the invention of the Bessemer process in the United Kingdom was in the form of crucible steel (Tweedale, 1986). Following the diffusion pattern in the USA, it took approximately 15 years for the Bessemer process to displace much of the market share of the crucible process, until the Bessemer process in turn was displaced by the open hearth process in the UK. Thus, it would be fair to assume the year 1870 (15 years after the introduction of the Bessemer process) as the approximate saturation period in the market share of the Bessemer technology. Since this question of the saturation date of the market share of the Bessemer process cannot in the absence of available statistics be resolved at present, the discussion will concentrate on the later period.

One can identify in the available data segment a regular diffusion pattern of open hearth technology versus the Bessemer process during the time period 1871 to 1920 when the open hearth process enters its phase of saturation (1920–1960). During World War I electric arc steel appears, and the market share of the various technologies shows no particularly regular pattern during the saturation phase of the open hearth and the emergence of electric arc steel. It is only with the appearance of the BOF technology around 1960 that a regular technological substitution pattern is resumed with the BOF replacing the open hearth process until it reaches its market saturation in the early 1970s. Along with this, one observes a steady decline in real steel prices which is only reversed the moment BOF technology reaches its saturation (and eventually its limit in cost reduction potential).

As a result, there exists evidence that the major turning points in the structural change of the system (especially with regard to the beginning of the saturation phase of the dominant technology and the emergence of new technologies) coincide with price flares, which are at least for 1870, 1920, and the present to



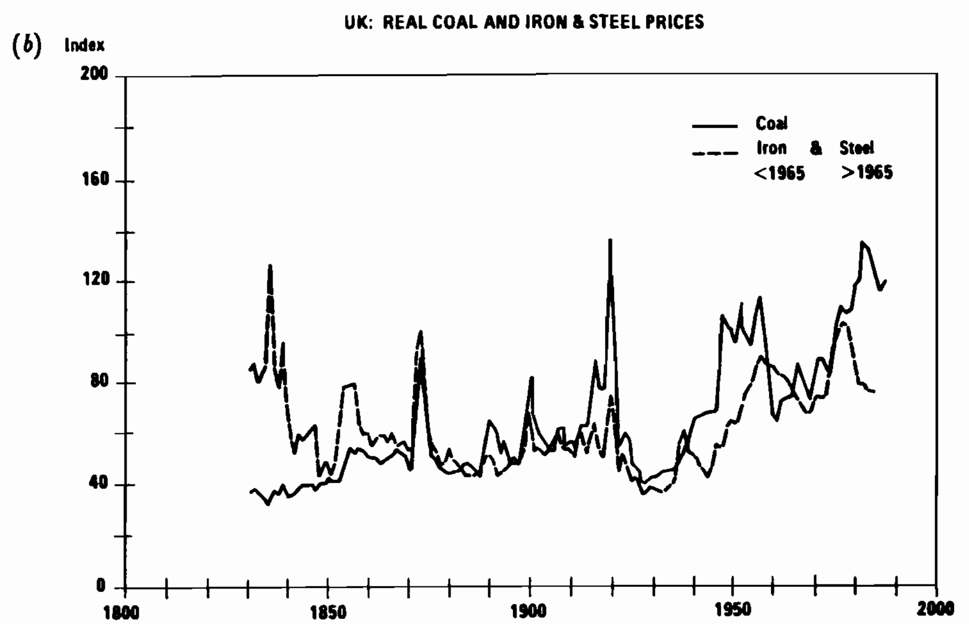


Figure 8.10. United Kingdom: nominal (*a*) and real (*b*) prices for coal, pig iron (1831–1964), and raw steel (1965–1985).

some extent consistent with phases of downswing or depression in a long-wave context.

Figure 8.12 shows the substitution of different steel production technologies in the USA. As already discussed in the examples of Germany and the UK, a strong correlation between major structural changes and turning points in technological change with the periods of downswing and depression in a long-wave context can be observed. This is illustrated by the saturation of the Bessemer process in the late 1870s, of the open hearth process at the end of the 1930s, and of the BOF process starting around 1975. During periods of stable economic growth (1890 to 1930 and 1955 to 1970), a regular pattern of technological diffusion and substitution in steel production technologies emerges from *Figure 8.12*.

Let us now consider whether these important dates of the technological change in the US steel industry can also be identified in another indicator of technological change, namely, in factor input (energy consumption).

Figure 8.13 presents the specific energy input per ton of raw steel production in the USA. Despite the fact that there is a considerable gap in the available statistics, one can still discern the historical tendencies both in energy input (specific energy consumption) and in the composition of energy input. The evolution of specific energy consumption [*Figure 8.13 (a)*] clearly exhibits a typical learning curve, where energy efficiency of steel making dramatically increases in the period between 1850 and 1920. After this time, energy efficiency still increases, however, not in the previously exponential manner but only in a linear trend, which is illustrated well in the available statistics from 1947 onward. This trend in the learning curve of energy requirements of steel production in the USA is confirmed also by European data (see Decker, 1976).

The year 1920 represents also an important date in the structural change of the *composition* of the energy inputs *Figure 8.13 (b)*. By that time the substitution process of bituminous coal for the traditional fuels charcoal and anthracite coal is completed and the system is almost entirely based on coal as fuel. It is interesting to note that prior to 1850 a similar substitution pattern of anthracite coal for charcoal can be observed, with the market share of anthracite reaching its saturation level between 1860 and 1870 with more than 50 percent of all energy requirements in steel production. This illustrates that it takes about 50 years between the shift in the market dominance of a particular fuel.

The situation characteristic for the beginning of this century, where practically all energy requirements were provided by a single source of primary energy, is in contrast to the situation after 1950. Here the energy is provided by four different energy sources: coal, oil (with a declining market share), natural gas, and electricity (with an increasing market share). Although the available statistics do not allow linkage between the two data series directly, one can by extrapolating structural change tendencies into the past observe that the important turning point in the introduction of new fuels into the steel sector and, thus the important structural change from a system relying on a single source of energy into a system with a diversified supply, occurred in the period from 1920 to 1930.

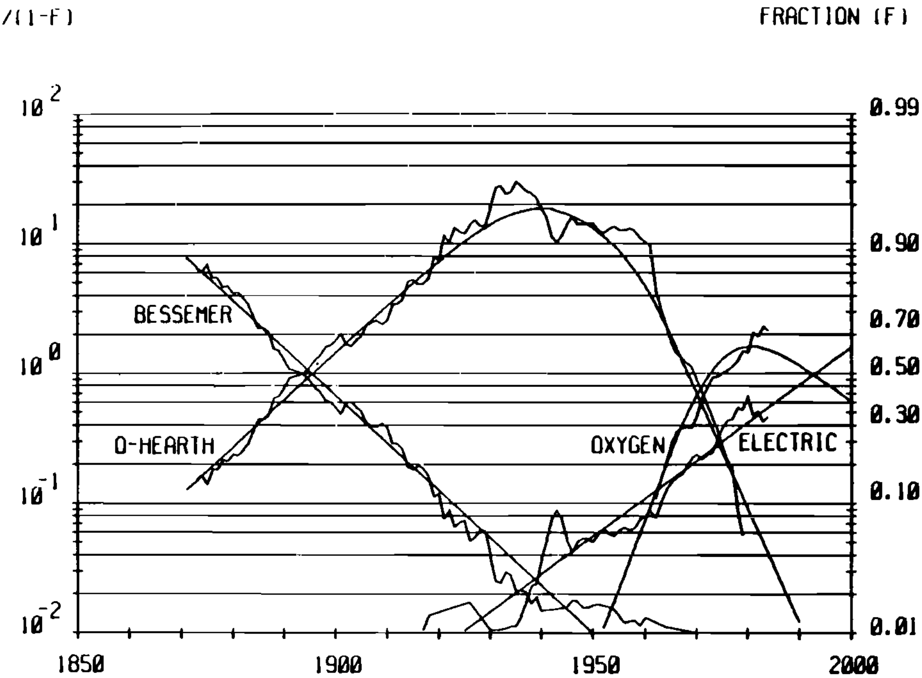


Figure 8.11. United Kingdom: substitution of steel production processes.

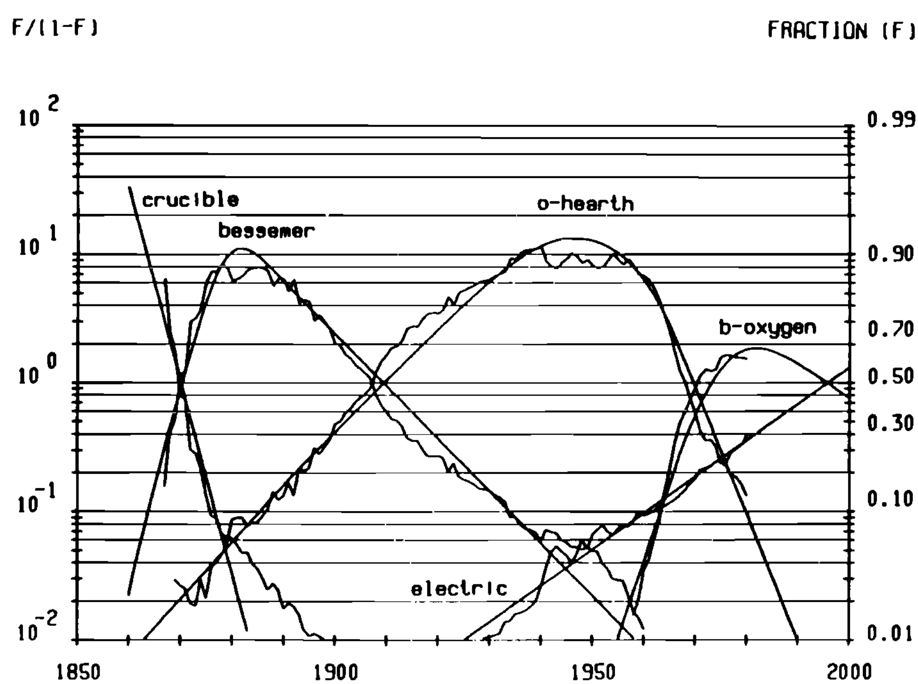
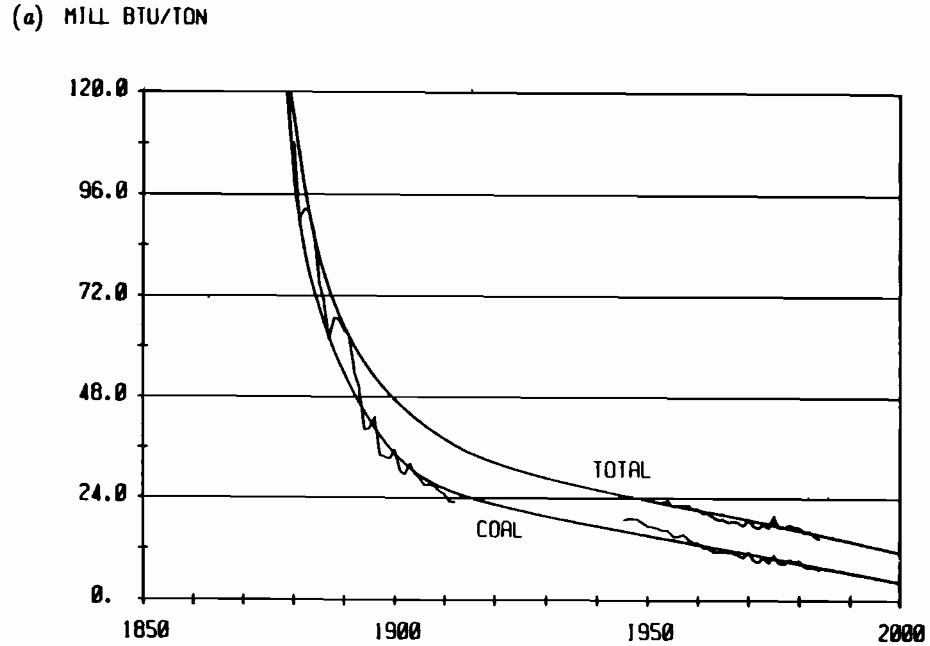


Figure 8.12. USA: substitution of steel production processes.



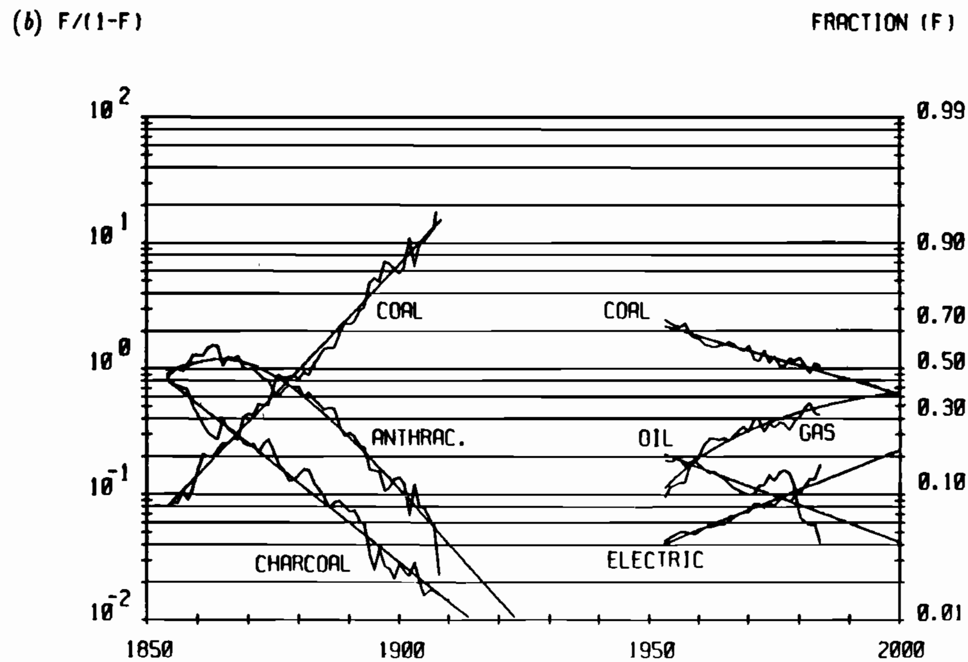


Figure 8.13. USA: specific energy consumption in raw steel production (*a*) and substitution of final energy carriers in the steel sector (*b*).

This structural change period coincides with another important date in the technological change of the system: the saturation of the open hearth process and the introduction of the electric arc process.

8.4. Technology Change in the Coal Industry (Mining and Use)

Figure 8.14 shows the evolution of world coal production from 1860 to the present. As in the case of world steel production regular periods of growth and fluctuation in world coal production can be identified. Long-term evolution can be described by a succession of two growth pulses that overlap during the period 1920 to 1945, where production oscillated between the boundaries defined by the two consecutive growth pulses.

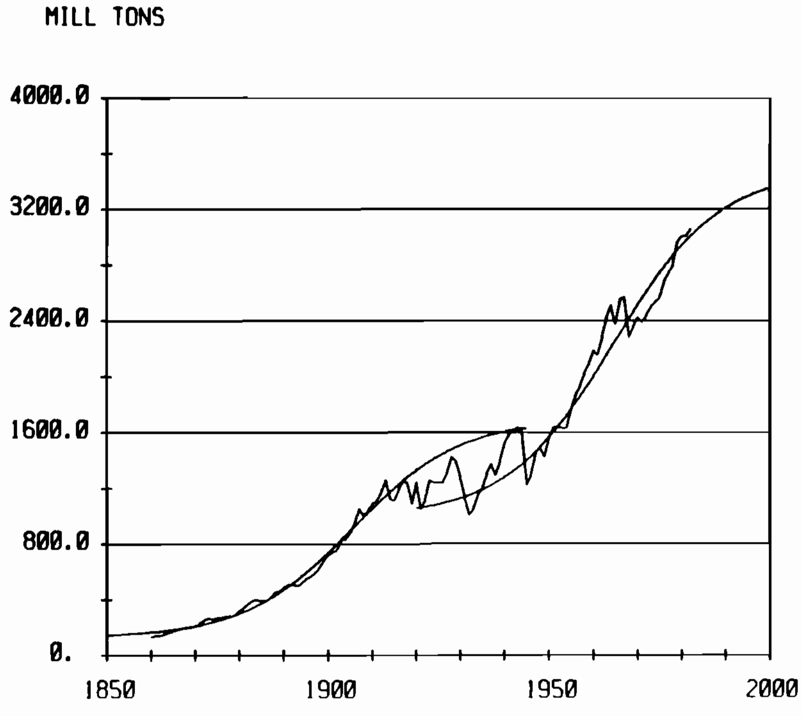


Figure 8.14. World: growth pulses in coal production.

A similar picture can be seen in the evolution of coal production and consumption figures for the UK, the country with the longest recorded history of coal production. *Figure 8.15* presents the history of coal production and consumption for the UK since 1700 and piecewise secular trends of homogeneous growth rates in the periods 1720 to 1760, 1760 to 1820, 1820 to 1870, 1870 to 1910, and then a less pronounced secular downward trend starting around 1960 becomes apparent. *Figure 8.16* illustrates these secular movements in terms of the deviation of coal consumption in the UK from a 50-year moving average.

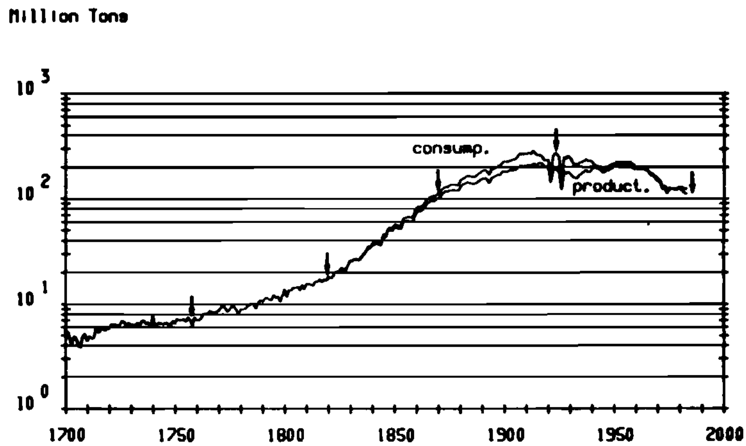


Figure 8.15. United Kingdom: coal production and consumption and periods of change in secular growth rates.

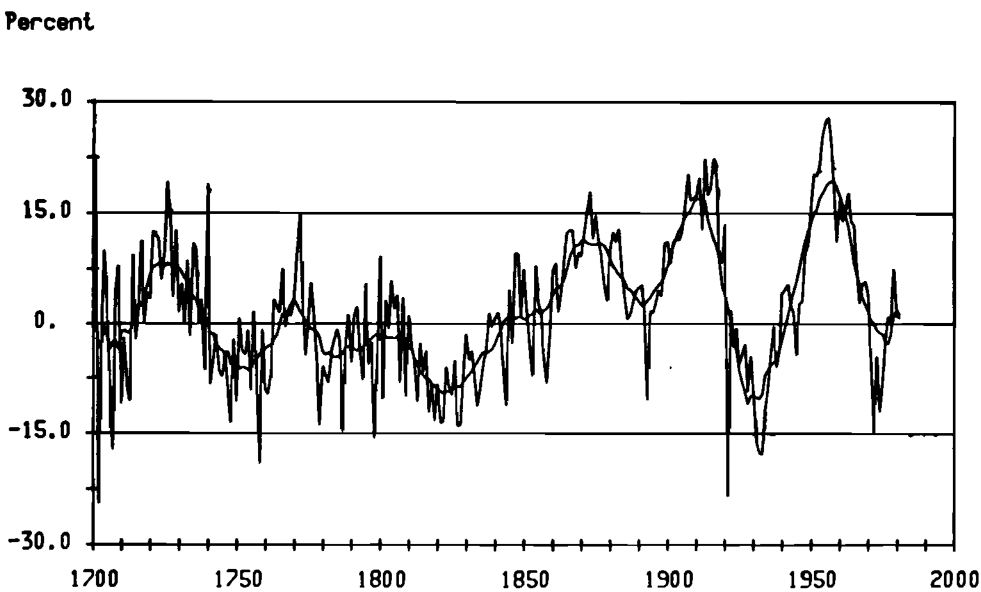


Figure 8.16. United Kingdom: long wave in coal consumption (percent deviation from 50-year moving average).

These deviations are smoothed by applying a 9-year moving average, with the smoothed secular deviations also being presented in *Figure 8.16.*

With regard to long-term price movements, which were already presented (in nominal and real terms) in *Figure 8.10* and which are further analyzed in *Figure 8.17* with regard to their deviations from a 50–year moving average

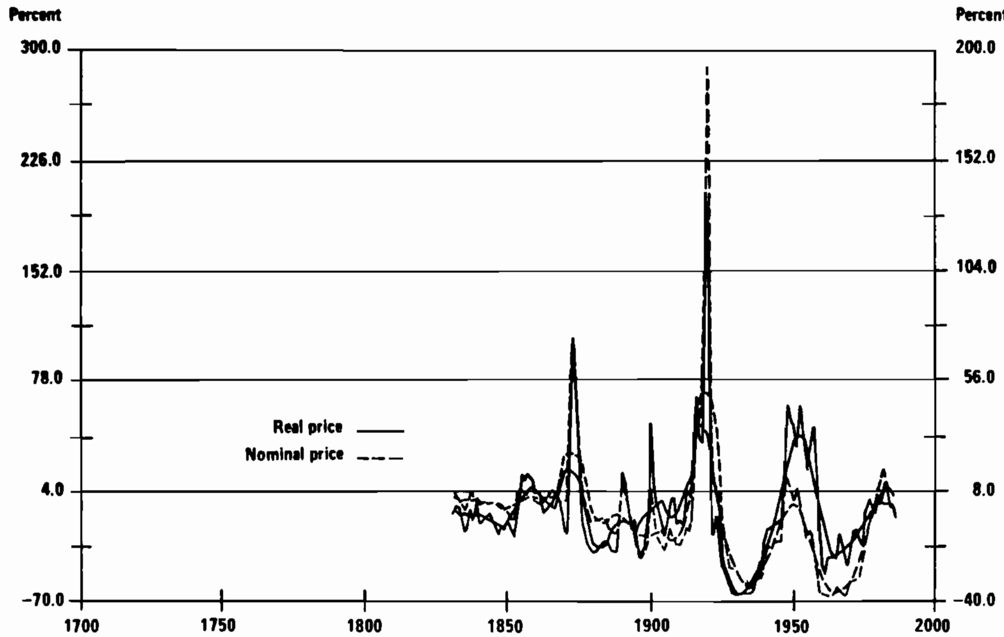


Figure 8.17. United Kingdom – long wave in nominal (left scale) and real coal prices (right scale), in percent deviation from 50–year moving average. (Source: Nakicenovic, 1987.)

(deviations are smoothed by a 9–year moving average), one observes an overlap of the secular movements in production and consumption and prices. With regard to peaks, the two secular movements harmonize for the period 1870, lag for 10 years in 1910 (consumption) and 1920 (price), and harmonize again during the 1950s. With regard to the troughs the two movements show a similar pattern only in the 1930s and a ten-year time lag between 1965 (price) and 1975 (consumption). Apparently, the frequency of the observed fluctuations has somewhat increased in the post–World War II period.

Before turning to a brief discussion on the long-term evolution of factor inputs (labor productivity) in coal production and its relation to technological change in the mining sector, another (social) indicator to determine periods of fluctuations and structural changes will be discussed. *Figure 8.18* presents the strike rates in the mining sector of the UK in the last 100 years. The original statistics refer to the number of workdays lost owing to working disputes; using this absolute measure the number of strike days in the 1980s amounts to 20 million lost workdays. This figure is not very high in comparison with the strikes in 1921 and 1926 with more than 70 million and nearly 150 million lost workdays, respectively. One has, however, to consider that the employed workforce in the coal-mining sector has drastically decreased between the 1920s and the present. Thus a relative measure, i.e., the number of strike days lost in relation to the total number of available workdays, was analyzed in *Figure 8.18*. As can be seen, the period 1912 to 1926 and the period since 1970 are particularly

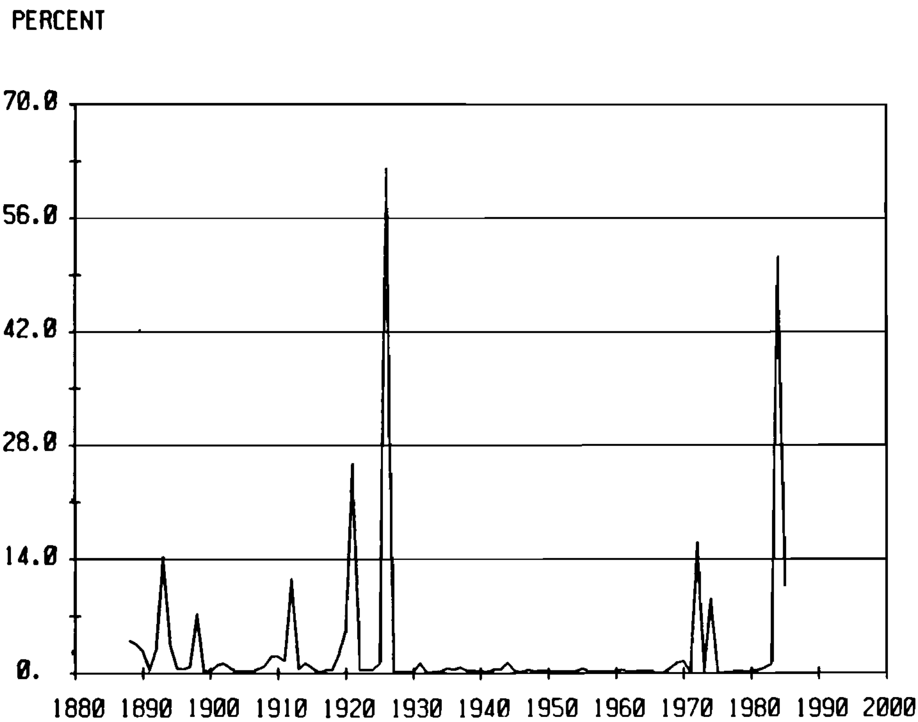


Figure 8.18. United Kingdom: percent of total workdays lost owing to strikes in mining and quarrying.

noticeable periods for intensive work disputes whereas during the period 1930 to 1970, the strike activities remained at a relatively low level. Incidentally the flares in strike activity occur at the same time as other indicators used so far to point out important turning points in the structure of the system.

Figure 8.19 shows the evolution of labor requirements (labor productivity) in UK coal mining. The productivity showed a downward trend over the period 1860 to 1925, with troughs in the productivity figures corresponding to the years of intense working disputes. This productivity decrease is primarily a consequence of resource depletion, as deeper deposits are mined, whereas the mining operation continues to be essentially a manual one.

Since 1925 labor productivity shows a reversal of the secular trend, in that 1925 productivity increased steadily (except for the wartime period 1939 to 1945) and reached a first level of saturation in 1970. This productivity increase is caused by the diffusion of mechanization technology (which will be analyzed later in more detail for the case of Germany), in terms of mechanization of winning operations (introduction of pneumatic picks and later shearers) and coal loading operations, as shown in *Figure 8.19.* Thus the massive introduction of pneumatic picks in the mid-1920s, followed by the increasing use of mechanized face operation schemes, is translated as an increase in the labor productivity, despite a continuation of the worsening in geologic conditions. Whereas the appearance

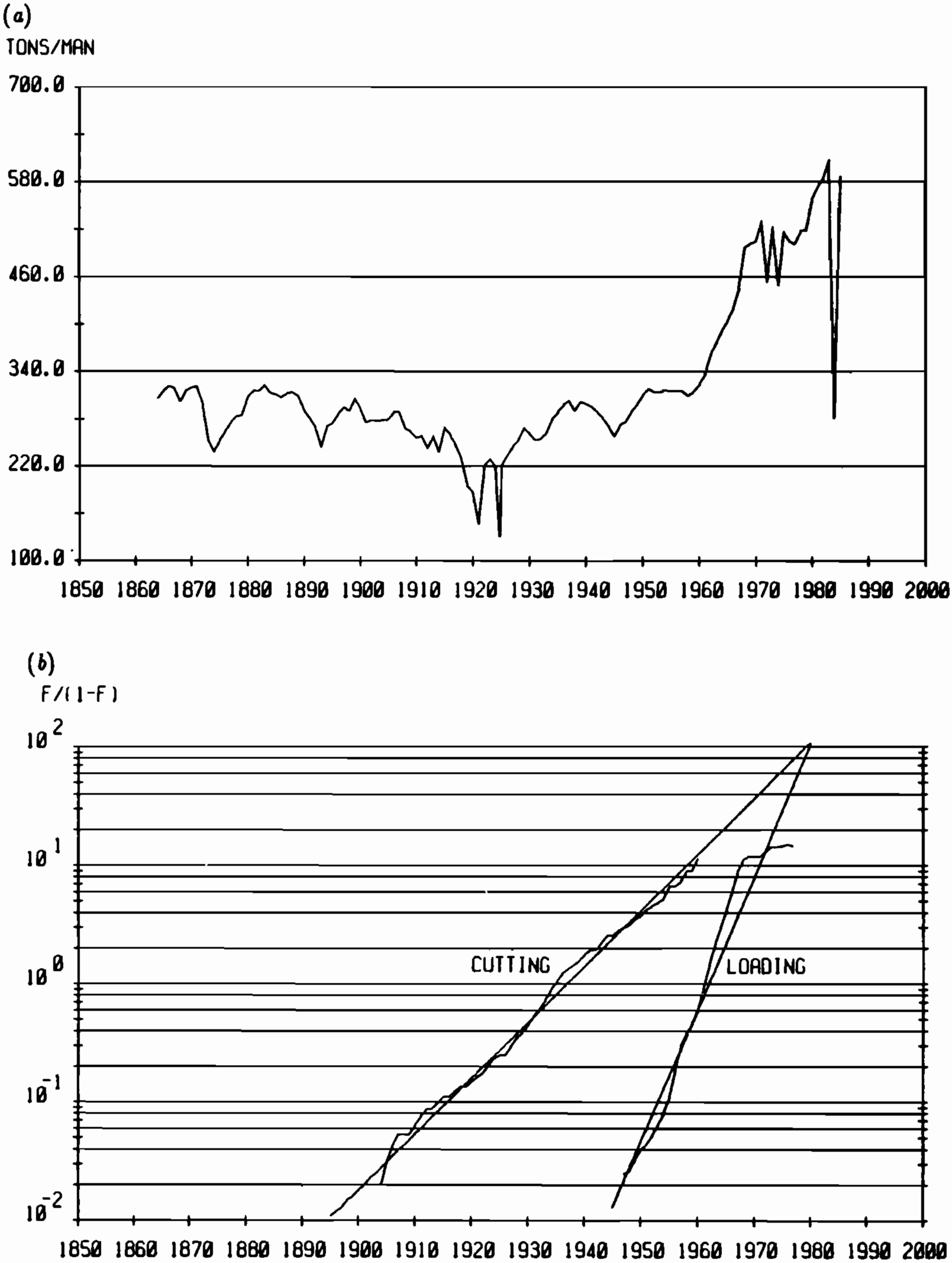


Figure 8.19. United Kingdom: labor productivity (*a*) and diffusion of mechanization (*b*) in coal mining.

of pneumatic picks and first shearers in the 1920s is consistent with earlier observations that important innovations tend to be introduced in periods of structural change and turbulence, one cannot follow a similar logic when analyzing the introduction of mechanical loading into the UK mining industry after 1945. However, it is interesting that the ultimate saturation of the two technology diffusion processes appears to be synchronized (1980). Also visible in *Figure 8.19* is the effect of the beginning saturation phase of mechanization (in passing more than 90 percent of total production) on productivity. This example would thus support the hypothesis of the importance of the simultaneous *saturation* of technology diffusion as a causal factor for economic depression periods.

Figure 8.20 analyzes the development of the underground labor productivity in coal mining of Germany (Federal Republic of Germany) similar to the previous analysis for the UK. Again two distinct phases in the development of labor productivity become apparent. The productivity, which remained practically constant up to the mid-1920s, is drastically increased as a consequence of the rapid diffusion of pneumatic picks as the first technology of mechanization in coal mining. This particular diffusion process saturates in the early 1930s when practically 100 percent of the total production is won with pneumatic picks. As the diffusion of the first mechanization in coal mining is saturating, the labor productivity (which doubled compared with the period when mining was essentially a manual process) is stagnating.

After World War II a regular pattern of diffusion of semi-mechanized mining technologies (mechanization of coal winning by shearers and of transport operations by conveyors) and later of completely mechanized mining schemes (mechanization of winning, transport, and development operations by means of self-advancing hydraulic roof supports) emerges, resulting in a doubling of the prewar productivity level. With the saturation of the diffusion of complete mechanization in the mid-1980s no further significant improvements in the labor productivity are to be expected (especially in view of ever worsening geologic conditions), unless a new technology (e.g., automation) is introduced.

The saturation of the diffusion of new technologies and its impact on factor input (labor productivity) and ultimately the costs of production can consequently be considered an important contribution to the stagnation of a particular sector. The simultaneous saturation in the technology diffusion of a number of key sectors in turn could provide an explicative model of long-term economic fluctuations.

At the end of this discussion of the technological changes in the coal sector, let us briefly consider the structural changes in the market outlets. The most important utilization of coal in the nineteenth century related to supplying the energy requirements of industry and households, the conversion of coal into secondary energy forms (in particular, coke and town gas), and fueling the transport sector (steam engines). Since 1900 one observes a shift in the market outlets especially for conversion into secondary energy carriers and in the transport sector.

Whereas the conversion to coke still constitutes an important market for coal, we observe a dramatic shift in the use of coal for the production of other secondary energy carriers. The production of town gas from coal has

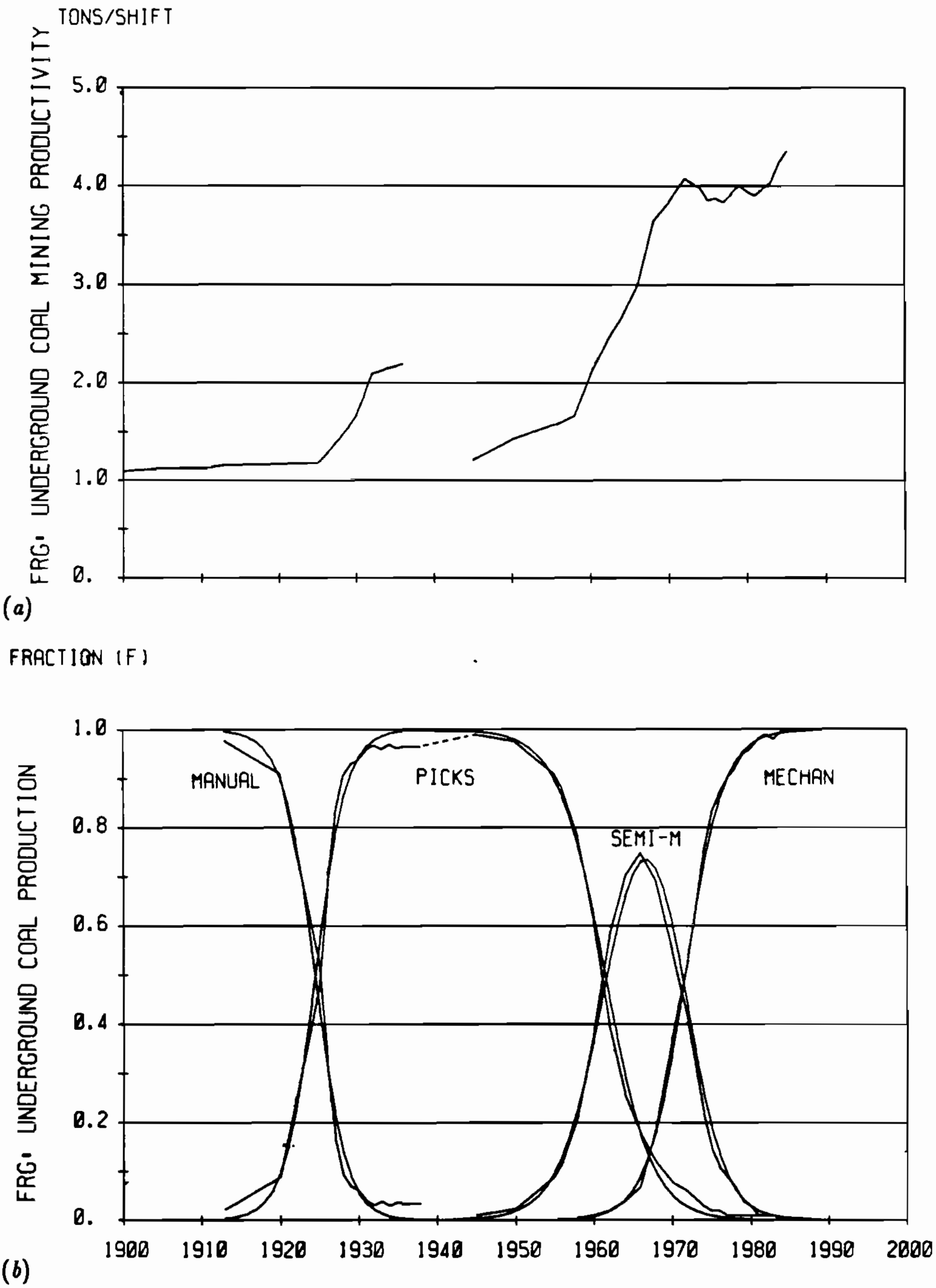


Figure 8.20. Germany (Federal Republic of Germany): underground labor productivity and substitution of mechanization technologies.

disappeared in all industrialized countries as a result of the substitution of gas for lighting purposes by electricity and by natural gas for other uses. Instead, the electricity-generating sector has in the period 1920 to present become the most important market for coal, both in absolute and in market share terms. Currently electricity production represents around 70 percent of total coal consumption in the UK and the Federal Republic of Germany, and around 80 percent in the USA.[4] The use of coal in the transport sector, which reached its peak around 1920 with more than 23 percent of total coal consumption in the USA and around 15 percent of total coal consumption in the UK, has entirely disappeared as a consequence of the substitution of the steam engine for the internal-combustion engine. This substitution occurs both in water transport (ships) and in ground transport (railways).

Figure 8.21 illustrates this substitution process of the coal-fired steam engine for the internal-combustion engine based on petroleum products in the case of the UK. In the figure the substitution process in terms of the total tonnage of ships registered in the UK fleet is presented, and a regular technological diffusion pattern emerges. Sail ships are replaced by steamships, which are introduced into the market in the mid-1820s and saturate their market share by 1920. Since that time period steamships are replaced by motor ships, resulting in the loss of the market outlet for coal for this application. For the USA a similar pattern can be observed.

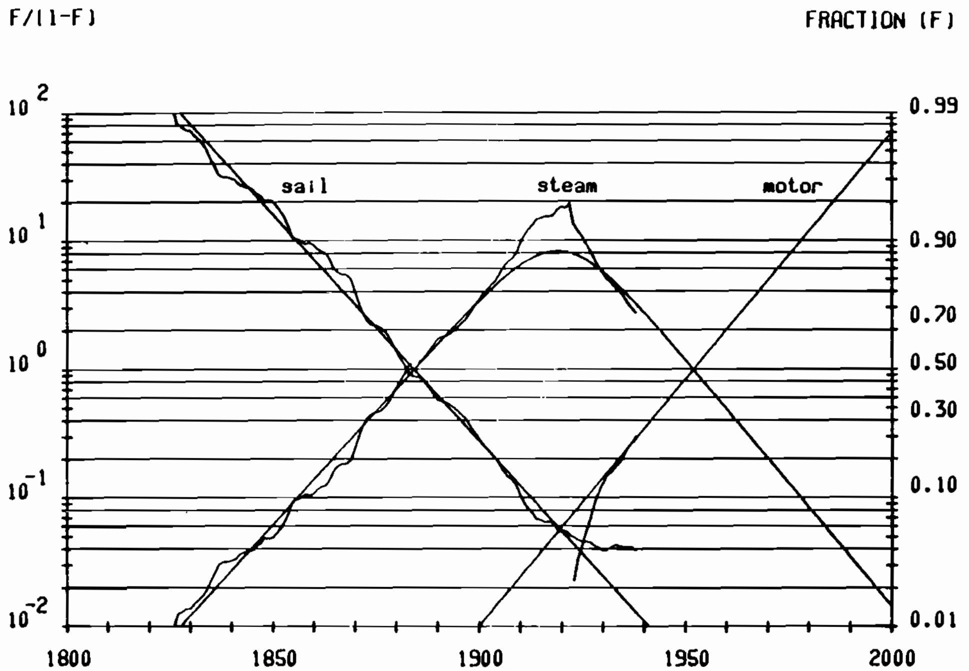


Figure 8.21. United Kingdom: substitution of propulsion in tonnage of ships registered. (Source: Nakicenovic, 1987a.)

In a similar way, coal-fired steam locomotives are replaced by diesel locomotives (illustrated in *Figure 8.22* for the USA). Since the end of the 1930s steam locomotives are replaced by diesel locomotives along a logistic substitution path. This substitution process was completed by the beginning of the 1960s and currently the entire locomotive fleet of the USA consists of diesel locomotives (contrary to Europe, electric locomotives have no practical importance in the USA). The result from these technological substitution processes in the transport sector initiated in the 1920s and 1930s was a complete disappearance of coal utilization in the transport sector and the loss of this important market outlet.

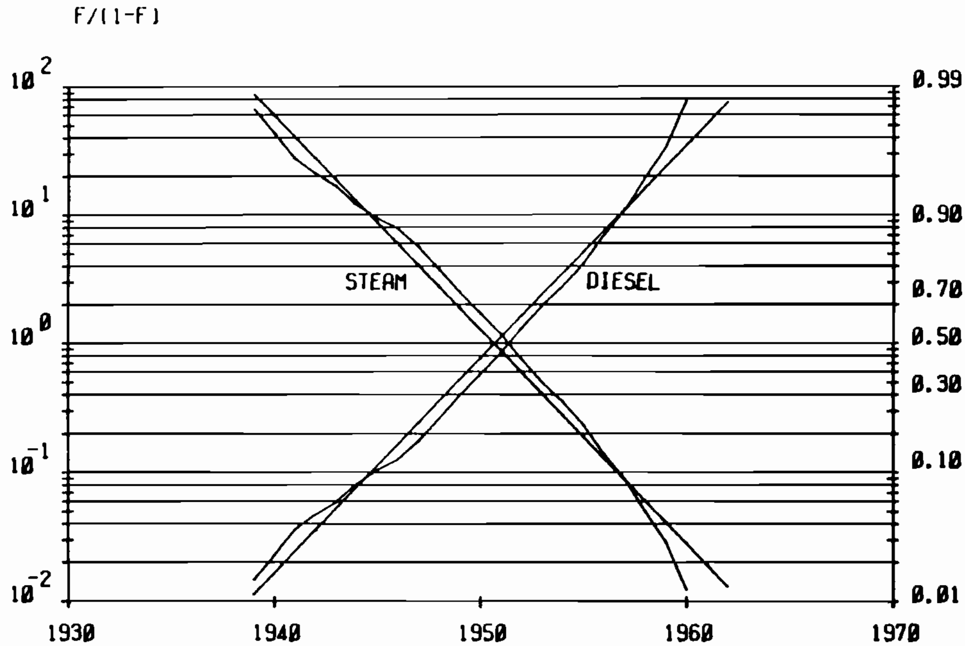


Figure 8.22. USA: substitution of steam for diesel locomotives. (Source: Nakicenovic, 1987b.)

8.5. Conclusion

The illustrative examples from the steel and coal sector presented in this paper provide evidence to support the hypothesis that technological change is closely related to *Wechsellagen* in long-term economic development.

A variety of different indicators has been used to illustrate that historically one can discern periods of turbulence from periods of stable economic development, where stable growth in output, regular technological diffusion patterns, and stable (real-term) prices are observed.

The regularity of growth and technological diffusion patterns during upswing phases seconds the contention that long-term economic fluctuations are to a large extent driven by the growth and the resulting saturation of technologies.[5] The regularity of technological diffusion processes (as captured quite accurately by the logistic substitution model), i.e., the *somnambulisme* in the structural change of a technological system during the phase of regular economic growth, is seen as a result of the *comparative advantage* (with regard to quality of output, economics of production, and response to societal needs) of new technologies over old ones.

Stable growth periods are followed by periods of stagnation and turbulence (depressions), during which time a high degree of *clustering of saturations* (saturation of market growth and subsequent oscillation of output levels and the saturation in market penetration and in efficiency improvements of technologies) can be observed. The *clustering of saturations* in the diffusion of different technologies is not necessarily related to the introduction date, i.e., technologies that are introduced at a later date tend to diffuse faster and reach saturation at a similar time period to technologies introduced much earlier into the system. During this *clustering of saturations* of leading technologies and markets/products, new technological innovations very often emerge and flares in nominal/real-term prices and social indicators can be observed.

Thus it is my contention that long-term fluctuations in economic activity are caused by the overlap of the main growth phases in output and technology diffusion *and* the subsequent *clustering of saturations* in markets, in output, and in the diffusion of technology, in particular, in those key sectors responsible for much of the previous economic growth.

The examples analyzed in this paper from the steel and coal industries are seconded by similar types of development patterns in the field of primary energy and transport infrastructures, which can be considered as (aggregate) "metatechnologies," and are analyzed by Nakicenovic in Chapter 9.

The *clustering of saturations* hypothesis thus draws on examples of the saturation of primary energy carriers (coal), infrastructures (railways) [6], steel technologies (open hearth processes), and mining technologies (first mechanization of coal mining with the introduction of pneumatic picks), among others, to explain, for instance, the depression of the 1930s. In that respect it is noticeable that the diffusion rates of these examples are widely different. In the case of the primary energy share of coal and the expansion of the railway network, the time constant is around 100 years; in the case of the diffusion of the open hearth process it is around 50 years; and in the case of coal mechanization the diffusion rate amounts to only 20 years. Still the dates of saturation of all these technologies fall within a relatively short time period, between 1920 to 1935.

It would be illusionary to expect from such a working hypothesis as the *clustering of saturations* a perfect synchronization with the historically observed depression dates. Consequently, similar saturation phenomena or price flares outside the 50-year pattern of Kondratieff cycles can be found. Examples for this saturation of technology diffusion and price flares *during* periods of economic growth include the saturation of the puddling process in Germany around 1860, the price flares and long-wave upswing in coal and steel prices in the UK around

1950, and the date of the trend shift in one steel intensity indicator for the Federal Republic of Germany (per capita steel production) around 1960. This, however, is not necessarily a contradiction to our working hypothesis, as the saturation dates should be considered in terms of a frequency distribution around the depression date of a Kondratieff cycle, rather than in terms of a rigid synchronization.

The regularity of growth rates, the regular technological diffusion patterns, and their overlap in time, as well as the coincidence of the (– in a larger sense – simultaneous) saturation in output and technology diffusion and the often-observed emergence of new technologies during these saturation phases would point to a Schumpeterian view of long-term economic development.

More phenomenological evidence has still to be assembled before the conclusions of this paper can decisively be confirmed and a causal model of economic *Wechsellagen* can be developed. The preliminary phenomenological cluster presented in this paper is intended to provide the basis for such a model, in which technology diffusion and growth, clustering of technology and market saturation, and a certain degree of clustering in the introduction of key technological innovations are considered as driving forces of economic development.

Notes

[1] For example, van Duijn, 1983; Freeman, 1983; Hartman and Wheeler, 1979; Kleinknecht, 1984; Kondratieff, 1926; Kuznets, 1930; Rostow, 1975; Schumpeter, 1939; Stewart, 1982.

[2] This particular model draws an analogy to the behavior of biological systems in using models of S-shaped growth patterns, like the logistic curve (Verhulst, 1838), used in biology to describe the growth and the dynamics of interaction of biological populations (Goell, Maitra, and Montroll, 1971; Lotka, 1910; Pearl, 1925). The theoretical basis for an application of such models in the technology and market area can be found in the concept of product life cycles as well as in the theories of the adoption and diffusion of new innovations, in both time (Rogers, 1962; Bass, 1969 and 1980) and space (Hägerstrand, 1967).

[3] Contrary to other countries (e.g., the UK and the USA) one cannot observe a displacement of the Bessemer process by the open hearth process in the case of Germany. The reason for the continuing importance of the Bessemer process can be found in the available iron ore resource base of continental Europe (e.g., the high phosphorous ores of the Lorraine) for which a variant of the Bessemer process, the basic Bessemer, or Thomas process was particularly suited.

[4] For a detailed overview of the historical changes of the market outlets of coal produced in the Rhur basin of the Federal Republic of Germany see Wiel, (1970).

[5] At this point it is necessary to stress the synergistic effects of the growth of a *number* of key sectors in explaining major upswings in economic development. Thus, a number of sectors would have to be studied under the "leading sector" hypothesis (Rostow, 1962; Hirschmann, 1967) in order to explain the extent of economic growth during certain time periods of intense economic development. See, for instance, Holtfrerich, 1973, and Fremdling, 1975, for an analysis of coal and railways in the economic development of Germany in the second half of the nineteenth century.

[6] For additional evidence see also Mothes, 1950.

References

Bass, F.M., 1969, A New Product Growth Model for Consumer Durables, *Management Science* **15**:215–227.

Bass, F.M., 1980, The Relationship Between Diffusion Rates, Experience Curves and Demand Elasticities for Consumer Durable Technological Innovations, *Journal of Business* **53**:51–67.

Decker, A., 1976, *Energy Accounting for Steel*, 9th International TNO Conference: The Energy Accounting of Materials, Products, Processes and Services, TNO, Applied Scientific Research Organization, Delft, The Netherlands.

van Duijn, J., 1983, *The Long Wave in Economic Life*, George Allan and Unwin, London, UK.

Fisher, J.C. and Pry, R.H., 1971, A Simple Substitution Model of Technological Change, *Technological Forecasting and Social Change* **3**:75–88.

Freeman, C., ed., 1983, *Long Waves in the World Economy*, Butterworths, London, UK.

Fremdling, R., 1975, *Eisenbahnen und deutsches Wirtschaftswachstum 1840–1879.* Gesellschaft für westfälische Wirtschaftsgeschichte, Dortmund, Germany, F.R.

Goell, N.S., Maitra, S.C., and Montroll, E.W., 1971, On the Volterra and Other Nonlinear Models of Interacting Populations, *Rev. Mod. Phys.* **43**:231–276.

Hägerstrand, T., 1967, *Innovation Diffusion as a Spatial Process*, University of Chicago Press, Chicago, IL, USA.

Hartman, R. and Wheeler, D., 1979, Schumpeterian Waves of Innovation and Infrastructure Development in Great Britain and the United States: The Kondratieff Cycle Revisited, *Research in Economic History* **4**:37–85.

Hirschmann, A.O., 1967, *Die Strategie der wirtschaftlichen Entwicklung*, Stuttgart, Germany, F.R.

Holtfrerich, C.L., 1973, *Quantitative Wirtschaftsgeschichte des Ruhrkohlenbergbaus im 19. Jahrhundert*, Gesellschaft für westfälische Wirtschaftsgeschichte, Dortmund, Germany, F.R.

Kleinknecht, A., 1984, *Innovation Patterns in Crisis and Prosperity: Schumpeter's Long Cycle Reconsidered*, Vrije Universiteit, Amsterdam, The Netherlands.

Kondratieff, N.D., 1926, Die langen Wellen in der Konjunktur, *Archiv für Sozialwissenschaft und Sozialpolitik* **56**(3):573–609.

Kuznets, S., 1930, *Secular Movements in Production and Prices*, A. M. Kelly, Fairfield, NJ, USA.

Lotka, A.J., 1910, *Journal of Physical Chemistry* **14**:271.

Marchetti, C. and Nakicenovic, N., 1979, *The Dynamics of Energy Systems and the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Mitchell, B.R. and Deane, P., 1971, *Abstract of British Historical Statistics*, Cambridge University Press, Cambridge, UK.

Mothes, F., 1950, *Das Wachstum der Eisenbahnen, Zeitschrift für Ökonometrie* (1950), **1**:85–104.

Nakicenovic, N., 1987a, private communication

Nakicenovic, N., 1987b, *Transportation and Energy Systems in the US*, WP-87-01. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Pearl, R., 1925, *The Biology of Population Growth*, Knopf, New York, NY, USA.

Roesch, K., 1979, *3500 Jahre Stahl, Abhandlungen und Berichte*, (1979) Vol. 2, Deutsches Museum, Munich, Germany, F.R.

Rogers, E.M., 1962, *Diffusion of Innovation*, Free Press, New York, NY, USA.

Rostow, W.W., 1962, *The Process of Economic Growth*, New York, NY, USA.

Rostow, W.W., 1975, Kondratieff, Schumpeter and Kuznets: Trend Periods Revisited, *Journal of Economic History* 35:719–753.

Schumpeter, J.A., 1939, *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*, Vols. 1 and 2, McGraw-Hill, New York, NY, USA.

Stewart, H.B., 1982, *Technology Innovation and Business Growth*, Nutveco, San Diego, CA, USA.

Tarde, G., 1895, *Les lois de l'imitation, étude sociologique*, 2nd Ed., Germer Baillière et Cie., Paris, France.

Tweedale, G., 1986, Metallurgy and Technological Change: A Case Study of Sheffield Specialty Steel and America, 1830–1930, *Technology and Culture* 27(2):189–222.

Verhulst, P.-F., 1838, Notice sur la Loi que la Popluation suit dans son Accroissement, *Correspondence Mathématique et Physique* 10:113–121.

Wiel, P., 1970, *Wirtschaftsgeschichte des Ruhrgebietes Tatsachen und Zahlen*, Siedlungsverband Ruhrkohlenbezirk, Essen, Germany, F.R.

CHAPTER 9

Dynamics of Change and Long Waves

Nebojsa Nakicenovic

"Evolution is the result of a sequence of replacements." Elliott W. Montroll, 1978

9.1. Introduction

Much of the long-wave debate centers around the existence of empirical evidence and possible explanations of the phenomenon, but even those who claim that such long-term fluctuations in economic life are real argue whether the waves and their timing are regular and to what extent various events are synchronized. We will argue that the long-wave phenomenon is one particular way of describing technological, economic, and social change and development. We show that this evolutionary development process is the result of a sequence of replacements of one technique (technology, artifact, product, practice, organization, tradition, or idea) by another.

Long waves emerge because a number of important replacements take place simultaneously, leading to a prolonged phase of economic growth. Likewise, these replacements and growth processes tend to saturate during relatively short periods resulting in prolonged recession, economic restructuring, and technological and social change. It is during these periods that innovations cluster and eventually lead to a new "wave" of replacements and growth. This hypothesis does not require a high degree of synchronization or "focusing" for either the innovation clusters or the saturation clusters. Instead, it is sufficient that a number of important replacements take place simultaneously resulting in the periods of prolonged growth.

Denoting the time interval that it takes to replace from 10% to 90% of old practice or technique by ΔT, all that is required is that the ΔTs overlap for a number of important substitutions. This requirement implies a certain degree of clustering of saturations and innovations, but the clusters need not be sharply focused during the troughs of the long wave. In fact, they are initiated during the end of the growth phase, which is followed by a period of turbulence when the first important replacements start saturating, and last well into the beginning of the upswing phase marked by a new pulse of replacements. This view of the long wave is also consistent with Kleinknecht's observation that innovation clusters can be observed around the troughs. However, they do not appear to be as strongly focused as originally maintained by Mensch (see Mensch, 1979; Kleinknecht, 1987; Chapter 4). We will show that important technologies saturate around the troughs and, more importantly, that the time intervals covered by their ΔTs encompass prolonged periods of growth and expansion. Long-wave pulsations (flares) in prices will be used as an indicator of the various phases, especially the sequence of growth and recession phases.

On the basis of our assumptions, the dynamics of change and economic growth can be decomposed into a sequence of replacements that generate periods of growth followed by recession. These are marked by turbulence caused by the lagged saturation of the replacement processes. Innovations eventually lead to a new phase of replacements and thus growth. The innovations should not be viewed in the narrow technical sense, because they also include new social and organizational forms and ideas and creation of new practices. At the aggregate level, e.g., physical output or gross national product (GNP), the indicators portray long-term increases interrupted by phases of turbulence. In other words, the aggregate indicators look like step functions where each step is a plateau with large fluctuations, and the periods of growth between the plateaux are S-shaped curves. Due to the inaccuracy of the data and the inherent difficulty of identifying exact resolutions of different phases in the long waves, we will use a number of different indicators simultaneously to describe technological and economic changes.

We will use both physical indicators, such as energy consumption or length of transport infrastructures, and monetary indicators, such as prices or output at given prices. A large number of indicators give a higher precision in identifying growth pulses, replacements of old by new, and various phases in the long wave. This higher precision is reached in the statistical sense and same way as a large number of synchronized clocks will give a better time measurement than a single one does (official time is in fact measured by an average). Thus, the multidimensional approach increases the accuracy of the results in spite of often insufficient data quality, especially for the records from last century or before.

Another reason for analyzing a number of different indicators for the same time periods and same processes of change (e.g., energy, transport, prices) is that not all can be described by the same secular patterns. For example, prices often portray fluctuations, some of them with relatively long periods, but rarely increasing or decreasing secular trends of more than a few decades. Growth and senescence of technologies or consumption levels, on the other hand, often have consistent secular trends with very long duration (compared with the long wave). Usually, the secular trend of growth and senescence processes can be described by S-shaped (often logistic) curves. However, not all can be described by simple (or single) S-shaped functions. Sometimes more complex patterns are observed, which are often described by envelopes that can be decomposed into a number of S-shaped growth or senescence phases.

Two typical cases are successive growth pulses with intervening saturation and a period of change, and simultaneous substitution of competing technologies. We argue that (1) the population, production, or performance of inanimate (man-made) objects or systems can be described by successive growth pulses that often have an S-shape, the same shape encountered in the growth of populations, "organisms," etc., as originally described by Verhulst (1844) and Pearl (1924); and (2) these growth pulses can be decomposed into a sequence of replacements, originally exploited by Fisher and Pry (1971), as the appropriate model for the dynamics of industrial replacements, and later extended to simultaneous replacement of more than two competing technologies (Marchetti and Nakicenovic, 1979). In the first case, successive S-shaped pulses usually represent, for example, the growth of energy consumption or successive improvements in performance, such as aircraft speed records. Here the first pulse is associated with an old technology, the piston engine, and the second with the new technology, the jet engine (see Nakicenovic, 1987a). In the second case, simultaneous substitution of competing technologies is usually described by increasing market shares of new technologies and decreasing market shares of old technologies, such as the replacement of sails by steam and later by motors.

Thus, we will describe the dynamics of change and the long waves with a number of different indicators (that is, by vector rather than by scalar measurements). Usually an asymmetrical view is offered, using either price or physical (technological) indicators as explanatory variables. Our description is not frozen into one or the other view, but rather we offer a symmetrical description by considering many dimensions of these dynamic processes.

First, we will illustrate the waves (flares) in prices. Second, we will show that the expansion of the most important means of transportation, starting with canals and ending with aircraft, can be seen as a replacement process where the substitution of older for newer technologies overlaps the growth phases in the long wave. This analysis starts, somewhat unconventionally, with the youngest technologies (aircraft and airways) and ends with the oldest (canals and waterways). Finally, we will show that the evolution of the energy systems and steel production also follows a similar pattern where important replacement processes take place during the growth phases of the long wave. In addition, we will show that total global production of energy and steel evolved through two growth pulses that are related to the last two long waves. Throughout the text we will use the price flares as a clock for the long waves.

9.2. Price Waves

The regularity of fluctuations in price data was the phenomenon that first stimulated Kondratieff and his predecessors to postulate the existence of long waves in economic life. These waves are most pronounced in the wholesale price indices for all commodities in the United States, but they can be observed in the price indices of other industrialized countries including the United Kingdom. *Figure*

9.1 shows the wholesale price index in the United Kingdom from 1560 to 1982 and *Figure 9.2* in the United States from 1800 to 1982. In both countries prices appear to be stationary with long fluctuations almost over the whole period. Only after the 1940s can a pronounced inflationary trend be observed that had a magnitude greater than any other previous fluctuation. In the United States prices reached pronounced peaks around the years 1780, 1815, 1865, and 1920,

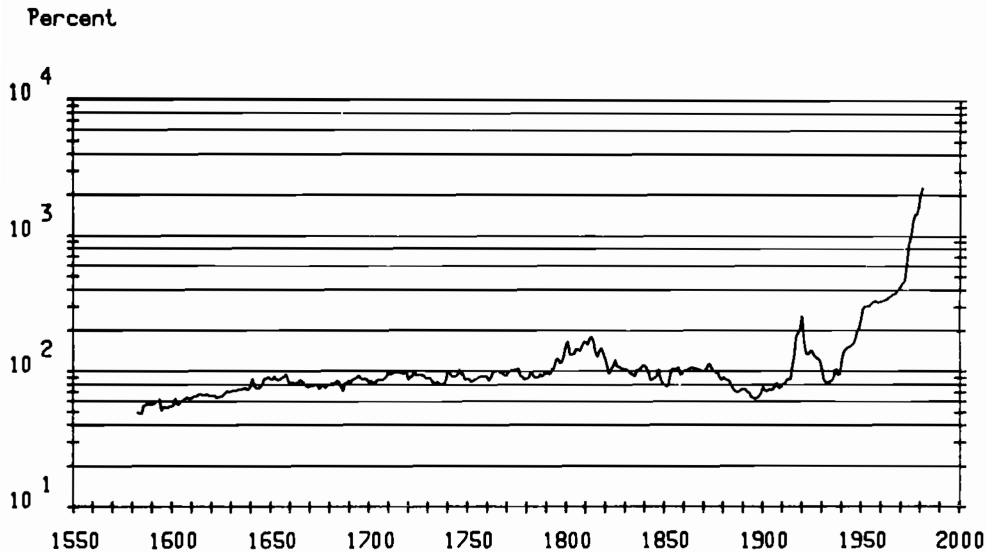


Figure 9.1. Wholesale price index, UK.

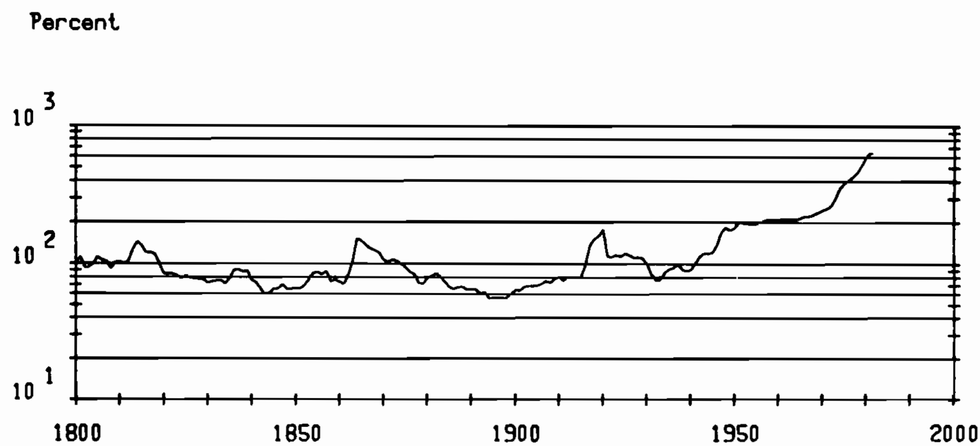


Figure 9.2. Wholesale price index, USA.

and there were also sharp increases during the last decade. In the United Kingdom, the fluctuations were subdued until the end of the eighteenth century. The first pronounced peak occurred around the year 1810, a weaker but prolonged peak around the year 1872, another pronounced peak in 1920, and a prolonged inflationary period during the last decade. Because all pronounced inflationary periods in both countries are associated with major wars, Hartman and Wheeler (1979) observed that the absence of strong inflation during the mid-nineteenth century in the United Kingdom could be partly due to the absence of such military conflict.

Clearly, price fluctuations in the United Kingdom and the United States display a broadly similar pattern, although the behavior in the United States prior to the mid-eighteenth century is not well documented and is uncertain. Prices in the United Kingdom portray a long decline from 1660 to about 1740 with two pronounced peaks in 1699 and 1710, and a long rise from about 1740 to about 1810, followed by another decline. The turning point between these two periods of rising and falling prices corresponds to the first pronounced peak. Especially large price rises occurred between 1785 and 1792 as the Industrial Revolution gained momentum. In the United States, a pronounced price peak occurred during the Revolution and the recovery period between 1775 and 1785. Although the two countries differed substantially in many respects, such as the level of industrialization, institutional development, energy use, and internal conflicts, the parallel in the pattern of price fluctuations through the nineteenth and twentieth centuries is striking. In the face of quickened industrialization, the Napoleonic Wars (UK), and the War of 1812 (USA), prices rose until the 1820s. A period of declining trend continued through 1850 (UK) and 1843 (USA), followed by a rising trend that was more pronounced in the United States, undoubtedly associated with the Civil War. The period from 1873 to 1896 is characterized by a declining trend in both countries, and the succession of rising and falling periods has remained almost exactly parallel until the present. Since World War II, prices in the United Kingdom and the United States have risen almost uninterruptedly and to unprecedented levels.

The price fluctuations in both countries indicate a regular and parallel pattern as will be elaborated shortly. Price peaks of the 1780s, 1820s, 1870s, and 1920s are spaced at intervals of four to five decades. These recurring long swings in prices are in our opinion not the primary causes of the long-wave phenomenon but rather a good indicator of the succession of alternating phases of the long wave. We consider the long swings in price movements to indicate the phases of rapid growth and saturation with the increasing level of prices and phases of recession, and regenerative destruction with decreasing price level (see Schumpeter, 1939).

To obtain a clearer picture of the timing of the long waves that are observable in the price indices of the two countries, we have decomposed the time series into fluctuations and a secular trend. Since the secular trend does not indicate a simple functional form, we have used a 51-year moving average method for its elimination from the time series. We have smoothed the resulting residuals (i.e., the relative difference between the actual price level and its secular trend expressed as a percentage) with a 15-year moving average. The

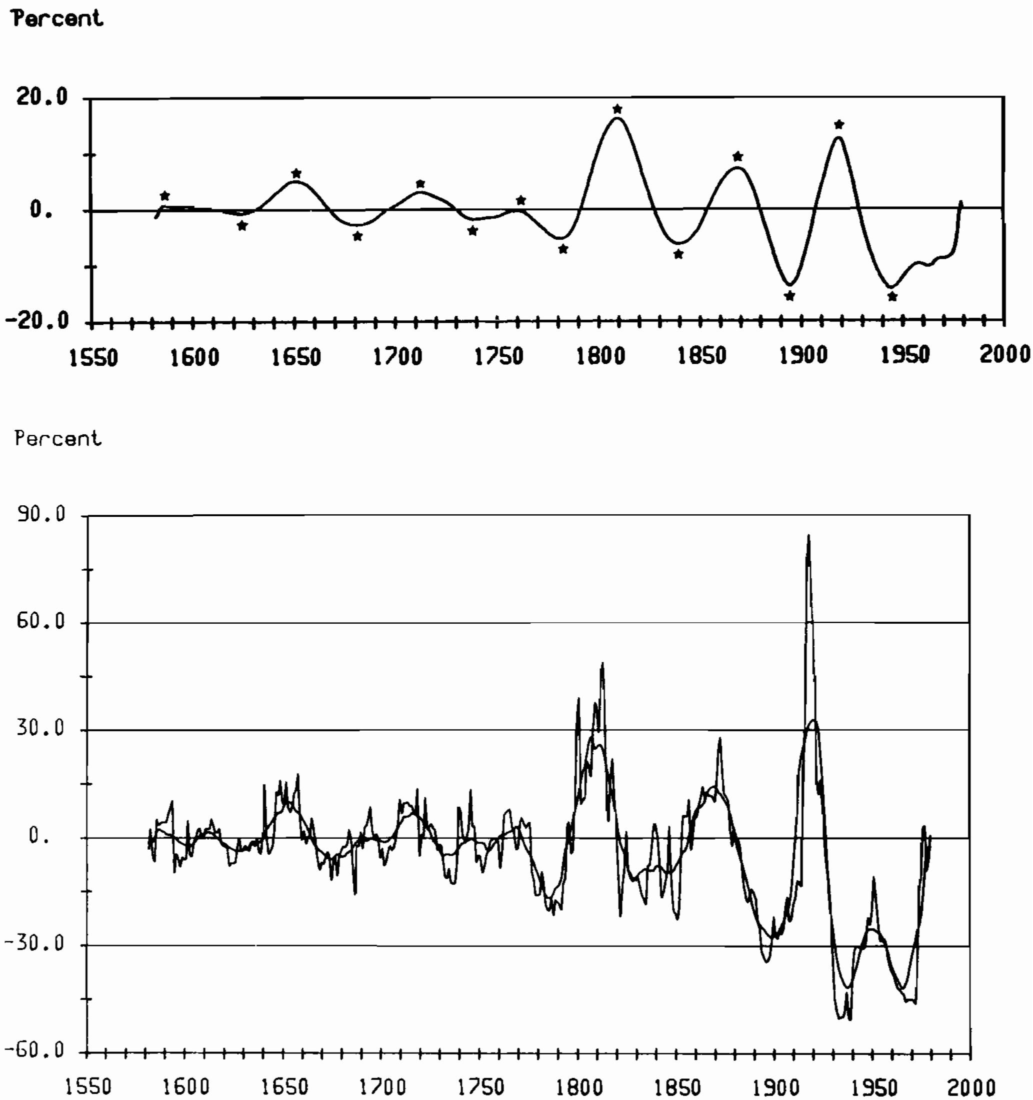


Figure 9.3. Long wave in wholesale prices, UK.

resulting stationary series (smoothed and unsmoothed residuals) are shown on the lower plots in *Figure 9.3* for the United Kingdom and in *Figure 9.4* for the United States. The upper plots in *Figures 9.3* and *9.4* show *stylized indicators* of the long swings in prices. The curves have been derived by smoothing the residuals by a 25- (instead of a 15) year moving average. We have chosen such a long moving average to eliminate some of the more pronounced fluctuations from the residuals that overlap the five-decade long swings. The stars approximately indicate the turning points of the long waves. We are aware that this kind of simple smoothing can introduce spurious fluctuations into the time series.

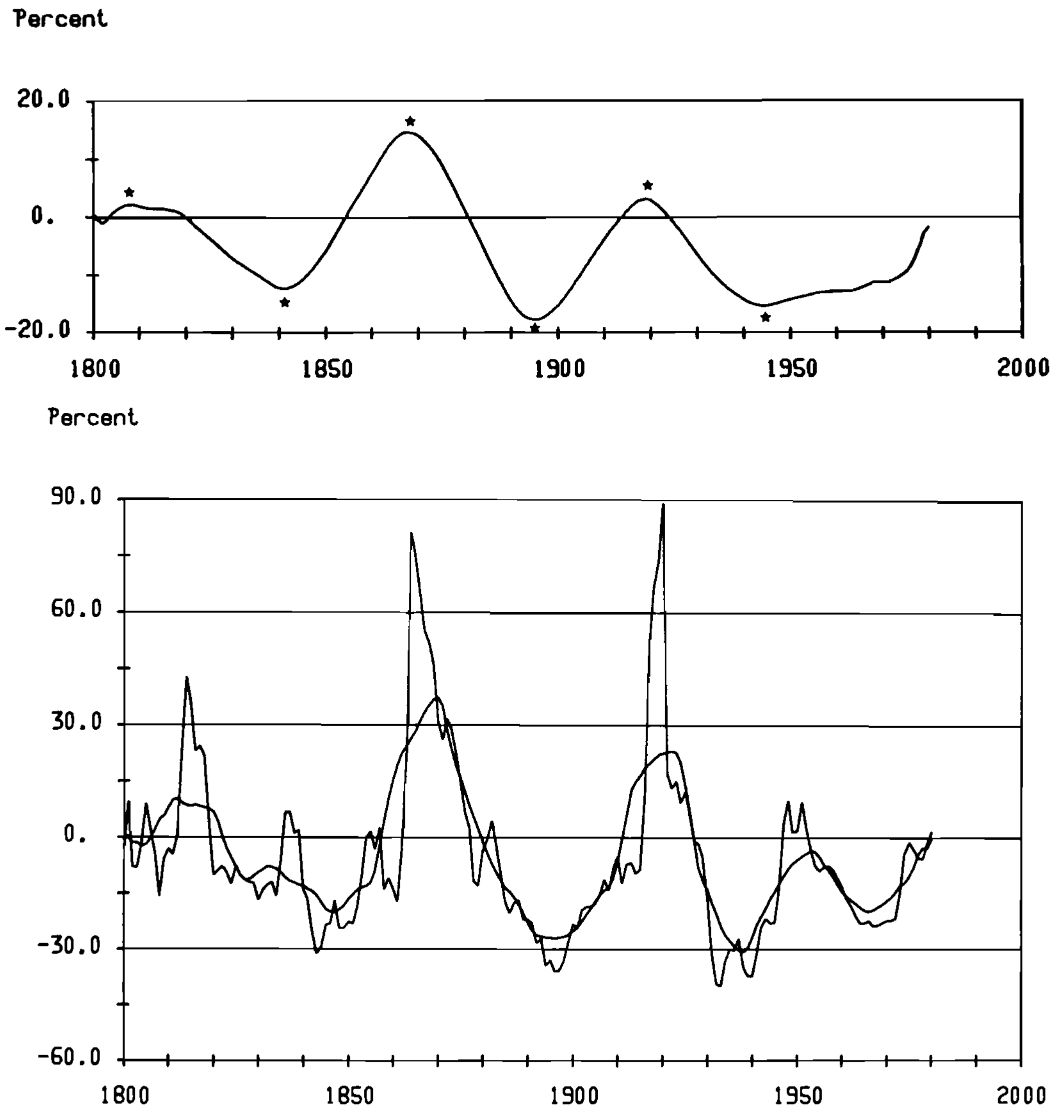


Figure 9.4. Long wave in wholesale prices, USA.

This empirical determination indicates not only a parallel development of the price movements in the two countries, but also a high degree of synchronization. For the period before 1800 in the United Kingdom, we dated turning points in the years 1623, 1651, 1681, 1712, 1739, 1753, and 1773. The intervals between the four succeeding troughs are 58, 58, and 44 years and between the three peaks 61 and 41 years. As the Industrial Revolution gained momentum, first in the United Kingdom and later in the United States, the long swings became more regular and the magnitude of the fluctuations increased. The average amplitude of the fluctuations rose from less than 10% in preindustrial United Kingdom to about 20% in both countries. *Table 9.1* shows the dates of the turning points and the duration of the long swings in prices for the two countries.

The average duration of the fluctuations is about 50 years and the occurrence of peaks and troughs varies by not more than one or two years. We consider *Table 9.1* a rough, empirical indicator of the timing of long waves in the two leading countries. This timing is similar to the stylized schemes derived by van Duijn (1983) and Bieshaar and Kleinknecht (1984). In subsequent examples we will use this empirical indicator of the long-wave turning points to determine the correspondence between the fluctuations that we will establish in other monetary and quantitative indicators of economic development.

Table 9.1. Chronology of the long wave, UK and USA.

	Price swings			
	United Kingdom		*United States*	
Phase	*Period*	*Duration*	*Period*	*Duration*
Downswing	1585–1623	48		
Upswing	1623–1651	28		
Downswing	1651–1681	30		
Upswing	1681–1712	31		
Downswing	1712–1739	27		
Upswing	1739–1753	24		
Downswing	1753–1773	20		
Upswing	1773–1810	37		
Downswing	1810–1840	30	1809–1841	32
Upswing	1840–1869	29	1841–1869	28
Downswing	1869–1895	26	1869–1895	26
Upswing	1895–1920	25	1895–1920	25
Downswing	1920–1945	25	1920–1945	25
Upswing	1945–		1945–	

9.3. Transport Systems

9.3.1. Aircraft

We will now show that the evolution of transport systems and infrastructures can be analyzed as a sequence of replacements of old modes of transportation by new ones. Furthermore, the substitutions of old for new technologies overlap the growth phases in the long-wave fluctuations as indicated by price flares. The analysis starts, somewhat unconventionally, with the youngest transport system (air travel) and ends with the oldest (canals and waterways).

The rapid expansion of air travel during the recent decades has its roots in developments achieved in aerodynamics and other sciences many decades ago, and especially in the engineering achievements made between the two wars. The DC-3 airliner is often given as the example of the first "modern" passenger transport because in many ways it denotes the beginning of the "aircraft age." The use of aircraft for transportation has increased ever since and its performance has improved by about two orders of magnitude. *Figure 9.5* shows the increase in air transport worldwide measured in billions of passenger kilometers per year

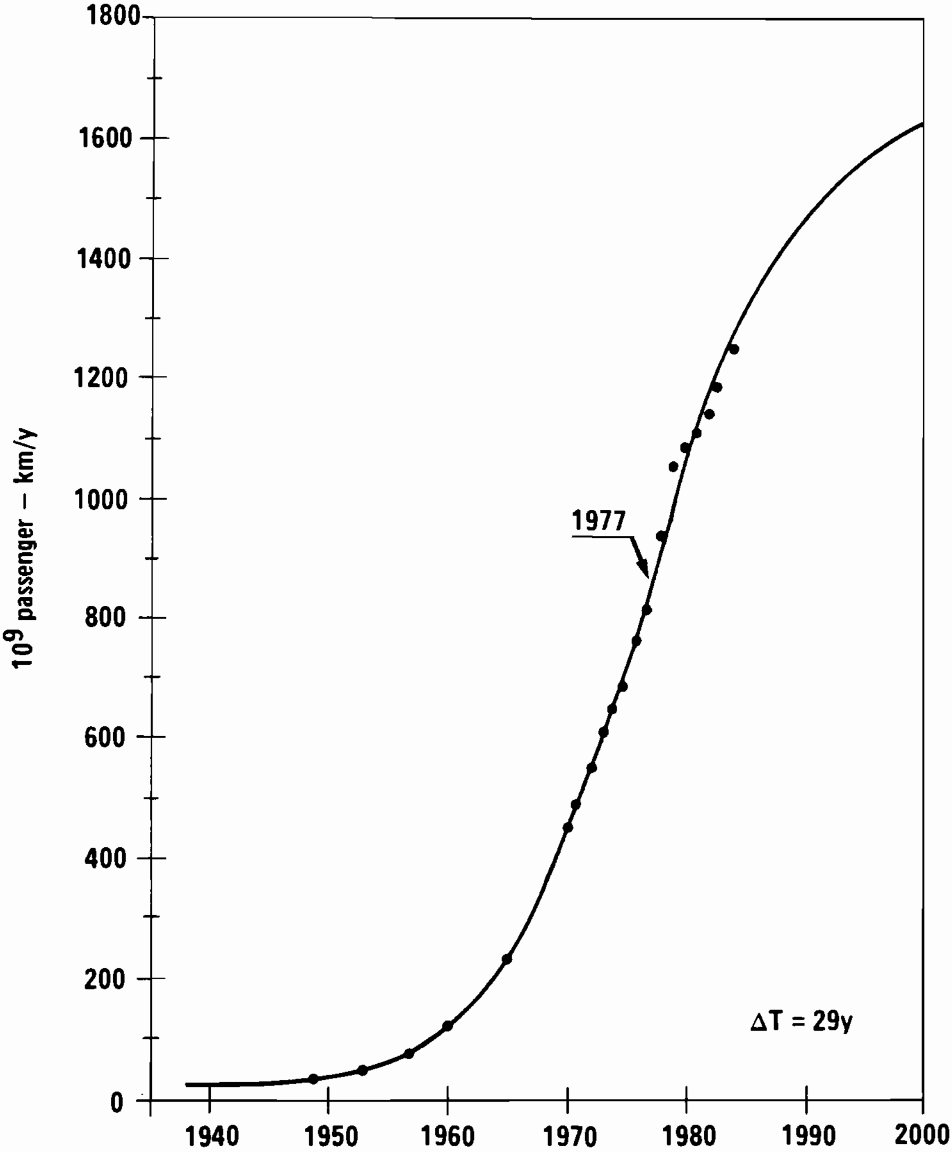


Figure 9.5. Air transport worldwide.

(passenger-km/yr). It gives all carrier operations including those of planned economies. The logistic function has been fitted to the actual data, and it indicates that the inflection point in the growth of air travel occurred about 10 years ago (1977).[1] Thus, after a period of rapid exponential growth, less than one doubling is left until the estimated saturation level is achieved after the year 2000. The figure shows that the most rapid expansion of air travel in the world lasted from the 1930s until the 1970s and that the growth rate has been declining for about the past 10 years. Therefore, air travel expanded in much the same way growth processes do in biology, as S-shaped growth patterns. The most rapid expansion of air travel took place during the growth phase of the long wave.

Figure 9.6 shows the same data and fitted logistic curve transformed as $x/(\kappa-x)$, where x denotes the actual volume of all operations in a given year and κ is the estimated saturation level. The data and the estimated logistic trend line are plotted in *Figure 9.6* as fractional shares of the saturation level, $F = x/\kappa$, which simplifies the transformation to $F/(1-F)$. Transformed in this way, the data appear to be on a straight line, which is the estimated logistic function.

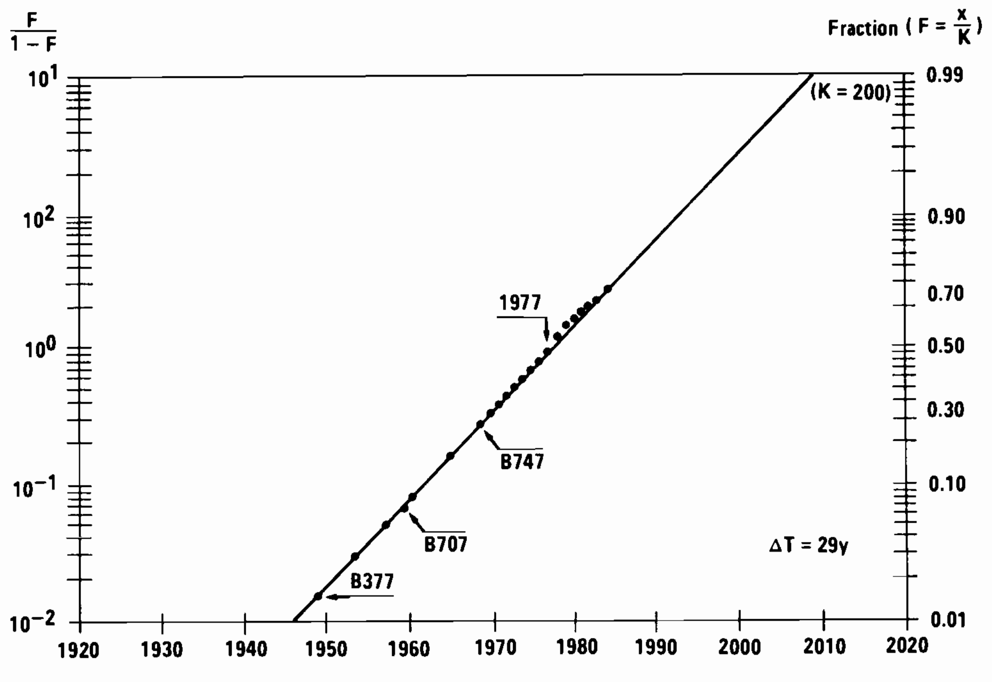


Figure 9.6. Air transport worldwide, logistic plot.

Perhaps the more interesting results are that it took about 30 years for world air transport to reach the inflection point (half the estimated saturation level) and that after two decades the saturation level will be reached. This raises a crucial question. What will happen after saturation? Can we expect another growth pulse, a decline, or the instability of changing periods of growth and decline? Most likely a new period of growth associated with new technologies will follow the projected saturation (see also Lee and Nakicenovic, Chapter 1).

During the same period, while air travel worldwide increased by two orders of magnitude, the productivity of the individual aircraft also increased by two orders of magnitude from the DC-3 to the Boeing 747 (see Nakicenovic, 1987a). Another growth phase of air travel, or some other new transport system, in the next century would require an analogous increase in the productivity of the vehicles. In the case of aircraft this would imply supersonic or hypersonic (extremely large subsonic transports) or both.

How probable is the development of a large cruise supersonic or hypersonic transport? S-pulses do not usually occur alone but in pairs. Usually structural change occurs at the saturation level, leading to a new growth pulse and in turn to new productivity and performance requirements for succeeding technologies. This logic would suggest the need for a more productive means of long-distance transportation for the next century than the current wide-bodied families of subsonic aircraft. It is questionable whether history repeats itself, but we will show that in the past each growth phase of the long wave is associated with the evolution of a number of important technologies that tend to saturate during the end of the prosperity phase and during the recession phase of the long cycle. Below we illustrate that the growth of the older technology, road transport systems, in the United States can be described by a pair of successive growth pulses with an intervening saturation during the 1930s marked by a period of change. Using the data available, we illustrate the evolution of road vehicles and other transport systems in the United States, but will return to analyzing long waves and technological change at the global level thereafter.

9.3.2. Automobiles

At the beginning of this century, few proponents of the automobile envisaged that its use would spread so rapidly throughout the world. As a commercial and recreational vehicle, the motor car offered many advantages over other modes of transportation, especially animal-drawn vehicles. Perhaps the most important advantage was the possibility of increasing the radius of business and leisure transport.

The most rapid expansion of the automobile was witnessed in the United States. It had a relatively late start in relation to European countries, such as France, Germany, and the United Kingdom. According to the records, four motor vehicles were used in the United States in 1894. This was followed, however, with an impressive expansion of the automobile fleet: 90 in 1897; 8,000 in 1900; almost half a million 10 years later; and more than one million after another two years. Thus the United States quickly surpassed European countries both in production and in the number of vehicles in use.

Figure 9.7 shows the rapid increase in the number of cars used in the United States. It also shows that the expansion of the automobile fleet is characterized by two distinct secular trends, with an inflection in the 1930s followed by less rapid growth rates. Since the two secular trends on the curve appear to be roughly linear on the logarithmic scale, the automobile fleet evolved through two exponential pulses. Thus, in this example, the growth of the automobile fleet did not follow a simple, single S-shaped growth pulse.

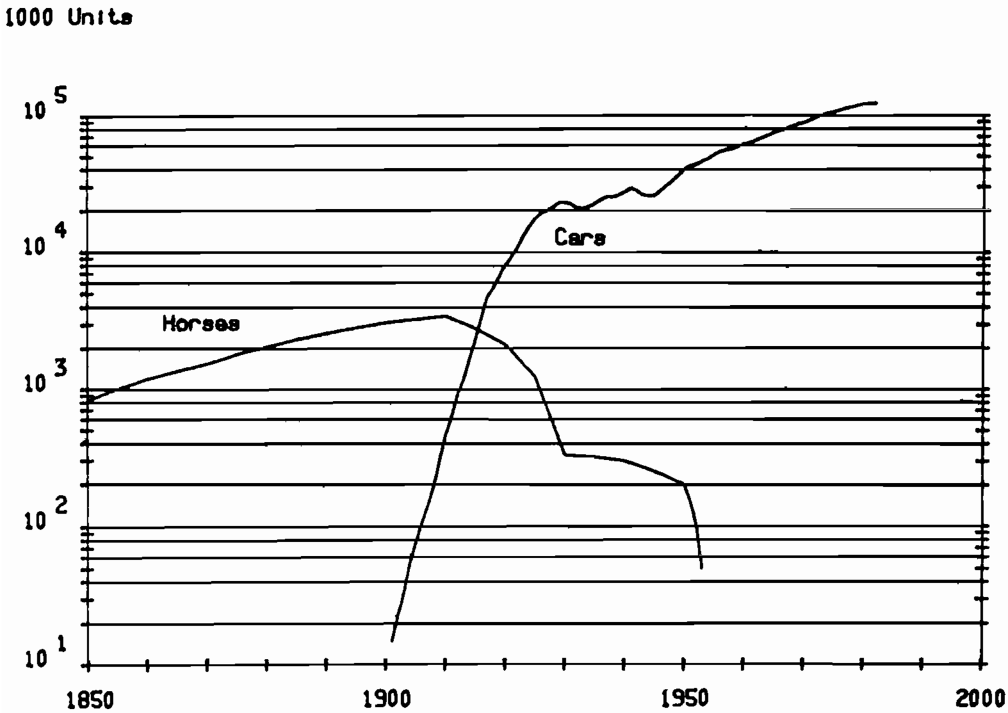


Figure 9.7. Number of automobiles and road horses (and mules), USA.

The working hypothesis here is that the two trends indicate two different phases of the dissemination of motor vehicles in the United States. The first characterizes the substitution of motor vehicles for horse-drawn vehicles and the second the actual growth of road transport after animal-drawn vehicles essentially disappeared from American roads. Thus the first expansion phase was more rapid since it represents "market takeover" or substitution for older means of transport, whereas the second represents the actual growth of road vehicle fleets and their associated infrastructure, such as highway systems. The inflection point that connects the two growth pulses coincides with the prolonged recession in the long-wave cycle. Thus our hypothesis implies that the motor vehicle fleets evolved differently in the two adjacent Kondratieff cycles.

The lack of historical records of the exact number of horse-drawn vehicles in the United States soon after the introduction of the automobile in 1895 makes it difficult to describe accurately the assumed substitution of the motor car for the horse during the first, more rapid expansion phase of the motor vehicle fleets. The number of draft animals (road horses and mules) and the automobiles given in *Figure 9.7* are therefore a rough approximation of this substitution process. *Figure 9.8* gives fractional market shares of horses and cars in all road vehicles (sum of horses and cars). Market shares, F, are plotted on a semilogarithmic plot transformed as $F/1 - F$, as a ratio of the market shares of one technology over the other since fractional market shares always sum to one.[2]

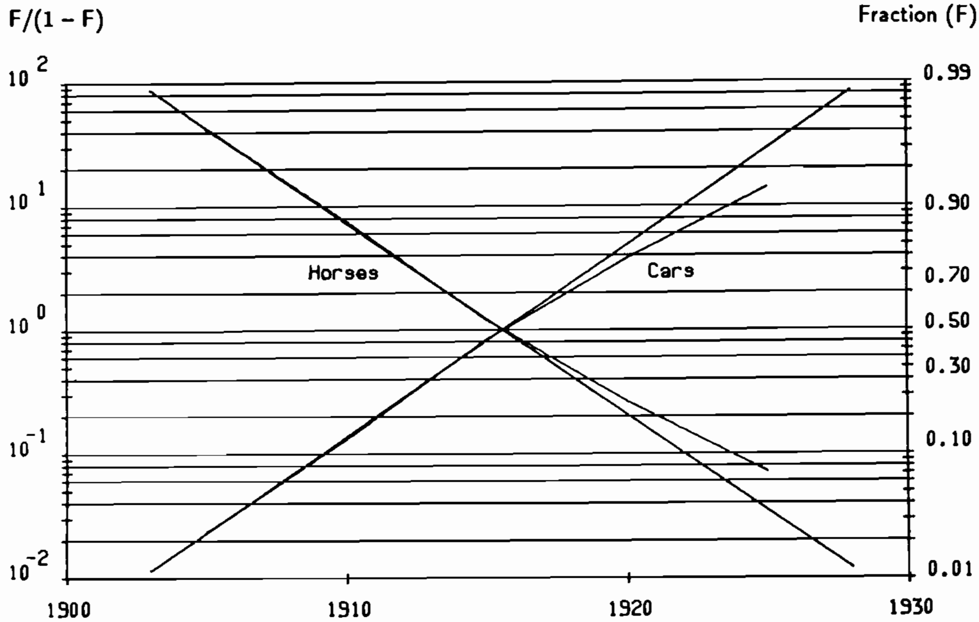


Figure 9.8. Substitution of horses for automobiles, USA.

Figure 9.8 indicates that the automobile replaced animal-drawn road vehicles during a relatively short process and proceeded along a logistic path. Motor vehicles achieved a 1% share of road vehicles shortly after 1900 and a 50% share in 1916. A complete takeover occurred in 1930 when there were 0.3 million road horses and mules and 23 million cars, an increase from less than 2 million cars 10 years earlier. Thus the inflection point in the growth of the automobile fleet from *Figure 9.7* actually coincides with the end of the replacement of animal-drawn road vehicles by motor cars and explains the apparent saturation in the growth of motor vehicles observed by many analysts during the late 1920s and early 1930s. This perceived saturation marks the beginning of a new phase in the motorization of America, with growth rates comparable with those of the expansion of horse-drawn vehicles before the automobile age. Seen from this perspective, the number of all road vehicles increased from 1870 to 1930 but from 1900 to 1930 the automobiles replaced the horses, whereas after the total replacement, only the number of cars was expanding during the last long wave. In most European countries, the rapid expansion of the automobile also started during the last Kondratieff wave. In this example we see two aspects in the dynamics of technological change: the growth of technological populations, in this case road vehicles, and the replacement of older by newer technological species. *Figure 9.9* shows the growth in the number of all road vehicles as a continuous growth process with an apparent saturation level of about 350 million vehicles after the year 2000 and a ΔT equal to about 100 years.[3] Because the growth of road transport in general and the substitution of automobiles for horse-drawn carriages and wagons overlap in time, together they produce two growth trends in the growth of the automobile fleet with an inflection point in the 1930s making the structural change in the composition of the road vehicle fleets. Therefore, this example illustrates that during the last two growth phases of a long wave, road vehicle fleets developed differently. One tentative conclusion from this example could be that the further development of road vehicles could take another new path after the 1990s.

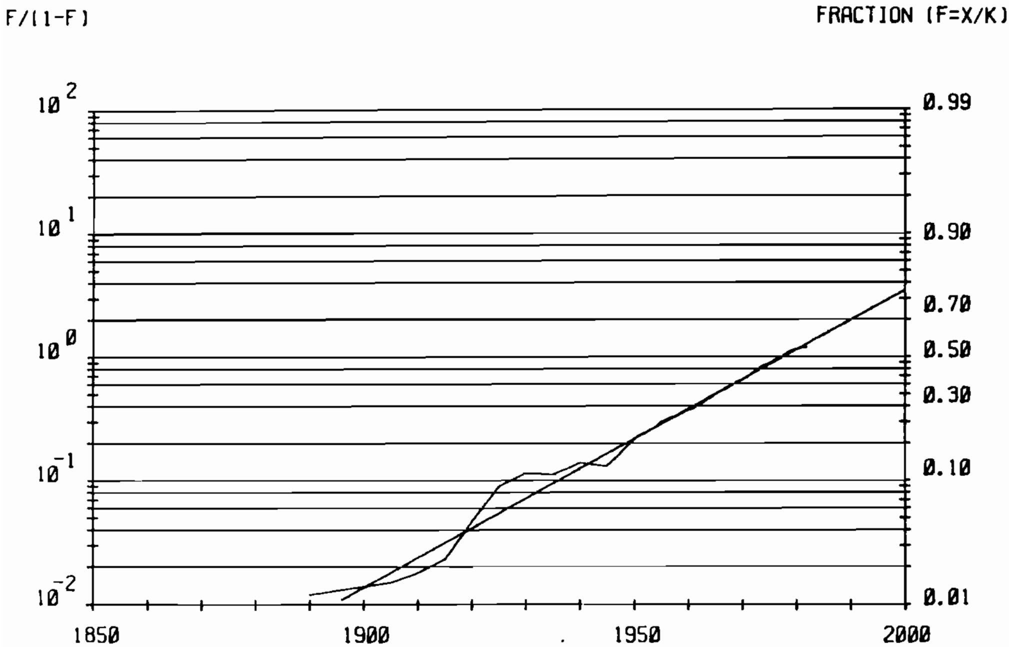


Figure 9.9. All road vehicles in use, USA.

The expansion of the road vehicle fleets in the United States and the growth of the global air travel illustrate two important aspects of technological and economic change. The expansion of air travel was shown as a S-shaped *growth pulse*. This example parallels growth processes in biology, for example, growth of a leaf or bacteria population. The replacement of horses by cars shows *substitution* of an old for a new technology as an S-shaped increase in market shares of the new competitor.

We have argued that during each expansion phase of the long wave a number of important technologies are developing simultaneously, growing, replacing old ones, and usually enhancing each other. Road transport systems require, in fact, elaborate and sophisticated infrastructure. The development of road transport vehicles (and infrastructure) illustrates the succession of replacements of old by new as one of the basic features of the development process. Furthermore, it shows that the different replacement processes are associated with the two long waves. Thus we have seen that the expansion of automobiles in the United States and global air travel took place during the growth phase of the last long wave, whereas the replacement of horses by cars was a feature of the previous long wave when railroads were the dominant form of transport in most of the developed world. Railroads are now in the post-saturation phase; their position as a means of passenger travel is being eroded in most industrialized countries and has become insignificant in the United States. A symbol of this decay is the discontinuance of the transcontinental railway service in the United States.

9.3.3. Transport infrastructures

Both air and railroad transport systems require elaborate infrastructures. In fact, airports and railroads were obviously constructed for the sole purpose of providing infrastructure for aircraft and trains. However, this distinction is not clear for roads, although we have shown that the construction of surfaced roads preceded the expansion of the automobile fleet. This similarity in the evolution of the transport systems is perhaps indicative of an invariance in the development process of transport systems and their underlying infrastructure. A serious problem arises, however, when comparing railroads and roads with other transport systems that do not depend exclusively on the rigid, man-made links between them. Airways and waterways, for example, rely less on man-made links between the nodes because they use the natural environment (air, rivers, coastal waters). Nevertheless, they require an elaborate infrastructure, such as airports, harbors, and canals. Thus it is difficult to compare the total length of the implicit airway and waterway routes with the total length of the main railroad tracks and surfaced roads. As an approximation in the analysis of the evolution of transport infrastructures, we will use the sparse accounts and probably inaccurate estimates about the construction of canals as an indicator for the total network of the waterway transport systems.

Figure 9.10 shows the length of the three successive transport infrastructures: canals, railroads, and surfaced roads. In the United States the first canals were built during the 1780s but in fact, canal construction really accelerated during the first decades of the nineteenth century, making that period the "canal era." It expanded very rapidly to about 2,000 miles until 1831, when 13 miles of the Baltimore and Ohio Railroad went into operation. Thus the canal era lasted until the railways became the main mode of long-distance transport a decade later. From this point of view, the 1830s were turbulent years: many turnpikes were abandoned; canal construction was reaching its peak; and important early railway projects were already completed. *Figure 9.10* shows that the railroads remained the longest transport network until 1920, when they were surpassed by the rapidly expanding road system and the automobile. The construction of railroads saturated during the following decade and has been decreasing ever since. This is analogous to the saturation of canals during the 1840s and their decline thereafter. *Figure 9.10* shows that the evolution of the transport infrastructure can be seen as a succession and replacement of older with newer transport systems, although at any given time more than two systems were actually in operation.

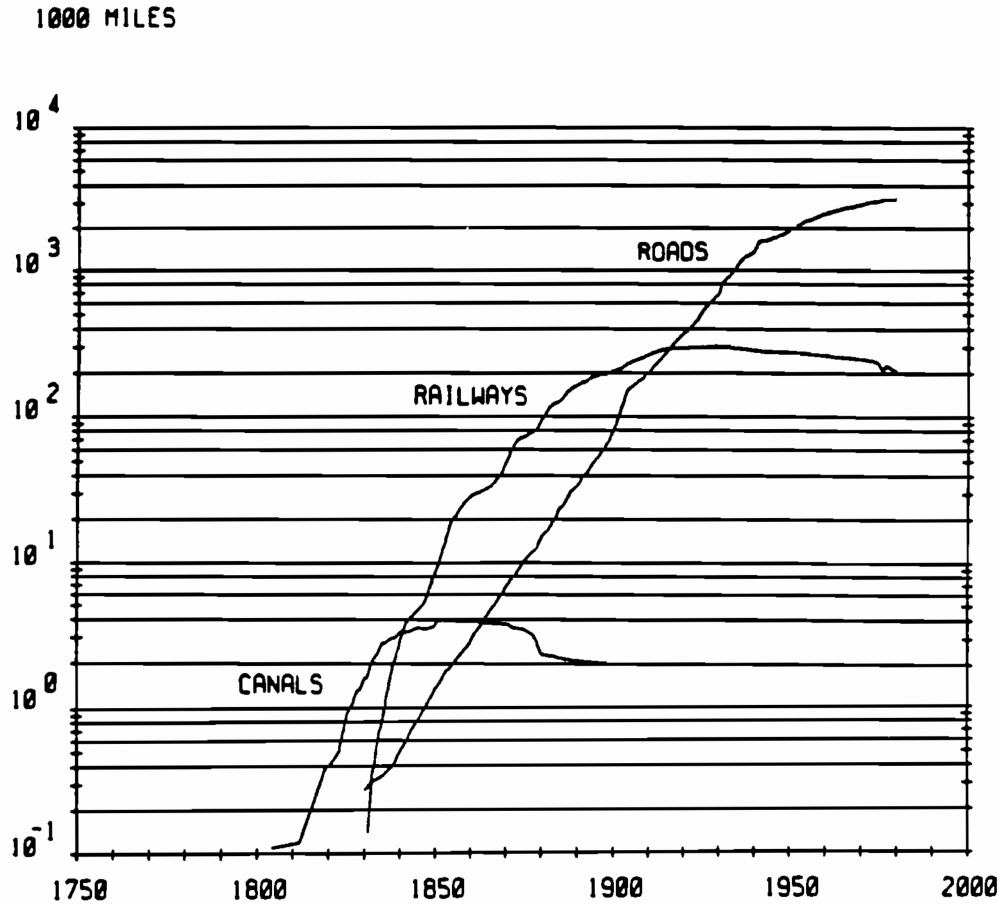


Figure 9.10. Length of transport infrastructures, USA.

This succession of the three transport infrastructures can be described in terms of three S-shaped growth pulses that are given together with the estimated logistic curves in *Figure 9.11*. Seen as successive growth pulses, the expansion of canals saturated during the 1860s at a level of about 4,000 miles and the expansion of railroads saturated during the 1930s at a level of about 300,000 miles, whereas roads will saturate during the coming years. Thus, the three transport systems saturated successively at intervals of about six decades. *Figure 9.12* shows the same growth pulses transformed so that the data and the S-curve appear as a straight line. This indicates that the development of canals was much quicker (with a ΔT of about 30 years) than the expansion of railroads and roads (with a ΔT of 54 and 56 years, respectively). Thus, canals have a time constant comparable with that of airways.

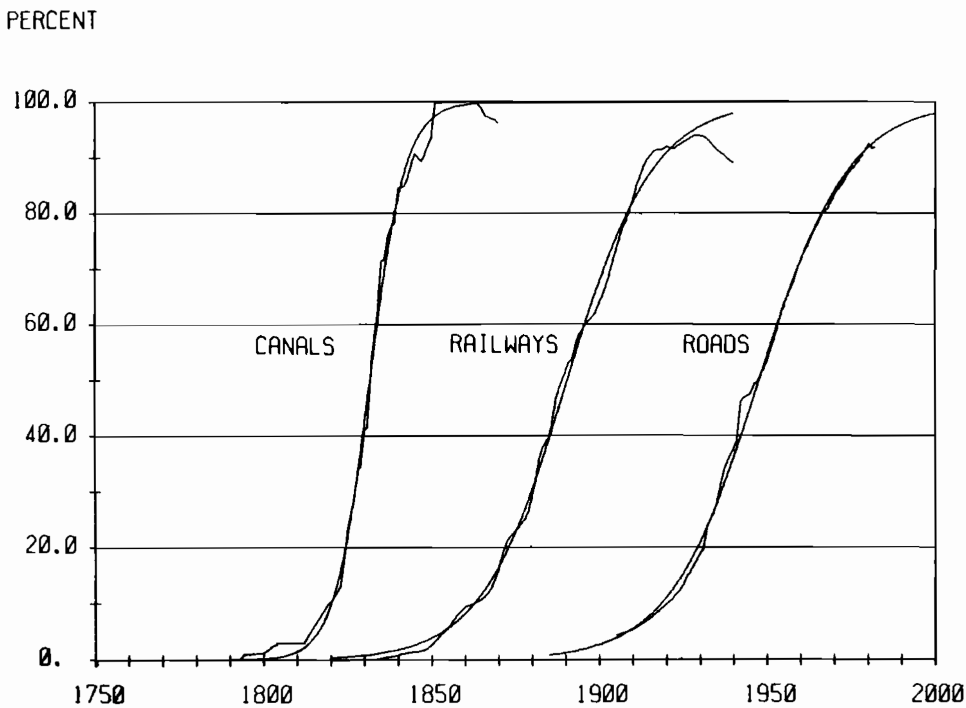


Figure 9.11. Growth of infrastructures as fraction of saturation, USA.

The difference in the time constant between air and inland water transport systems, on the one hand, and rail and road transport, on the other, indicates that at least at this level of comparison transport systems having more extensive infrastructures may take longer to expand, and possibly to complete the whole life cycle from growth to saturation and later senescence. Thus, it is remarkable that in spite of these differences, the saturation in the growth of these three infrastructures coincides with the beginning of the prolonged recessions in the last three long waves.

To assess whether the time constants are really different, *Figure 9.13* shows the successive substitutions of the three transport infrastructures and the federal airway route miles. The substitution process is shown as relative market shares (F) of a given transport infrastructure to the total length of all of the infrastructures together.[4] From this perspective the replacement of the four systems over time appears as a regular process.

This result may appear to contradict the earlier observation that the total length of railway tracks and surfaced roads took longer to construct than water and airway routes. In fact, the timetable associated with the substitution dynamics of infrastructural length is surprisingly consistent in relation to the duration of growth pulses of the four transport modes during the past 180 years. The apparent inconsistency results from the different ways of measuring the

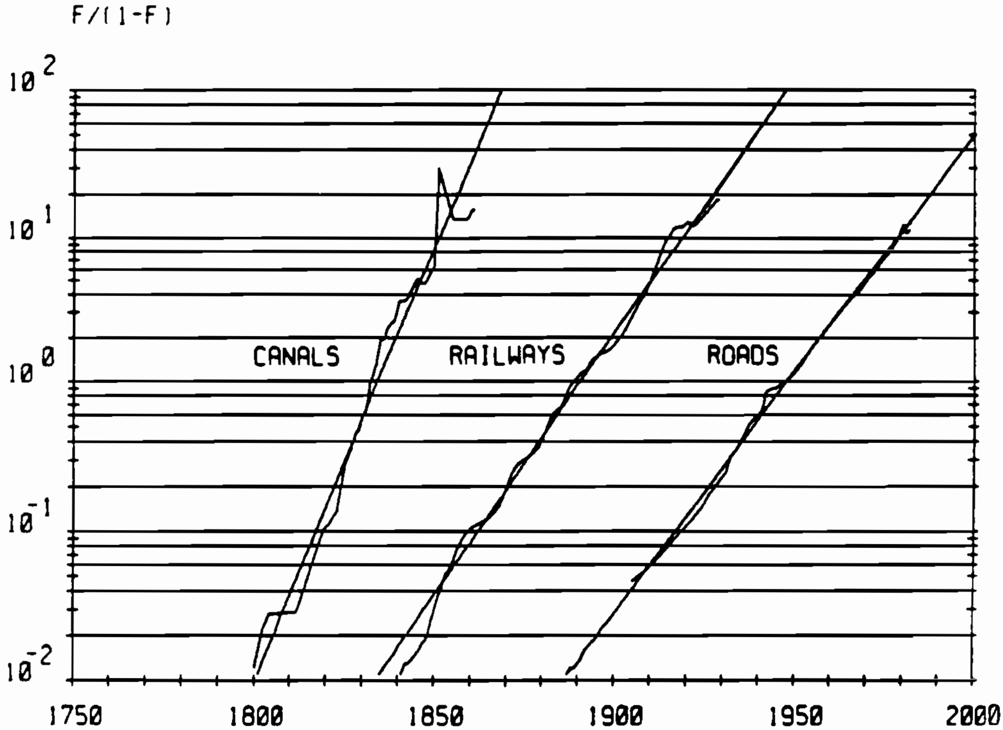


Figure 9.12. Growth of infrastructures as fraction of saturation, transformed.

growth rates and life cycles of the respective infrastructures. In the case of market shares the increase in a particular transport infrastructure is analyzed in terms of the length of all networks. Thus, even the rapid growth rate of airway route mileage is translated into a comparatively long time constant because at the same time the total length of all transport networks is also growing rapidly. As a result of these rapid growth rates, the share of surfaced roads has been declining since the 1970s, whereas the total length of surfaced roads is still growing toward the ultimate saturation level.

Thus, the total length of a transport infrastructure (in this case, canals, railroads, and surfaced roads) can still be growing even decades after ultimate saturation and final senescence, while the share of its length in all transport infrastructures is declining. The saturation and decline of market shares therefore precede saturation in absolute growth in a growing market, meaning that in those cases the eventual saturation of any competing technology can be anticipated in the substitution dynamics.

This description of the evolution of transport systems and infrastructures shows that during each growth phase of the last three long waves one of the important transport systems developed in the United States. Thus, Schumpeter's association of the last three long waves with canals, railroads, and automobiles can be confirmed from the empirical point of view.

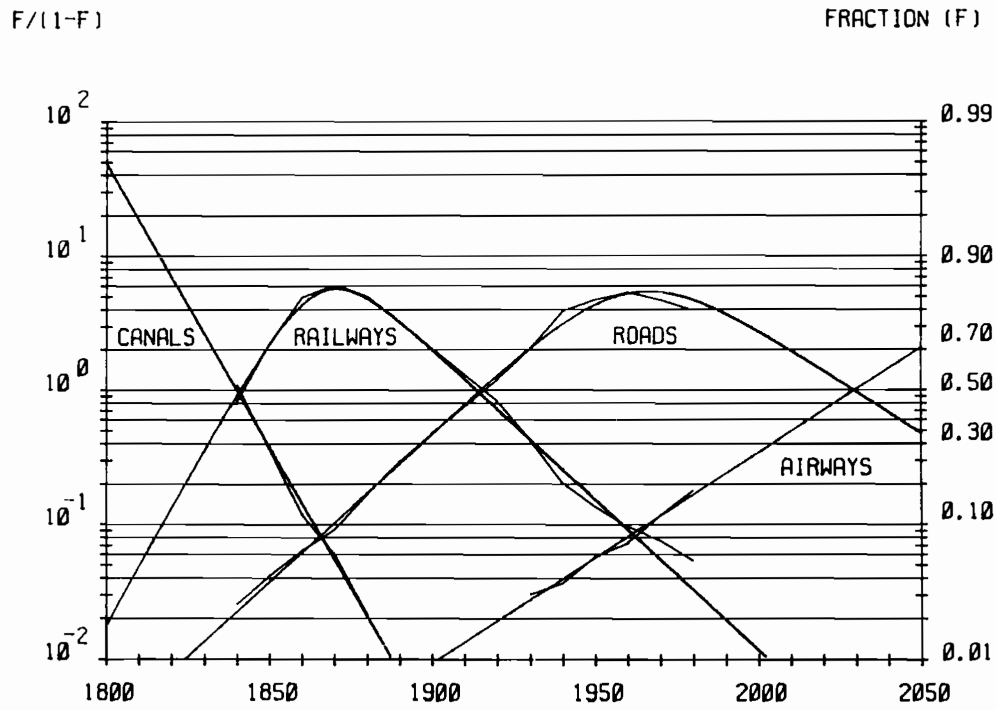


Figure 9.13. Substitution of transport infrastructures, USA.

9.4. Steel and Energy

9.4.1. Steel production

A widely accepted view is that the iron and steel industry constitutes a more mature sector of developed economies and an important sector in developing countries. In other words, iron and steel production is approaching or already has reached saturation and is declining in the industrialized countries, but it is still growing in most of the developing world. Subsequently, the iron and steel industry worldwide is in different phases of development, ranging from the early development and expansion phase to that of maturity and decline. We will attempt to give empirical evidence for this broad and long-term development of global steel production. In Chapter 8 Grübler analyzes specific changes of individual countries and technologies in greater detail.

Metallurgy dates back to the dawn of human civilization, but, because metals were precious, wood, stone, and sometimes bones were the dominant materials to help accomplish a difficult task. In spite of a wider and more sophisticated use of metals (initially mostly copper and bronze, and later also iron and some steel), the use of traditional materials prevailed through antiquity and the Middle Ages. The voracious use of iron and steel evolved parallel with the so-called Industrial Revolution. *Figure 9.14* illustrates the enormous increase in steel production worldwide since 1860. Production and growth has been especially rapid since the end of World War II, increasing from about 160 million tons to more than 850 million tons in less than four decades. *Figure 9.14* also shows wide fluctuations in steel production during the 1920s and 1930s and again during the last decade.

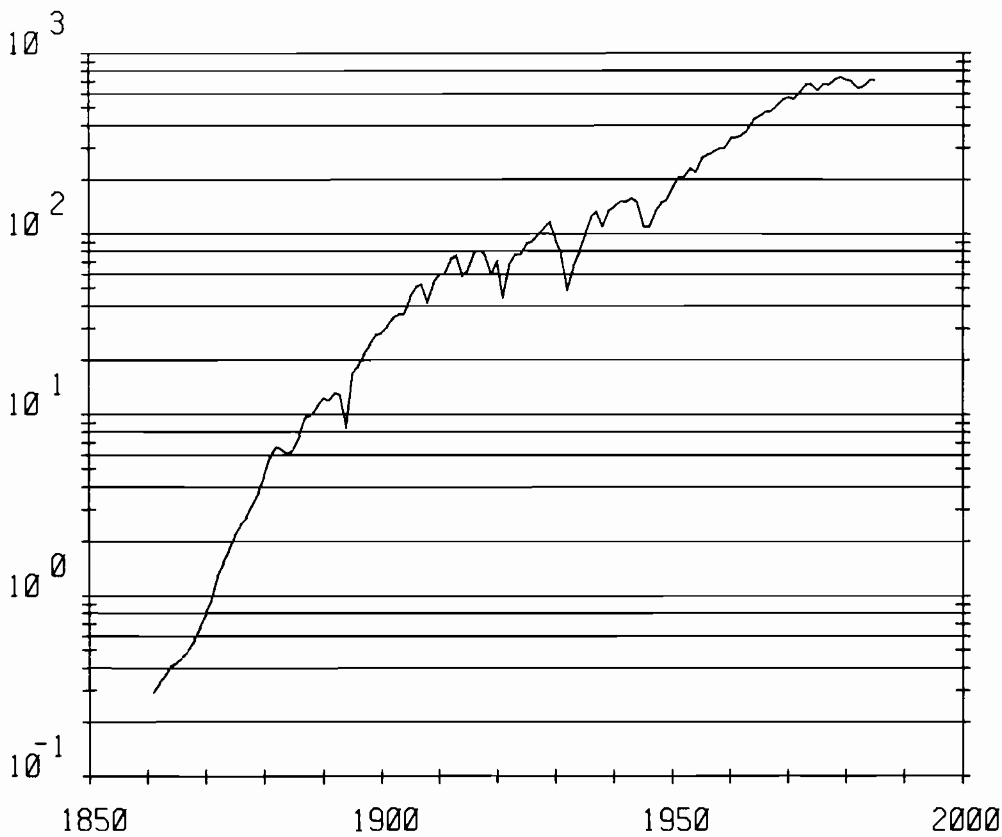


Figure 9.14. Global steel production. (Source: Grübler, 1987.)

This enormous increase in steel production is due to the crucial role that materials have in the development of industrial societies. The precise details of this strong coupling between economic development in general and materials in particular is truly complex, but the basic reason is actually obvious and transparent – the effectiveness and design of machines, equipment, and infrastructure depend to a substantial degree on the materials of which they are made. The discussion about the development of transport infrastructures illustrated this point very vividly: successive replacements of turnpikes, canals, railways, roads, and airways paralleled major improvements in construction materials from wood and stone to iron, steel, and concrete, and during recent decades to more advanced alloys and materials.

The introduction of better materials was instrumental for the development of new manufacturing techniques, energy sources, and transport systems. Metallurgy, manufacturing, energy, and transport all developed owing to numerous cross-links as one improvement or breakthrough made another possible and sometimes necessary. It so happens that during the last two centuries iron and steel were perhaps the most critical of all widely used materials. While stone and wood continued to be important materials, the advent of steam, railroads, and the coal "age" would not have been possible with wood and stone. In fact, a more efficient technology for producing high-quality steel was required, and, after a series of major innovations ranging from the substitution of charcoal for coke to new casting methods, it culminated in the invention of the Bessemer steel process in 1857. *Figure 9.15* shows the simplified representation of the technological changes in steel production since 1860, starting with puddle steel as the oldest production method and ending with electric arc steel. *Figure 9.15* shows the fractional market shares (F) of the five competing steel technologies transformed as $F/(1-F)$ on a semilogarithmic plot.[5]

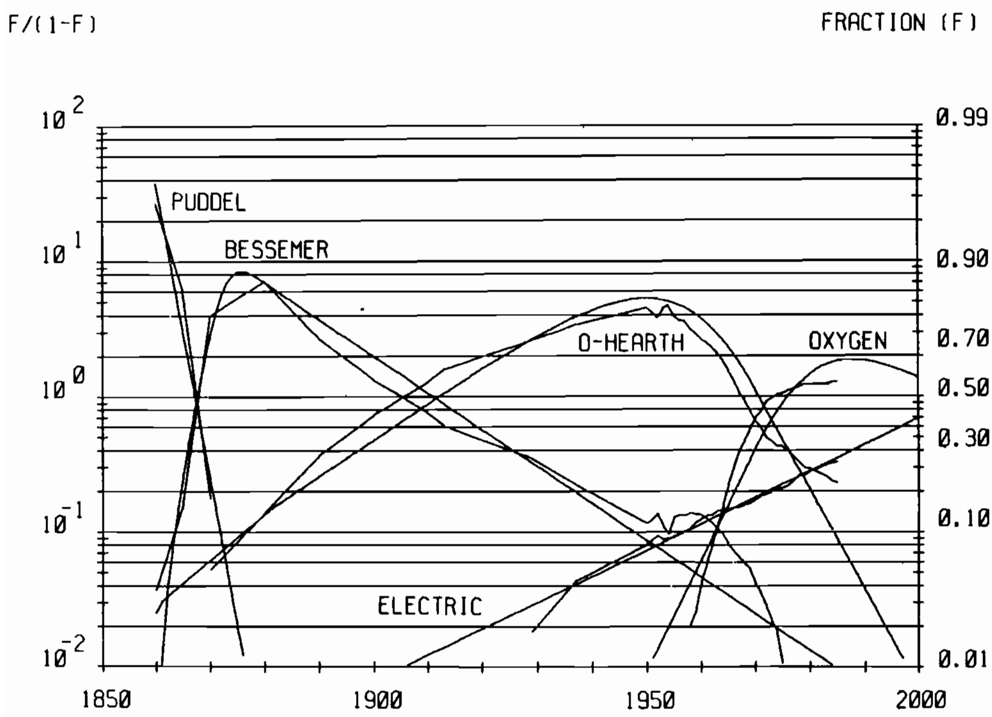


Figure 9.15. Steel substitution, world. (Source: Grübler, 1987.)

In spite of the very different nature of the five steel production technologies, their chemistry, energy sources, share of scrap iron, etc., they all appear to be in competition with each other, the newer technology eventually displacing the older. The linear trends indicate where the replacement of old with new technologies followed logistic curves. Grübler (Chapter 8) shows that the same historical trend can be observed in most industrialized countries. The replacement of crucible and puddle steel was a very fast process, whereas the open hearth method developed into the dominant steel technology over many decades. Bessemer became the dominant steel technology during the 1870s and thereafter its importance declined, while the open hearth process expanded. From 1870 to 1950 most of the increase in global steel production was achieved by the improvements and expansion of the open hearth method. The electric arc steel process was introduced during the 1920s, and its market share is still expanding with increasing amounts of recycled steel in the total production. After 1950 basic oxygen steel expanded vigorously, but now its relative contribution to steel production is saturating.

Thus, we have identified three distinct phases in the evolution of steel technologies. The first ended with the swift introduction of Bessemer steel, the first industrial process that could achieve high-quality and large-scale production of steel. The second is congruent with the development of the open hearth steel process, and the third marks the expansion of electric and basic oxygen methods. Another way of describing this succession of replacements in the evolution of the steel industry is to decompose the aggregate steel production (from *Figure 9.14*) into appropriate development phases. *Figure 9.16* shows total steel production (from all five technologies) in per capita terms as two distinct S-shaped growth pulses. The first starts with the dominance of the Bessemer method and mirrors the expansion of the open hearth process, while the second starts with the introduction of the electric arc process and accelerates with the expansion of basic oxygen steel.

Figure 9.17 shows the same growth pulses transformed so that the two S-shaped curves appear as straight lines.[6] Transformed in this way, the two pulses appear as parallel lines indicating equal duration of the two pulses with a ΔT of about 45 years. The two pulses overlap during the period of highest turbulence indicating that another period with large fluctuations in per capita (and therefore total) steel production may have started. These two pulses coincide with the growth phases of the last two long waves (see the rough timing for the long waves in the United Kingdom and the United States given in *Table 9.1*). *Figure 9.18* shows that in addition to broad fluctuations in the wholesale prices, the prices of iron and steel also show pronounced flares during the depression years in these countries. Thus, the iron and steel prices reflect important structural changes in the global steel industry.

This example confirms our original hypothesis that each growth phase of the long-wave cycle is associated with the expansion of a number of important new technologies through successive replacements of the old ones. Furthermore, the evolution of transport infrastructures and steel technologies is a very regular replacement process, but the saturation of dominant technologies does not correspond exactly to the timing of the depression phases in the long wave.

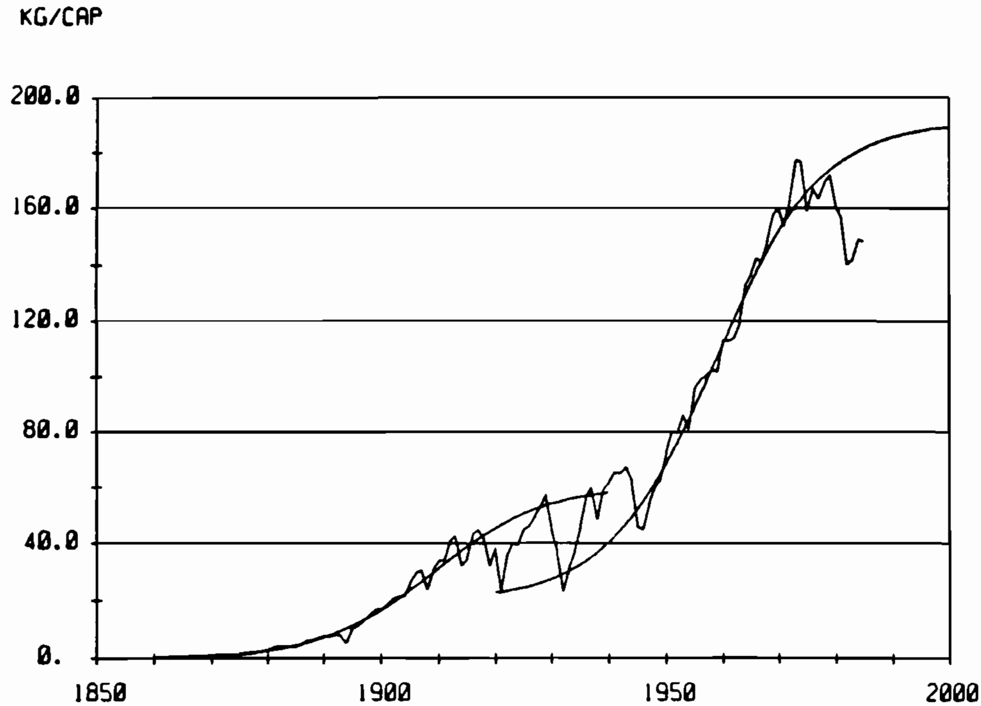


Figure 9.16. Global per capita steel production. (Source: Grübler, 1987.)

Instead, the saturation periods are more dispersed in time so that some occur before the depression years (e.g., canals in the United States) and some at the beginning of the upswing (e.g., saturation of open hearth steel in North America and West Europe).

9.4.2. Energy consumption

At the beginning of the nineteenth century, fuelwood, agricultural wastes, and mechanical wind and water power supplied most of the inanimate energy in addition to animal and human muscle power. We have seen that a considerable infrastructure of roads (turnpikes) and canals was in place for timber and later coal transport, although the widespread use of coal became possible with the emergence of railroads. Thus, like today, the use of energy in the early industrial development phase also depended on the transport system, and energy was an important component of goods transported on turnpikes, canals, waterways, and railroads. The development of energy and materials, and in particular steel technologies, is related in the same way as energy and transport systems. Better fuels made better steel processes possible, while higher-quality metals were crucial in further improving the whole energy system.

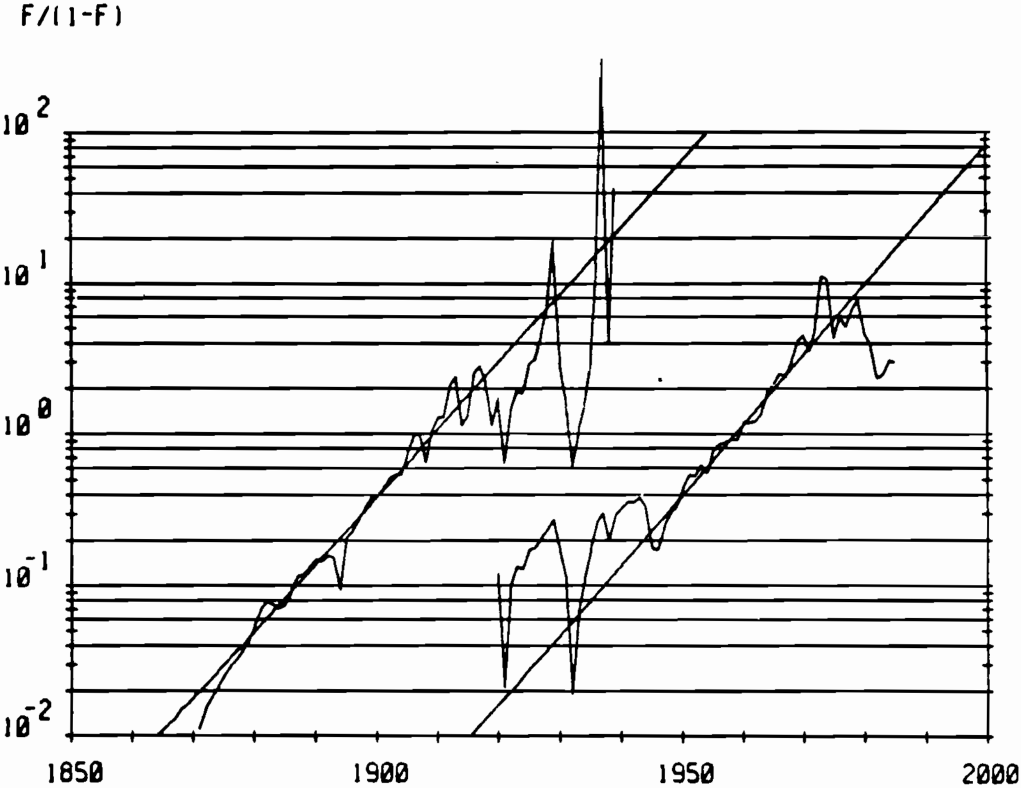


Figure 9.17. Global per capita steel production, transformed. (Source: Grübler, 1987.)

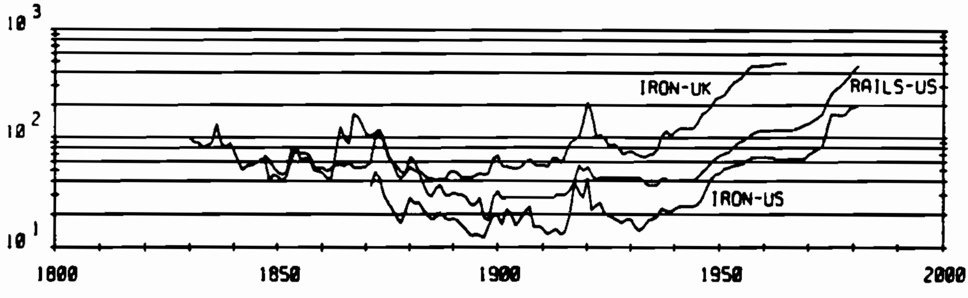


Figure 9.18. Irons and steel price indices, UK and USA. (Source: Grübler, 1987.)

Fuelwood represented most of the commercial primary energy inputs during the last century. *Figure 9.19* shows the annual consumption of fuelwood, fossil, and nuclear energy sources in the world since 1860. Data are plotted on a semilogarithmic scale and show the exponential growth phases in consumption by piecewise linear trends.

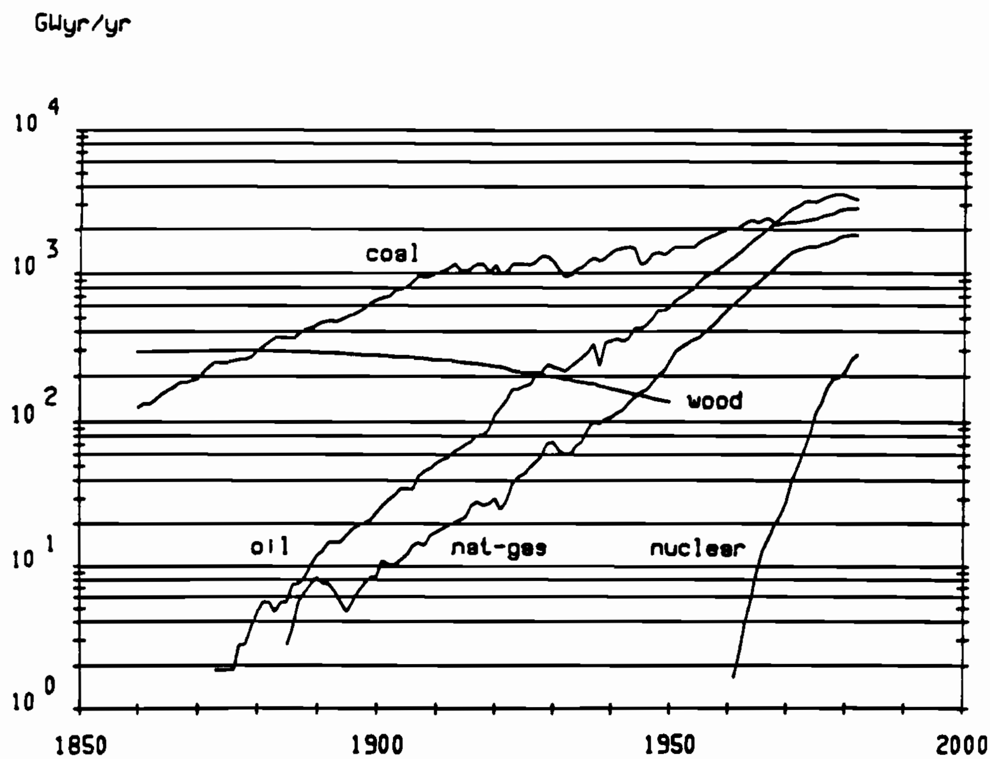


Figure 9.19. Global primary energy consumption.

Since the beginning of the century the consumption of fuelwood at the global level has declined as a commercial energy source, although it is still used widely, especially in the developing world. With the expansion of railroads and the steel industry, as well as the application of steam in general, the use of coal increased exponentially until the 1910s when a new, less rapid growth phase started. Since their introduction in the 1970s, oil and natural gas have been consumed at even more rapid rates. In fact, oil and natural gas curves have the same slope and thus almost identical growth rates; they are shifted in time by about 15 years. The increased use of oil and natural gas paralleled the growth of the petrochemical and electrical industries, and the expanded use of internal combustion and electric prime movers. Because nuclear energy is still in its early phase of development, the steep growth of the last two decades may not indicate the possibility of rapid expansion in the future. During the last few years, the growth of nuclear energy has declined worldwide to more moderate rates.

Primary energy consumption (including fuelwood) increased exponentially at an average growth rate of about 2% per year. The decline of older energy sources was more than compensated for by the rapid growth of the new ones. Thus, energy systems, like transport infrastructures and steel processes, evolved through a sequence of replacements of old by new technologies, practices, and

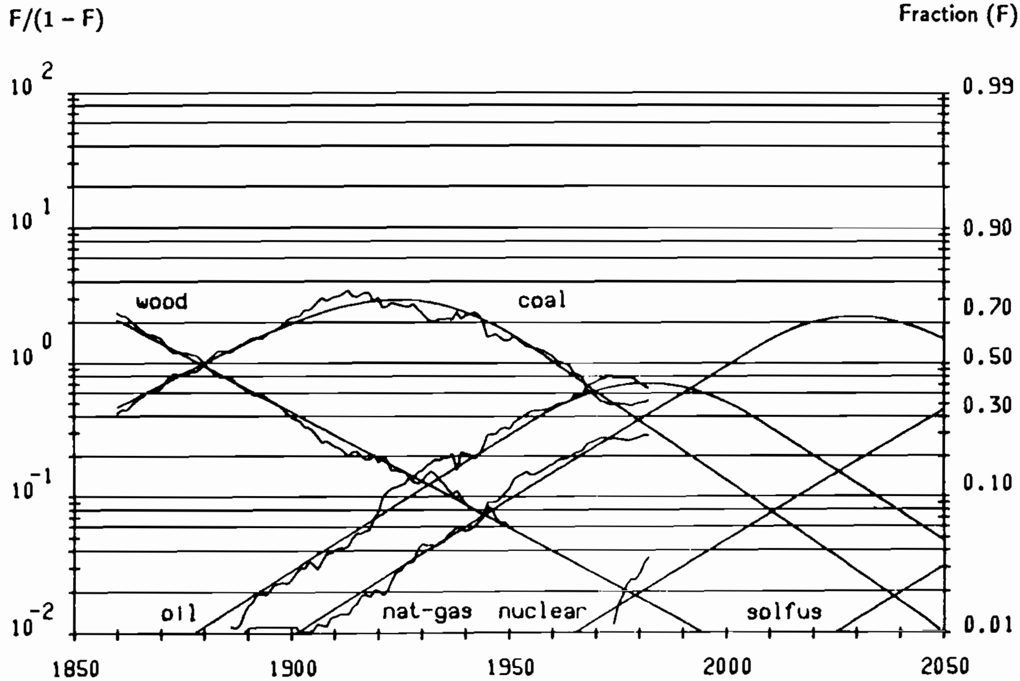


Figure 9.20. Global primary energy substitution.

methods. *Figure 9.20* shows the primary energy substitution process in terms of the fractional shares (F) held by each of the five energy sources in total consumption and plotted as $F/(1-F)$ on a semilogarithmic scale.

Compared to the substitution of transport infrastructures and steel technologies, the replacement of energy sources is a remarkably regular process. The slopes of the linear segments in the substitution process (logistic curves) are nearly the same, indicating that all four older energy sources have almost the same ΔT of about 100 years. Furthermore, the market shares do not reflect important historical events such as the world wars – the long-term trends are remarkably stable.

Based on these historical trends, we have used a scenario to project nuclear energy shares into the future. We have assumed the same slope as the expansion of oil and natural gas, implying a 5% market share by the year 2000. This indicates the possibility of a larger growth of nuclear energy in the next century, but also means *very few* additions to the current generating capacity during the next decades.

Coal saturated during the 1920s and oil during the 1980s. This again corresponds well to the timing of the last Kondratieff cycle. The growth phase until the 1920s is characterized by the expansion of the coal, railroad, and iron and steel industries, while the next growth pulse corresponds to the expansion of the oil, petrochemical, electricity, and road transport industries. These two growth pulses can be seen more explicitly at the aggregate level in total energy consumption. *Figure 9.21* shows the per capita global primary energy consumption divided into two growth pulses that reflect the substitution of primary energy sources. The first one was initiated with the rapid expansion in coal consumption after the 1860s and ends during World War II, by which time coal's share curved into a phase of decline. The second pulse is initiated with the onset of coal saturation and the beginning of the oil expansion phase (oil surpassed fuelwood in 1925) and accelerated after both fuelwood and coal were in decline. This second growth pulse is apparently nearing completion as the oil market shares in primary energy saturate.

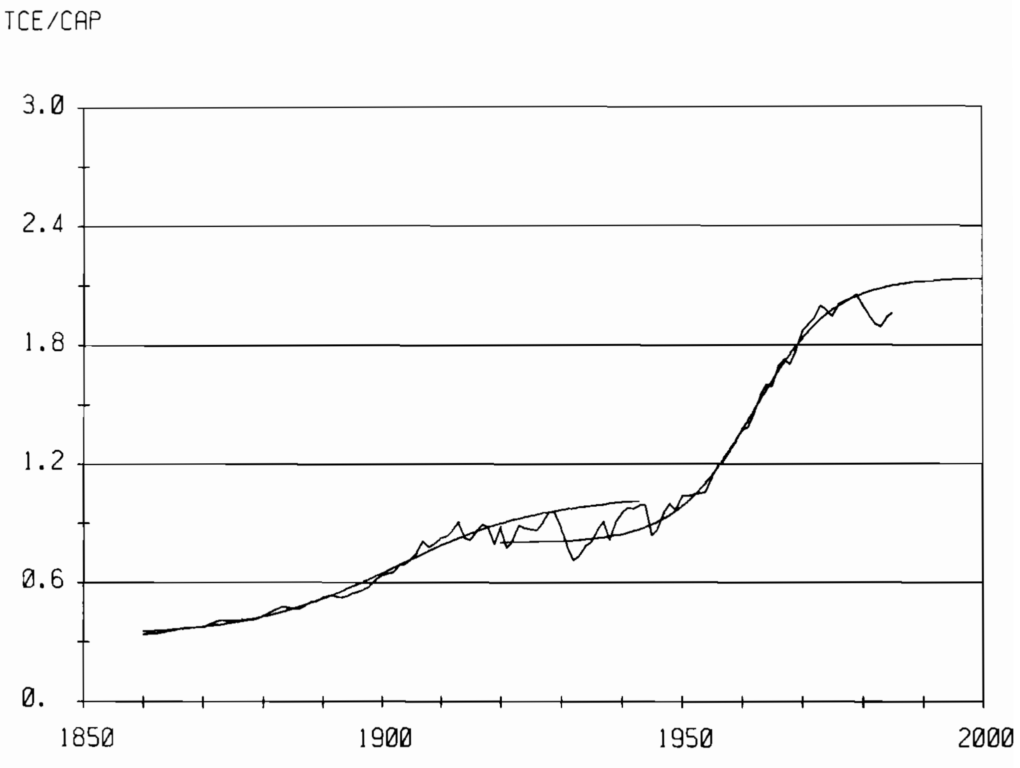


Figure 9.21. Global per capita energy consumption. (Source: Grübler, 1987.)

Figure 9.22 shows the same growth pulses transformed so that the two S-shaped curves appear as straight lines.[7] In this way the differences between the two growth pulses are more clear. The ΔT of the first pulse is longer than 60 years whereas it is shorter than 40 years for the second one. In contrast to this asymmetry the two growth pulses in per capita steel production had almost identical slopes and a ΔT of about 45 years (see *Figure 9.17*). In spite of these obvious differences between the energy and steel pulses, the analogy of the two processes is very strong and their timing almost identical. In fact, the deviations of the actual growth pulses from the estimated logistic curves show almost identical patterns. Fluctuations are strong at the beginning of the pulses until more than 10% of the saturation level is reached and they increase again above the 50% level. This result indicates that in addition to the regular pattern in the substitution of old for new technologies during the last two long waves, energy and steel use evolved through equivalent growth pulses that are concurrent with the upswing phases of the two last waves.

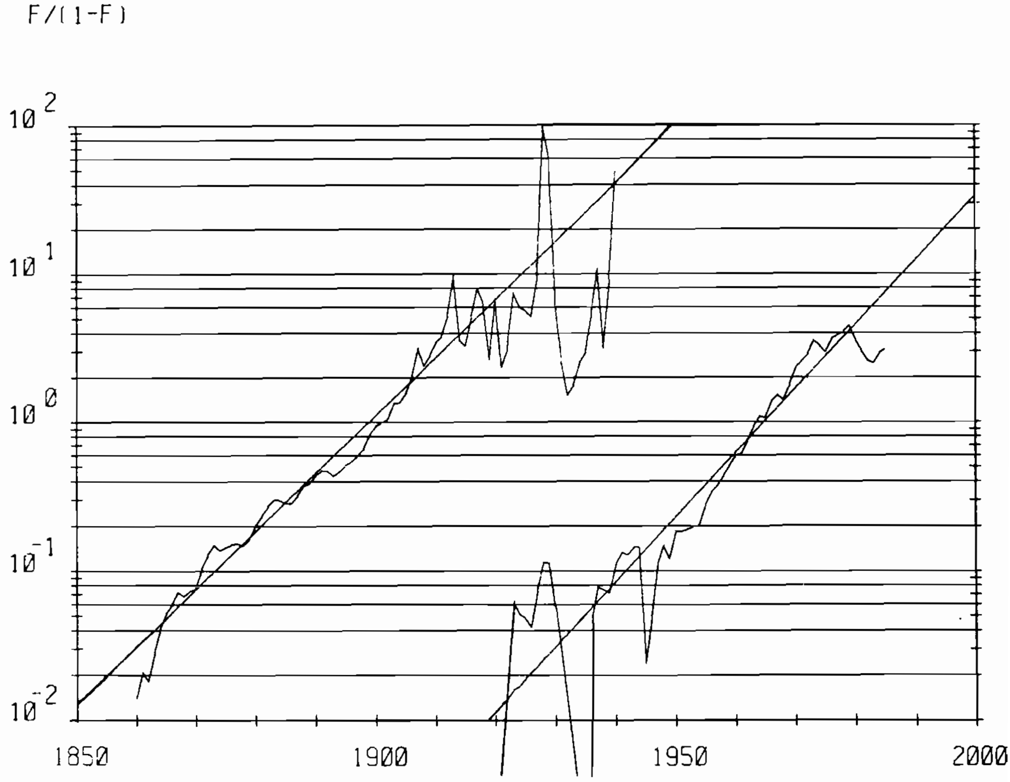


Figure 9.22. Global per capita energy consumption, transformed. (Source: Grübler, 1987.)

9.5. Energy and Prices

The above description of the complex process of technological and economic change is obviously incomplete. Certainly, there are many other, perhaps better, ways of describing the dynamics of transport infrastructures, steel production, and energy consumption. The intriguing aspect of the replacement dynamics and price fluctuations is that they appear to be interwoven with regular features and related to both the invariant pattern of substitution dynamics and the long waves in economic life. To show that this is not a feature unique to the last two long waves, we will show similar congruence in replacement dynamics of primary energy in the United States and price fluctuations over a period of about two centuries.

9.5.1. Primary energy

Energy use is one of the rare quantitative indicators that can, at least in principle, be compared over long periods of time in spite of many technological changes and substitutions of old for new sources of energy. This is possible because different energy sources can be expressed in common energy units. Fortunately, it is possible to reconstruct the history of the more important sources of energy for the United States since 1800, because a more complete record of energy consumption exists for this country than for the world. *Figure 9.23* shows the annual consumption of fuelwood, fossil energy, mechanical water power, and hydroelectric power in the United States since 1800. As in *Figure 9.19*, data are plotted on a semilogarithmic scale and show the exponential growth phases in consumption by piecewise linear trends. The growth of total energy consumption was on average about 3% per year in the United States. The general pattern in the evolution of the energy in the United States, however, is not different from that in the world as a whole. In a way this is not surprising since the United States is the largest energy consumer during this century.

Figure 9.24 shows primary energy substitution in the United States. Mechanical water power (mostly water and some windmills) and hydroelectric power are not plotted in the figure because of their small contribution to total energy supply: they barely exceeded the 1% level during short periods and were otherwise under that critical level. This shows that the omission of these energy sources due to the lack of data in the example of energy substitution at the global level is probably not too crucial, although the share of these noncommercial energy sources must be higher worldwide than in the United States.

Before the 1820s fuelwood fulfilled virtually all energy needs in the United States. Coal entered the competition in 1817 at the 1% level and until the 1880s it was essentially a two technology market – whatever gains coal made were translated into losses for fuelwood. Crude oil and natural gas were first used in the United States at the beginning of the nineteenth century, and both held a 1% market share during the 1880s. From then on the use of crude oil expanded, and by 1950 the consumption of crude oil surpassed that of coal. (Even as late as the 1920s, however, the consumption of crude oil was not much larger than that of fuelwood.) The use of natural gas surpassed the use of coal nine years later. In comparing this period with earlier periods, it is remarkable that the structure of energy consumption changed more during the period of oil dominance than ever before.

Natural gas exploration, production, and transport significantly indicate different trends from oil technologies, although natural gas was associated with the oil industry ever since its first commercial use. Nevertheless, most energy accounts bind natural gas to oil because of the large production of associated natural gas from oil wells. Except at the point of production, associated natural gas, or oil-technology gas, is indistinguishable from gas produced from natural

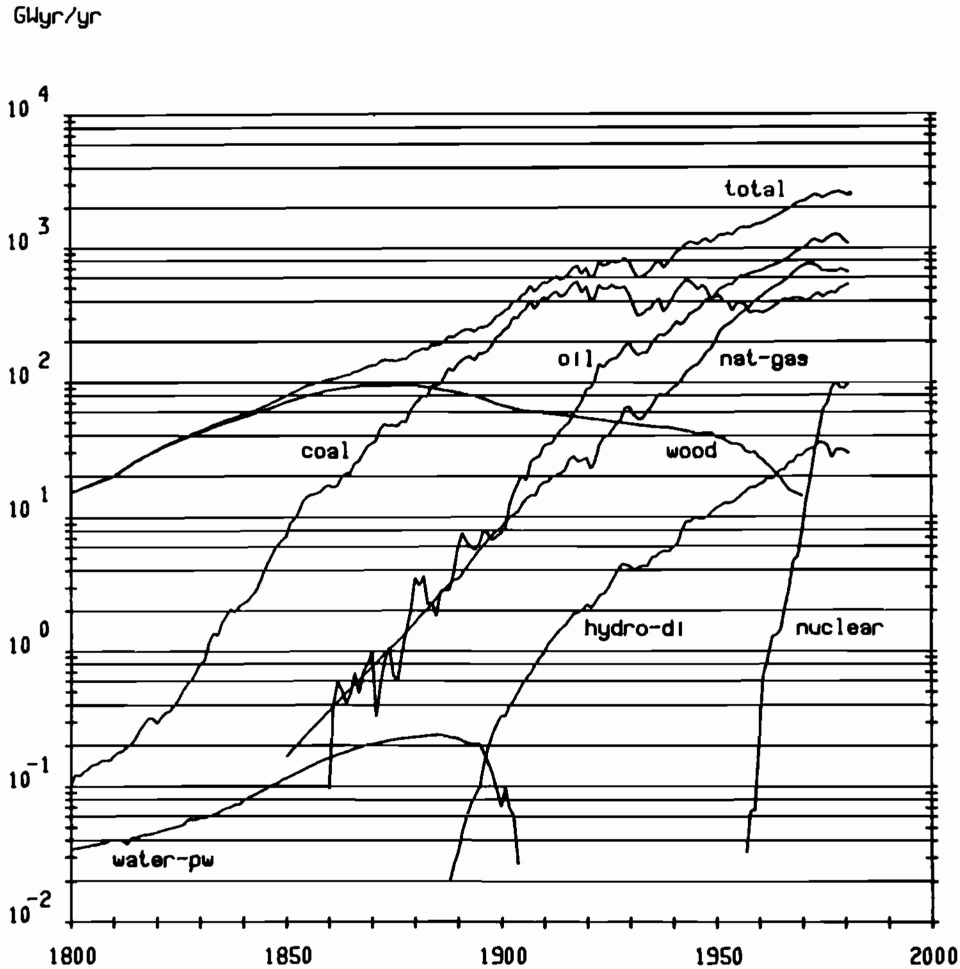


Figure 9.23. Primary energy consumption, USA.

gas wells. The fact that this distinction is difficult to make, and is consequently ignored in historical data, is to an extent misleading since oil and gas technology have portrayed distinctly different trends during the last century. This is however not reflected in oil and natural gas consumption data given in *Figure 9.23.* The distinction between associated gas and crude oil, in terms of primary energy accounting, is desirable and is consistent with the tradition of adding city gas produced from oil or coal to these primary energy sources rather than to natural gas.

To illustrate that natural gas is becoming increasingly uncoupled from oil technologies, we have attempted to reconstruct the primary energy balances by adding associated gas to crude oil and subtracting the same amount from natural gas consumption (but leaving net imports with natural gas balances). The revised market shares are given in *Figure 9.24.* The resulting replacements can

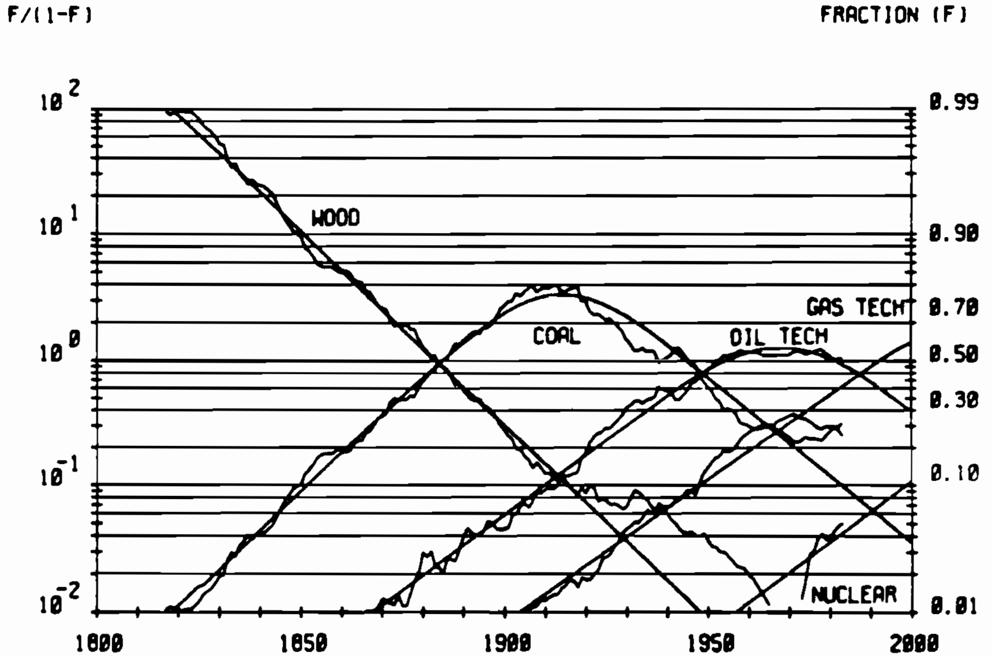


Figure 9.24. Primary energy substitution, USA.

be characterized by very regular time constants because the historical data are apparently accurate enough to provide the information required for this further analytical resolution. This is possible in the case of primary energy consumption in the United States because associated (oil technology) and nonassociated (gas technology) natural gas production is accounted for in the historical records.

This result shows that although associated gas has long been available as a by-product of oil, its use does not represent the actual evolution of gas technologies. *Figure 9.24* shows that the resulting substitution process improves the regularity to the extent that the time constants (ΔTs) now cluster at about 70 years for all energy sources and that the saturation intervals between coal, oil, and gas technologies are all separated by about 50 years. During the saturation periods of the dominant energy sources, new ones are introduced. Gas technologies are introduced during the saturation of coal, and nuclear energy during the saturation of oil. Thus, the evolution of the energy system reflects with perfect symmetry the long waves in prices and the associated periods of growth and structural change.

The slope of the nuclear energy penetration is determined by a scenario to have a ΔT of about 70 years by specifying a 10% market share by the year 2000. After the 1990s gas technologies clearly emerge as the most important energy source in the projection given in *Figure 9.24*. However, the natural gas (gas technology) shares have been below the trend line during the last few years while the coal shares have exceeded the projected market shares. It remains to be seen whether this disparity between our projections and actual development will be reabsorbed in the coming years like the "under- and overconsumption" of coal and oil, compared with trend lines during the 1920s and 1930s, which were eventually absorbed. It is conceivable that the uncoupling of oil and gas industries may provide a vehicle for wider use of natural gas in the future.

Although the saturation periods of coal and oil technologies are separated by about 50 years and during these periods of saturation new energy sources are introduced (gas technologies and nuclear energy, respectively), we still have yet to probe further into the past to test whether an even older energy source saturated during the previous long wave, i.e., during the 1870s.

Figure 9.25 shows that it is possible to include a partial reconstruction since 1850 of an even older energy source in the United States – animal feed. It represents an energy source equivalent to the amount of food consumed by the working animals (mostly horses and mules used in transportation and agriculture).

Animal feed reached its highest market share in the 1880s indicating that draft animals provided the major form of local transport and motive power in agriculture in spite of the dominance of railroads and steamships as long-

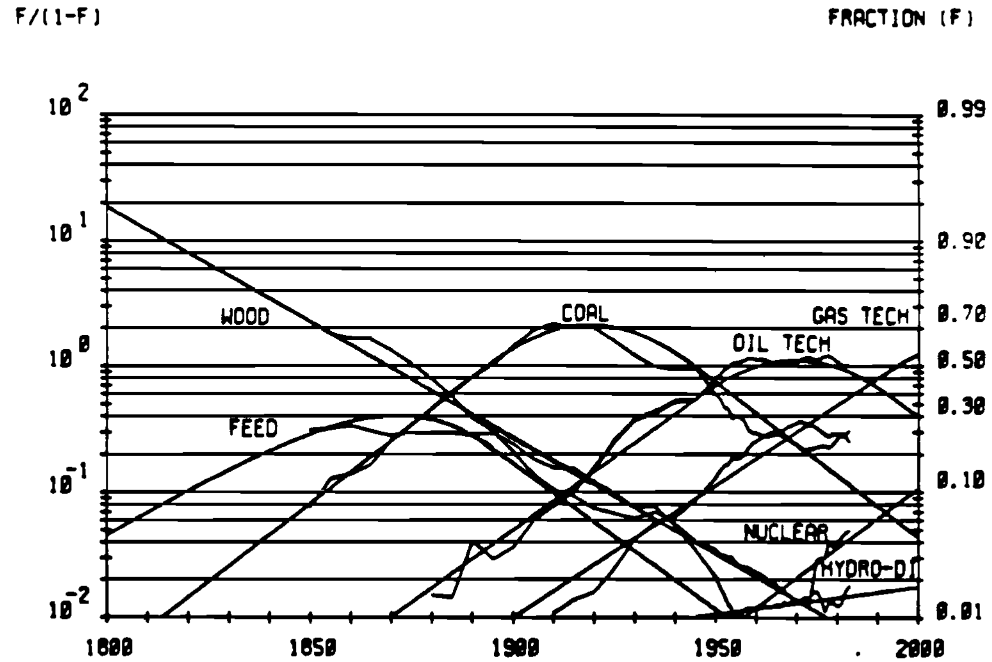


Figure 9.25. Primary energy substitution (with feed), USA.

distance transport modes (see the discussion on evolution of transport systems above). It is curious that the feed and oil technologies cross in the 1920s as if to suggest the simultaneous substitution of the horse carriage and wagon by the motor vehicles (see *Figure 9.8* and Nakicenovic, 1987b).

Figure 9.25 indicates that the replacement dynamics can result in a perfect symmetry of the successive substitutions of older for newer energy forms, providing a rather complete reconstruction of past patterns in energy use. Three emerging saturation periods are separated by about 50 years starting with animal feed in the 1880s, coal in the 1920s, and oil technologies during the 1980s. Each economic growth phase that connects the periods of energy saturation and prolonged recession throughout the economy is characterized by the expansion of two energy sources, one with large market shares "attacking" the saturating energy source and the other just emerging during the periods of saturation and structural change in the energy system. Next we will investigate another feature of the evolution of energy use in the United States that reflects the long waves in terms of the energy intensity of the whole economy.

9.5.2. Efficiency of energy use

There are many ways of determining the efficiency of energy use. The most obvious indicators are the efficiencies of primary energy conversion to secondary and final energy forms. Another possibility is to estimate the efficiency of energy end use. Examples include the amount of fuel needed for travel or for space conditioning. All of these efficiences have improved radically since the beginning of the Industrial Revolution along with the introduction of more efficient technologies. In some cases the improvements span almost an order of magnitude. For example, in 1920 the average efficiency of natural gas power plants in the United States was 9%, whereas today the best gas turbine power plants can operate with efficiences of almost 50%. Over longer periods the improvements are even more impressive. For example, the second law efficiency of prime movers increased by two orders of magnitude since 1700, that of lamps by almost three orders of magnitude during the last century, and so on (see Marchetti, 1979). All of these efficiency improvements of individual technologies are translated into more effective uses of energy and other materials at the level of the overall economic activity. Some efficiency increases result from improved technologies and others from substitution of the old for new technologies. In general, replacement follows when the saturation in additional improvements of an established technology is reached.

The extent of these changes and improvements can be expressed at an aggregate level by the amount of primary energy consumed per unit of gross national product in a given year. *Figure 9.26* shows the total primary energy consumption (from *Figure 9.23*), per capita consumption, and the ratio of energy consumption over gross national product (energy intensity) for the United

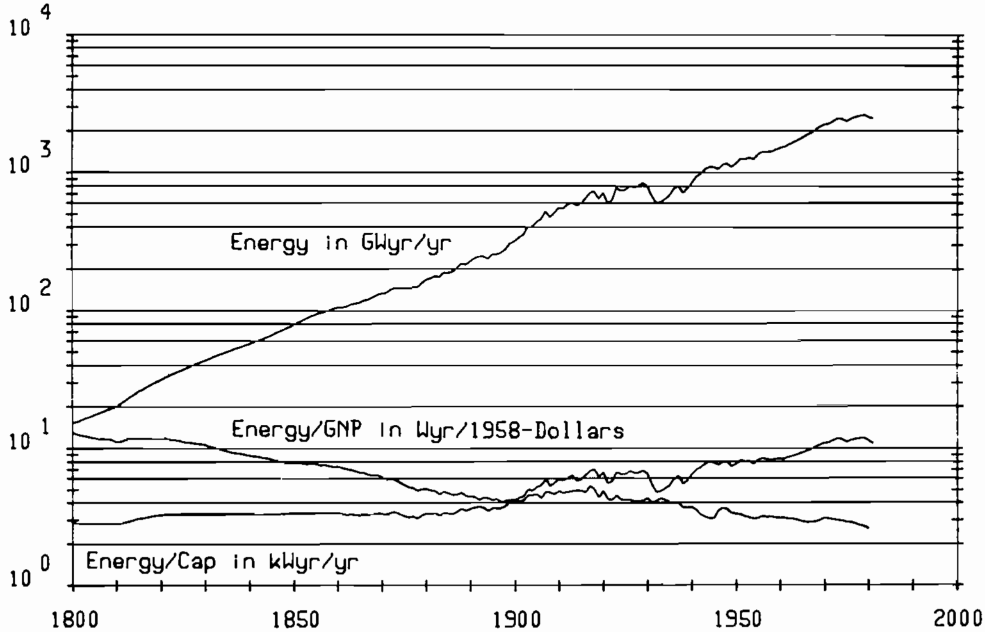


Figure 9.26. Primary energy, gross national product, and energy intensity, USA.

States. The average reduction in energy consumed to generate one dollar of gross national product was about 0.9% per year during the last 180 years. The ratio decreased from 10 kilowatt years per (constant 1958) dollar in 1800 to slightly more than two kilowatt years per dollar in 1982. Thus, a regular decline in energy intensity of the whole economy prevailed over a long historical period indicating that energy conservation is a historical replacement process that was discovered as a concept only during the last decade.

Figure 9.27 shows the fluctuations in energy intensity in the United States after the elimination of the secular trend by a 51-year geometric moving average. The fluctuations show pronounced long-wave movements and a high degree of synchronization with the price swings (see *Table 9.1*). *Figure 9.28* shows the evolution of energy prices in comparison with the wholesale price index from *Figure 9.4* and the heat and lighting price index (the last two indices are almost indistinguishable from each other), and in *Figure 9.29* the oil prices are also shown as an index to indicate the concurrent changes in energy and wholesale prices.

During the downswings in prices the energy intensity of the economy decreased more rapidly and during the upswings less rapidly. This illustrates relatively high energy price elasticity since energy prices changed in unison with general price movements as shown in *Figure 9.29*. This means that during the downswing in economic activity general rationalization measures of individual enterprises cause larger energy savings compared with the average historical reductions.

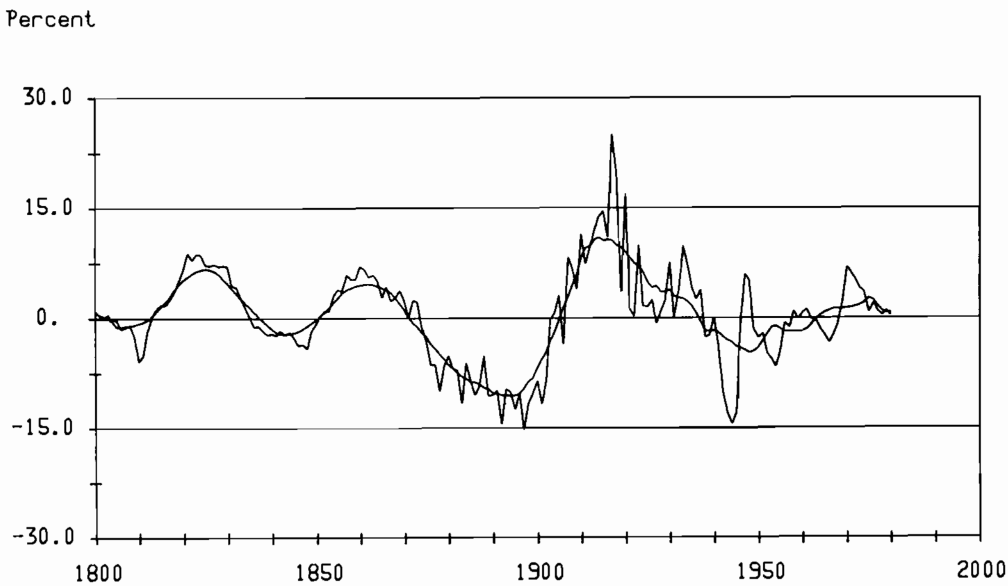


Figure 9.27. **Long waves in energy intensity, USA.**

As the competition intensifies during the recession and depression, energy savings become an important factor in cost reduction, also because of generally higher energy prices in addition to the overall price inflation. With recovery, new demands and prospects of continued economic growth release many pressures associated with saturating markets. Price levels are also much lower at the beginning of the new growth phase. Most entrepreneurs in the new growth sectors must intensify their activities to meet new demands, and low energy intensity ceases to be an important competitive criterion. New technologies and energy forms offer possibilities for continued expansion in new markets so that relative energy use intensifies. Toward the end of the prosperity period the growth process encounters limits once more. These are reflected in saturating demand and general price inflation illustrated by the long wave of wholesale price movements (see *Figure 9.4*). Thus, during the downswing energy use reductions become important.

These reductions are not only due to efforts to cut costs as a reaction to saturating demand, but also due to a host of social constraints. Many energy technologies, along with other economic activities, become socially and environmentally unacceptable toward the end of prosperity. This means that some diseconomies that were socially acceptable during the growth phase become internalized as additional economic costs or as explicit limits to further expansion. These causes of additional costs appear to offset the benefits of the economies of scale achieved during the expansion phase. In fact, with the demand reductions during the downswing the large capacities that offered economies of scale become sources of additional costs as excess capacity.

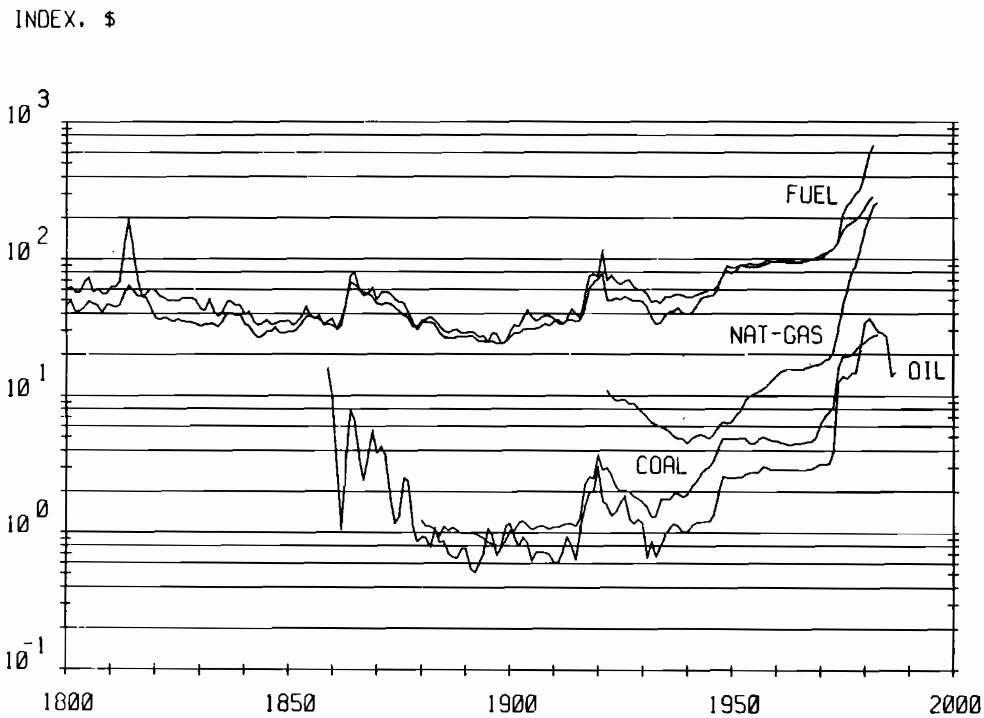


Figure 9.28. Energy and wholesale prices, USA.

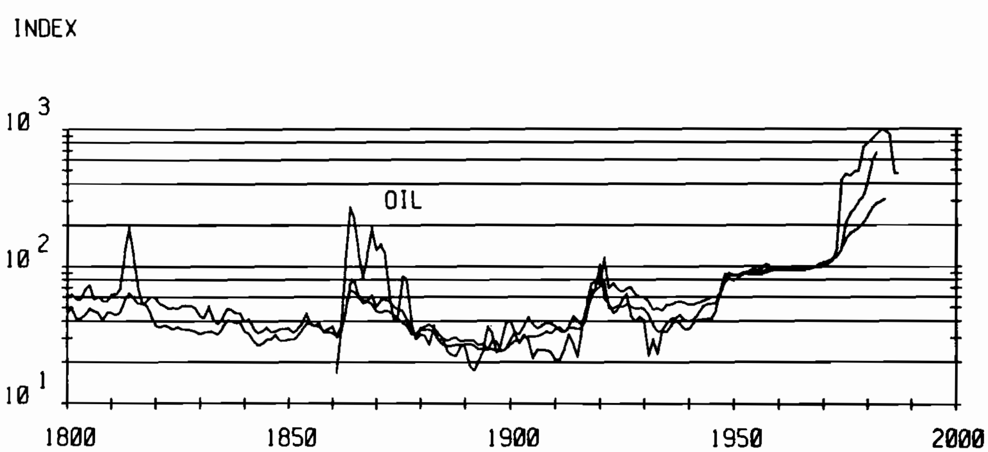


Figure 9.29. Oil – heat and lighting – and wholesale price indices, USA.

The relationship between primary energy consumption patterns and the long wave appears to extend beyond the parallel changes in the per capita level of energy consumption and energy intensity with the fluctuations of other long-wave indicators such as wholesale prices. Comparing *Figure 9.25* with *Figure 9.27* indicates that the upper turning points of energy intensity fluctuations correspond to the saturation points of primary energy sources. The upper turning point that occurred in the 1870s is related to the saturation in animal feed substitution, the 1915 turning point with the saturation in coal substitution, and the turning point of the 1970s with the saturation of crude oil. In addition, new energy sources reached 1% market shares during the times of *high* energy intensity (during the 1870s, 1910, and 1973). Thus, the succession of the long waves indicates a similar timing as the dynamics of energy substitution and the changes in steel production and transport infrastructures.

9.6. Dynamics of Change

At the risk of generalizing, we can state that there is strong evidence that symmetric or at least similar changes in the patterns of energy consumption, steel production, transport infrastructures, and price *niveau* occur from one long wave to another, although the historical content and individual manifestations change profoundly so as to make the symmetry apparent only at the higher level of abstraction. Thus, only patterns of change are similar while the symmetry breaks with an attempt to relate individual events that are always unique to a particular historical situation and thus different from one long wave to another. This is the essence of the analysis presented here – recurring patterns emerge as we consider sequences of replacements of old by new technologies, growth pulses, and deviations from smoothed secular trends. In all three cases the individual distinctions are removed from the data. It is the individual distinctions that "condense" the time series to actual (original) indicators removing the symmetry in patterns among periods of growth and periods of saturation and change.

To understand the actual mechanisms behind the long-wave phenomenon and changes in technology, economy, and society, we must acquire better analytical and statistical descriptions of the causal relationships we generally call experience. This would also imply that we need to understand the course of specific events and their individual manifestations that led, for example, from a period of rapid growth after World War II to the oil shocks of the 1970s, saturating world markets, changing industrial structure, increasing national debt in many quarters of the world, and developing the economic slowdown of the last decade. For the time being we can only observe that the particular circumstances change from one long wave to another, but that the sequence of fluctuations and structural changes at a higher level of abstraction indicate a striking regularity. Thus, the symmetry is destroyed in the transition from a sequence of replacements and fluctuations to particular circumstances and individual events. In other words, the individual event, such as an innovation, is unique, but the pattern of change appears to be regular for the emergence of innovations taken together as a dynamic process.

The annals of business cycles (see, for example, Thorp and Mitchell, 1926; Mitchell, 1927) show that the severe crises or so-called Great Depressions occur regularly during the downswing of the long waves. It suffices here to mention that the Great Depressions and financial panics of 1819, 1874, and 1929 in the United States with small variance occurred throughout the rest of the world. This immediately suggests an obvious historical manifestation of the prolonged periods of stagnation, but this does not answer the question whether these Great Depressions are a necessary characteristic of the downswing.

The analysis of technological substitution in steel production, energy consumption, and transport infrastructures showed that the same basic approach can be applied to describe the structural changes. In all three cases older technologies were replaced by new ones with regular recurring patterns. Besides the now obvious similarity in the substitution patterns, it should be observed that the timing of the saturation phases is to some degree synchronized in the three examples.

In all three cases technologies that have saturated before 1850 (canals, crucible and puddle steel, and fuelwood) are declining, although at different rates. The next "wave" of technologies to reach saturation in terms of market shares between 1870 and 1875 are railways, Bessemer steel, and animal feed as a source of energy. During the 1920s coal saturated, in 1950 open hearth steel, in 1970 length of roads and oil technologies, and in 1980 oxygen steel. This indicates that the saturation periods measured in terms of relative market shares are not perfectly aligned, but that they are grouped around the upper turning points in the long wave as indicated by price fluctuations (see *Figure 9.4*). What is more important however is that during period of saturation of some technologies and prolonged economic recession, the next generation of growing technologies surpasses the declining ones, e.g., crucible and puddle steel by open hearth process, canals by roads, fuelwood by oil and gas, railroads by airways, and Bessemer by electric and oxygen steel. This illustrates that the periods of prolonged recession are indeed periods of structural change when dominating technologies are either close to saturation or already declining and when older declining technologies are surpassed by the emerging ones. Another way of stating this symmetry is that the dominating technology in each example bridges the period between the end of growth in one wave and the beginning of growth in the next as its market shares curve from increase through saturation to decline.

A possible explanation of this similarity in the substitution patterns is that the specific changes that led to the replacement of old by new technologies and energy sources were interrelated. For example, the new steel processes and transport infrastructures were dependent on new energy technologies. On the other hand, the new energy sources could only be developed with increased intensity of energy use, such as in the new industrial and urban complexes that emerged as the availability of transport possibilities and basic materials increased (symbolized here by transport infrastructures and steel production). This kind of interdependent lacing of technological development and growth of demand indicates that a certain degree of synchronization in the substitution processes could be expected. This of course still leaves the question of the precise nature of the 50-year time constant unanswered. Since we have already shown that the three substitution processes appear to portray similar features and relatively close the timing of crucial market saturation and takeover events, we will now consider the timing of long-wave fluctuations and energy substitution.

Figure 9.30 shows energy substitution (from *Figure 9.25*) on the lower plot and long wave in energy consumption, energy intensity, and wholesale prices (from *Figures 9.4* and *9.27*) on the upper plot. Here we have implicitly assumed that energy substitution is indicative of other replacement processes because the timing of saturation periods, introduction of new energy sources, and ΔTs are perfectly symmetrical in this example. Thus, *Figure 9.30* summarizes the results of the phenomenological analysis of the dynamics of technology and long waves. A careful examination of the timing and patterns of changes shows that they are all in tune. The saturation periods of energy technologies coincide with the peaks in prices and energy intensity. The period of decline from saturation to loss of dominance (i.e., loss of the highest market shares) lasts on the order of 25 years, or about as long as the downswing phase of the long wave, which is characterized in *Figure 9.30* by the fluctuations of energy consumption, intensity, and the price index. By symmetry, the upswing of the long wave is paralleled by the growth of the new energy source from newly acquired dominance to saturation.

9.7. Conclusions

A number of important technologies and growth sectors emerge within relatively short periods of time: they expand because they are interrelated, and they enhance each other because they are interlaced and interdependent. Long waves emerge because the replacement of old technologies and methods and the growth of new ones take place simultaneously, leading to a prolonged phase of economic growth. Likewise, these replacements and growth processes tend to saturate during relatively short periods of time resulting in prolonged recession and turbulence, restructuring of the economy, and technological and social change.

The prolonged recession period reduces economic opportunities, and also leads to excess production capacities, unemployment, and lower intensity of other factor inputs. This change in the economic environment from relatively continuous growth to stagnation, recession, and turbulence makes the application of financial innovations attractive as other opportunities for return on capital decline. The prolonged stagnation and recession period is characterized by volatility and expansion of speculative instruments and decline in productive investments. As volatility increases so do the risks, and it is conceivable that this diverts some capital and human resources into innovative activities.

Some evidence exists that innovations cluster around the recession periods and eventually lead to a new "wave" of replacements and growth. This basic pattern can recur and be repeated in completely different circumstances provided that the growth of a number of important technologies is to an extent synchronized. All that is necessary is a distinction between periods of growth and periods

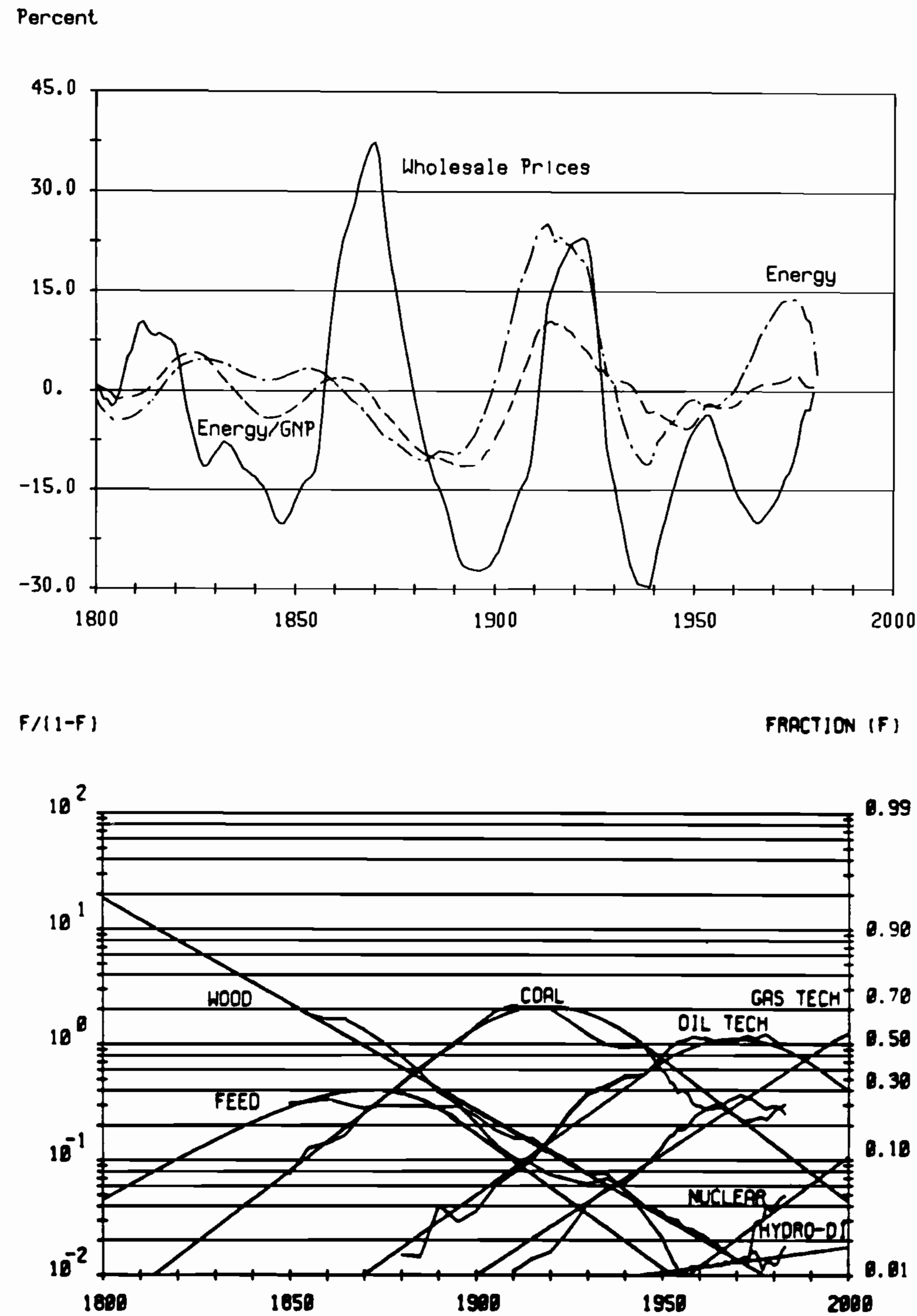


Figure 9.30. Long waves and substitution dynamics, USA.

of change. This hypothesis does not require a high degree of synchronization or "focusing" for either the innovation clusters or the saturation clusters. It is sufficient that a number of important replacements take place and cause a period of prolonged growth. Certainly this is a very simplistic and stylized scheme of technological change and economic and social development, but it is supported by the empirical evidence of technological substitution in energy, steel, and transport and in the overall price movements. This evidence supports the existence of weak clusters in the growth phases of different technologies, and perhaps weaker evidence for saturation clusters. A larger "sample" of case studies would be required to confirm this result in the statistical sense. For the time being it is not clear whether the phenomenological evidence is representative enough to allow generalization of the results or not. The example of primary energy substitution in the United States (see *Figure 9.25*) is remarkable in that is portrays a perfect symmetry of introduction, growth, and saturation of technologies and price movements over a period of almost two centuries.

The fact that some events that characterize profound changes in technology and economic structure occur in tune is striking, but it leaves many questions open. For instance, we have observed that technological replacements in steel production and transport infrastructures do not portray perfectly symmetrical patterns as primary energy substitution. This means that technologies saturate in broad waves that start at the end of the prosperity phase and extend to the beginning of the next Kondratieff wave. Perhaps this is an artifact of the choice of technological substitution processes in that they are very closely related to each other and represent the most important growth sectors. Yet, given the sparse statistical records, it is difficult to find other examples that span equivalent historical periods.

Nevertheless, the importance of the energy system and related infrastructural developments appears to be crucial with respect to the observed pulses in economic activity. For example, the construction of great canals throughout Europe and the United States during the eighteenth and beginning of the nineteenth centuries was initiated by the ever increasing need to transport timber and other goods in larger quantities over longer distances. Later, railroads were associated with a similar boom period basically due to the same reasons – the concentration of production in urban areas required a more efficient transport system that also helped in the acquisition of new and larger markets. Thus, canals and railroads expanded existing markets and "created" new ones for new products. In terms of the energy system, large canals were associated with the transport of fuelwood, which at that time was the primary source of energy for many industrial activities such as iron smelting. The railroad era was very closely related to the widespread diffusion of steel, steam, and coal-related industries.

In terms of the long-wave fluctuations, we will name the upswing phase from 1773 to 1810 the "age of canals" and the upswing from 1840 to 1869 the "age of railroads." Accordingly, we name the upswing from 1895 to 1920 the "age of electricity" because of its significant contribution to the rapid development of new industries and communication technologies. The last upswing, from 1945 to the 1970s, we symbolically identify with motor vehicles, roads, aircrafts, and petrochemical industries. Unfortunately, it is not possible to time this last turning point with any precision, but in view of the empirical evidence in the synchronization of technological substitution processes, energy efficiency, and other indicators, it probably occurred after the "oil crises" of the early 1970s, which marked the saturation of crude oil and its eventual replacement as the dominant source of primary energy. Let us assume, for the sake of naming a particular reference year, that it occurred in 1973. If this were actually the case, and assuming the continuation of the long-wave fluctuations, the next turning point could be expected sometime around the turn of the century. Going further into the future on the basis of this scheme, the next upswing phase could be expected to last until the 2030s.

The overall picture that emerges suggests that each upswing phase is associated with large infrastructural development. This development first opens many new product and factor markets and toward the end of the prosperity phase leads to eventual saturation of these markets and full adoption of the technologies that were introduced during the recovery period. This was the process that occurred between the end of World War II and the initiation of a downswing five to ten years ago. We can already anticipate some developments of the current downswing period. For example, the energy intensity curve in *Figure 9.27* indicates that during this and the next decades we can anticipate further relative improvements in the energy efficiency of the economy (i.e., reductions in the amount of primary energy consumed per monetary unit of gross national product in real terms). Thus, we can expect further dissemination of energy efficient technologies and institutional measures during the downswing phase until the end of the century. As far as energy technologies are concerned, the market penetration analysis suggests natural gas as the best candidate for eventual dominance as the major energy source during the upswing period after the 1990s. Natural gas is the cleanest fossil fuel and from that perspective alone it is attractive. It could also become a very efficient source of electricity and clean fuels. Widespread use of natural gas would require new infrastructures for production, long-distance transport (e.g., by pipelines or by superconductive cables via electricity), conversion to fuels and electricity, and distribution to the final consumer. Thus, construction of large grids and new industries based on natural gas would be required.

Another growth sector discussed in terms of technological change are transport systems and infrastructures. One possible development could be an advanced aircraft with substantively higher productivity than the Boeing 747. It could be a super large subsonic transport, a cruise supersonic, or a hypersonic spacecraft. In any case, such advanced means of transport would require new infrastructures, such as airports and "feeder" aircraft, and other transport modes to and from airports. Cruise supersonic and hypersonic transports would also require new sources of energy and new materials. Methane and hydrogen are obvious candidates to replace kerosine, so that the coupling to natural gas is obvious. It is analogous to the simultaneous development of the petrochemical and automobile industries. New materials are related to a possible replacement for the saturating steel industry, which in the long run will have to expand to other materials as recycled steel and the electric process exhaust their market potentials.

These are just some possible candidates, but they are consistent with the apparent requirements that emerge from the overall pattern of economic pulses and technological substitution dynamics since the beginning of the Industrial Revolution. Before these and other new technologies could expand during the next upswing, the next decades would bring a period of renewal and "gales of creative destruction" (Schumpeter, 1935). A period of rapid (relative) deflation can be expected together with prolonged unemployment and further economic slowdown. These are the selection mechanisms that in the past distilled the successful from a wide range of promising new technologies and entrepreneurial innovations. The existing patterns will have to be destroyed before new ones can emerge and their destruction will mark the beginning of renewal and a promise of prosperity.

Most speculations about the nature and timing of future events are based on the dynamics of equivalent changes in the past. The perfect symmetry between changes in the primary energy and price fluctuations shows that certainly price mechanism alone cannot explain all of the dynamic changes. The secular trends in prices are however consistent with the structural changes in energy. There is an intertemporal price elasticity of energy use. As prices increase energy is saved (intensity decreases) and new energy sources are introduced. In some way prices appear to have short range as a "force" in a dynamic economy. Savings are introduced in the face of a recession, but they only create unemployment and excess capacities. It is the innovations that can "tunnel" through the Kondratieff "barrier" and create a new period of growth. From a large number of innovative activities only a few are successful, most cannot escape from the recession and decline. In the past the few successful innovations were important enough to create a new period of growth. Thus, from the dynamic perspective prices appear to have a shorter range while innovations have little importance in the short run, but are of fundamental importance over periods longer than the business and the inventory cycles.

This illustrates why the Kondratieff wave is a long cycle, but perhaps the most important question is why the clock that times such events as the dynamic changes in technology and long waves in economic activity operates on a 50-year scale. Since we have shown that the events that mark structural changes are synchronized and follow a logical order, the question of the time scale and invariance is crucial. If it were answered all the other events, since they occur in logical order apparently as required, would fit the grand pattern like pieces of a puzzle.

Notes

[1] One general finding of a large number of studies is that many growth processes follow characteristic S-shaped curves. Logistic function is one of the most widely applied S-shaped growth curves and is given by:

$$x/(\kappa-x)=\exp(\alpha t+\beta),$$

where t is the independent variable usually representing some unit of time, α,β, and κ are constants, x is the actual level of growth achieved, while $\kappa-x$ is the amount of growth still to be achieved before the (usually unknown) saturation level κ is reached. Taking logarithms of both sides gives the left side of the equation to be expressed as a linear function of time so that the secular trend of a logistic growth process appears as a straight line when plotted in this way. Substituting $F=x/\kappa$ in the equation expresses the growth process in terms of fractional share F of the asymptotic level κ reached, i.e., the equation becomes:

$$F/(1-F)=\exp(\alpha t+\beta).$$

[2] One general finding of a large number of studies is that substitution of an old technology for a new one, expressed in fractional terms, follows characteristic S-shaped curves. Fisher and Pry (1971) formulated a simple but powerful model of technological substitution by postulating that the replacement of an old by a new technology proceeds along the logistic growth curve:

$$F/(1-F)=\exp(\alpha t+\beta),$$

where t is the independent variable usually representing some unit of time, α and β are constants, t is the fractional market share of the new competitor, while $1-F$ is that of the old one.

[3] We define ΔT as the time elapsed between the achievement of one and 50% of the saturation level κ, i.e., in this example $\Delta T = 95.5$ years. Because of the symmetry of the logistic function, the same time is required for the increase from 50 to 99% of the saturation level. An alternative definition of ΔT is the time elapsed between the achievement of 10 and 90% level. In this case ΔT would be slightly different from the other definition, but for all practical applications both definitions can be used interchangeably.

[4] The fractional shares (F) are not plotted directly but as the linear transportation of the logistic curve, i.e., $F/(1-F)$ – in this more general case, as the ratio of the market share taken by a given transport infrastructure over the sum of the market shares of all other competing infrastructures. This form of presentation reveals the logistic substitution path as an almost linear secular trend with small annual perturbations. Thus, the presence of some linear trends in *Figure 9.13* indicates where the fractional substitution of transport infrastructures follows a logistic curve. In dealing with more than two competing technologies, we must generalize the Fisher and Pry model, since in such cases logistic substitution cannot be preserved in all phases of the substitution process. Every competitor undergoes three distinct substitution phases: growth, saturation, and decline. This is illustrated by the substitution path of rail tracks, which curves through a maximum from increasing to declining market shares (see *Figure 9.13*). In the model of the substitution process, we assume that only one competitor is in the saturation phase at any given time, that declining technologies fade away steadily at logistic rates, and that new competitors enter the market and grow at logistic rates. As a result, the saturating technology is left with the residual market shares (i.e., the difference between 1 and the sum of fractional market shares of all other competitors) and is forced to follow a nonlogistic path that joins its period of growth to its subsequent period of decline. After the current, saturating competitor has reached a logistic rate of decline, the next oldest competitor enters its saturation phase

and the process is repeated until all but the most recent competitor are in decline. A more comprehensive description of the model and assumptions is given in Nakicenovic (1979).

[5] As in *Figure 9.13*, the fractional shares (F) are not plotted directly but as the linear transformation of the logistic curve, i.e., $F/(1-F)$ – as the ratio of the market share taken by a given steel technology over the sum of the market shares of all other competing technologies. Also in this figure, this form of presentation reveals the logistic substitution path as an almost linear secular trend with annual perturbations. The presence of some linear trends in *Figure 9.15* indicates where the fractional substitution of steel technologies follows a logistic curve.

[6] Transformation $F/(1-F)$ is used where $F = x/\kappa$, κ is the estimated saturation level and x the steel production in a particular year, so that F represents the fraction of the saturation.

[7] Transformation $F/(1-F)$ is used where $F = x/\kappa$, κ is the estimated saturation level and x the energy consumption in a particular year, so that F represents the fraction of the saturation.

References

Bianchi, G., Bruckmann, G., and Vasko, T., eds., 1983, ***Background Material for a Meeting on Long Waves, Depression and Innovation***, Siena/Florence, 26–29 October 1983, CP-83-44, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Bieshaar, H. and Kleinknecht, A., 1984, Kondratieff Long Waves in Aggregate Output? An Econometric Test, ***Konjunkturpolitik*** **30**(5): October.

van Duijn, J., 1983, ***The Long Wave in Economic Life***, George Allan and Unwin, London, UK.

Fisher, J.C. and Pry, R.H., 1971, A Simple Substitution Model of Technological Change, ***Technological Forecasting and Social Change*** **3**:75–88.

Grübler, 1987, private communication, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Hartman, R. and Wheeler, D., 1979, Schumpeterian Waves of Innovation and Infrastructure Development in Great Britain and the United States: The Kondratieff Cycle Revisited, ***Research in Economic History*** **4**:37–85.

Kleinknecht, A., 1987, ***Innovation Patterns in Crisis and Prosperity: Schumpeter's Long Cycle Reconsidered***, Macmillan, London, UK, and St. Martin's Press, New York, NY, USA.

Kondratieff, N.D., 1926, ***Die langen Wellen der Konjunktur***, Archiv für Sozialwissenschaft und Sozialpolitik, Verlag Von J. C. B. Mohr, Band 56:573–609, Tübingen, Germany, F.R.

Kondratieff, N.D., 1979, ***The Long Waves in Economic Life***, Review (Binghampton), Spring 1979: 519–62.

Lee, T.H. and Nakicenovic, N., 1987, ***Technology Life Cycles and Business Decisions***, Draft, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Marchetti, C., 1977, Primary Energy Substitution Models: On the Interaction Between Energy and Society, ***Technological Forecasting and Social Change*** **10**:345–356.

Marchetti, C., 1979, Energy Systems – The Broader Context, ***Technological Forecasting and Social Change*** **14**:191–203.

Marchetti, C., 1981, *Society as a Learning System: Discovery, Invention, and Innovation Cycles Revisited*, RR-81-29, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Marchetti, C. and Nakicenovic, N., 1979, *The Dynamics of Energy Systems and the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Mensch, G., 1979, *Stalemate in Technology*, Ballinger, Cambridge, MA, USA.

Mitchell, W., 1927, *Business Cycles: The Problem and Its Setting*, National Bureau of Economic Research, Inc, New York, NY, USA.

Montroll, E. W., 1978, *Social Dynamics and the Quantifying of Social Forces*, Proceedings of the National Academy of Sciences, Vol. 75, October 1978, Applied Mathematical Sciences, Washington, DC, USA.

Nakicenovic, N., 1979, *Software Package for the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Nakicenovic, N., 1984, *Growth to Limits, Long Waves and the Dynamics of Technology*, International Institute for Applied Systems Analysis, Laxenburg, Austria,

Nakicenovic, N., 1987a, *Transportation and Energy Systems in the U.S.*, WP-87-01, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Nakicenovic, N., 1987b, *The Automotive Road to Technological Change: Diffusion of the Automobile as a Process of Technological Substitution*, RR-87-1, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Pearl, R., 1924, *Studies in Human Biology*, Williams and Wilkins, Baltimore, MD, USA.

Schumpeter, J.A., 1935, The Analysis of Economic Change, *Review of Economic Statistics* **17**:2–10.

Schumpeter, J.A., 1939, *Business Cycles, A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, Vols. 1 and 2, McGraw-Hill, New York, NY, USA.

Stewart, H., 1981, *Transitional Energy Policy, 1980–2030*, Pergamon Press, New York, NY, USA.

Thorp, W. and Mitchell, W., 1926, *Business Annals*, National Bureau of Economic Research, Inc., New York, NY, USA.

Verhulst, P.F., 1844, Mem. Acad. R. Bruxelles **28**:1.

CHAPTER 10

Techniques and Labor in Long-Term Fluctuations: A Study of Underground Transport in Mines

Louis Fontvieille and Anita Prigent

10.1. Introduction

After the 1973 crisis, the return to Kondratieff was generally effected with reference to Schumpeter under the stimulus of Mensch in West Germany and then of Freeman in England. Mensch (1979) studied the lists of innovations produced during the history of science and technology and showed that innovations appear to be more numerous during the depression phase of the long cycle. The flood of innovations produced would appear to overcome the depression by stimulating investment.

A few years later, Freeman's team (1982) studied the relationship between technical innovations and unemployment. The Brighton group considered that innovations appeared to develop during the second half of the prosperous phase, thus explaining the increase in unemployment. The most advanced form of this approach is expressed by the concept of "New Technological Style" proposed by Carlota Perez (1983) who intelligently combined Schumpeterian concepts with the institutionalist theories of D. Gordon and the regulation theories of R. Boyer. These approaches form the greater part of current research on the long-fluctuation phenomena and have the common feature of making innovation play a major role.

In general, innovation is considered a qualitative change in terms either of product (creation of a new product) or of process (new manufacturing technique). When an inventory of innovation is drawn up *ex post facto*, one generally tends to favor visible innovations that are concretized by clearly distinct new products or by characteristic material installations (the Bessemer furnace, for example). In addition, in most cases only the first application is recorded, and the following ones are ignored.

This manner of operation leads to isolating the innovation from the economic and social context in which it was produced. It appears as a purely technical, essentially material, fact that is produced directly by capital or its engineers. The changes contributed by the workers themselves, who often planned them, are generally forgotten and never figure in the inventory. In addition, studying technical change – in terms of innovation – tends to take it out of its historical context. The innovation – cut off from its past – appears to be the result of chance and not the result of a process that is part of a technical, social, and economic rationality.

Long fluctuations thus appear to be caused essentially by technical progress that is broadly exogenous with regard to the field of economy. The crises in which they appear are forced on men like fate, forming the ransom that must be paid to progress.

We obviously do not intend to negate the role played by innovations in long movements but intend to give it a different position in the set of structural processes that engender fluctuations.

Our theory of long fluctuations has been developed elsewhere. Very briefly, fluctuations express the regulation of the growth of productive forces in their link with the social relations of production in whose framework they are developed. These production relations periodically block this development. Rate of profit is the instrument through which regulation operates. A high rate of profit results from developing production methods that favor the material means of production at the expense of the workers who put them into operation. In time, the resulting imbalance leads to a fall in the efficiency of the production system and production relations. This fall in profit weighs on behavior, thus leading to development of new productive forces and encourages the adaptation of production relations.

In terms of productivity, the first type of development leads to economizing live labor and excessive expenditure on past labor. In contrast, the second type encourages the economizing of past labor and the development of live labor.

The coal survey started in 1981 verifies these hypotheses and the data at the same time. In the first part we drew up a set of accounts describing the evolution of the productive forces and the economic relations within whose framework they were developed. We then attempted to analyze the qualitative changes brought about by this development in terms of innovation and technical change, on the one hand, and in terms of skill, qualification, way of life, and consumption on the part of the labor force, on the other.[1]

It is doubtless excessive to try to establish a continuum between the measurable economic transformations of production relations and the qualitative evolution of the productive forces. However, if one wants to understand what is happening one cannot just draw up accounts of economic activity. It is necessary to enter the question, observe the transformations that take place, and actually see if they have measurable economic results. Thus, for example, the hypothesis regarding the development of productivity that takes place during the long fluctuation, sometimes by privileging the live labor economy and sometimes by privileging past labor, can be measured in terms of result through the evolution of labor and capital productivity. However, this measurement is soulless and lacks history. Its only perspective is mechanistic if it is isolated from the transformation of man and the tools that he fashions and from the social relations within whose framework these transformations take place.

We aim to set down a few markers along this pathway by means of studying underground haulage in mines. It is of course difficult to limit the field of investigation to a single aspect of a complex technological process. In any evolving productive activity, all the technical processes used interact; any change that occurs in a particular domain has frequently originated in another.

The coal industry is no exception to this principle. It serves as an excellent example of this principle owing to the increasing complexity of the operation, the variety of the situations faced, and the techniques used to adapt to individual situations. Indeed, the characteristic of the mining industry is the need to integrate nature as a direct, immediate, productive force. This can only be achieved by adapting to nature's own requirements.

This permanent, direct affront to nature gives the coal-mining industry a special stamp. It consists of exploiting a natural, nonrenewable, fossil resource. Apart from the discovery of new deposits, and if it is assumed that exploitation begins with the most accessible part, the complexity of the problems increases with time. Nature's productivity thus decreases. Increasingly remote deposits must be sought at greater and greater depths, and this in itself is a factor of technological evolution.

Underground coal mining consists of removing coal from the seam and conveying it to the surface. From the technical point of view, this operation involves the solving of successive problems:

- Searching for, identifying, and characterizing the seams to be mined.
- Access to the seams mined.
- Management of cutting and the manner of exploiting the seams.
- Support of the workings.
- Lighting.
- Ventilation.
- Breaking down coal.
- Haulage of coal to the surface.

This last point can be divided into horizontal haulage or haulage along various slopes at the bottom of the pit or in galleries leading directly to the surface and vertical lifting in the shafts themselves.

These different aspects of mining are obviously interdependent and any progress made in one field will modify the equilibrium and bring about changes in the others. The study of a particular sector thus assumes at least overall knowledge of the situation in the others. From this point of view, the evolution of underground transport is conditioned by two essential factors: the increase in the quantities mined and the increase in haulage distances. These two factors exert continuous pressure with regard to modification of techniques to increase the speed and capacity of transport.

However, it would be a mistake to consider this evolution purely from its technical aspect and with regard to the physical constraints that weigh on it. Techniques are put into practice by labor. This assumes *savoir-faire*, skill, and creativity on the part of the workers concerned, and the labor itself involves cost. Combining these factors produces a technical change. This is what we would like to demonstrate through study of underground haulage.

10.2. Mining at the End of the Eighteenth Century

Coal production in France remained at less than one million metric tons per year until the end of the eighteenth century. Small operations were a general rule, consisting of a few miners and an operator who managed the work. In many cases it was an alternating activity, and the workers were sometimes in the field and sometimes down the mine. Techniques were rudimentary, and were generally carried out using narrow galleries, in which a child could hardly stand, leading directly to the open air. There was sometimes no propping.

The main technique for hauling was carrying the coal on miner's backs. This is the most primitive type of haulage, characterized by the fact that everything is supplied by human energy – both lift and displacement depend on the carrier's muscular strength alone. It is the most simple method but also the most flexible and the best adapted to thin seams, poor tunnels, and steep slopes. Rouff described the workings in the Saint-Etienne region toward the middle of the eighteenth century:

> Men and women and children of both sexes work in these diggings, the men use pick-axes to break off corners and blocks of coal from the seams, boys bring up the large coal and women and girls put the slack in long, narrow bags holding a hundred to a hundred and fifty pounds. . . . Thus loaded, these wretches climb the rampart of the dim galleries without lights. The effort that they have to make to overcome the weight of the burden, the difficulty of the route and the discomfort of their position forces them to make violent inhaling and exhaling movements of the chests, whence emerge plaintive, broken sounds.[2]

More complex forms of organization were required in certain workings by the depth of the seams and the difficulties to be overcome to reach them. In 1716, Desandrouin in the Nord département was obliged to have recourse associates to assemble the capital necessary to dig the first shafts. An early steam engine was used in 1730 to pump out water; there were four in 1751. However, the shallow seams meant that the galleries were only 1.20 meters high, and the coal was generally hauled.

> Since it was impossible to use workers of normal height to go into these narrow caverns, young people went in, sorted the stones from the coal and loaded each separately on small sledges which they took to the shaft, crouching and sometimes on all fours the whole length of the seam.[3]

Hauling was an improvement over carrying. The energy required was divided, and workers provided only the displacement. However, in galleries where the slope was greater than 12° the hauler had to be helped by a child who pushed. This method could not be used when the slope was more than 20°. It can thus be seen that the change to the new method of transport assumed progress in management of operations and in the driving of galleries whose gradient had to be fixed according to the method of haulage used.

To improve the yield of the method and to increase the load of the sledges, gallery floors were boarded or even lined with logs on slopes and the sledges were fitted with iron skids.

Thus designed, underground haulage used the greater proportion of the labor force. The observations available report that three workers were occupied in conveyance for every two involved with breaking down coal. Measurement of the useful work carried out each day by a worker showed that a carrier could take loads of 50 to 75 kilograms as far as 50 to 100 meters up slopes of up to 50°. He could thus move 200 to 300 kilograms in one day over one kilometer. The hauler could carry loads of 60 to 160 kilograms for distances of 60 to 200 meters. In all, he moved 350 to 800 kilograms over one kilometer.

In 1790, Dieudonné estimated that the average wages of the Compagnie d'Anzin was 270 francs per year, i.e., approximately 90 centimes per day. At that time the company mined 75 metric tons per worker per year. This was sold for 10 francs per metric ton. Under these conditions, any increase in production implying more than a proportional increase of the distances to be traveled soon reached a limit.

10.3. The Changes of 1820–1850

In spite of a significant increase in production, this haulage system remained largely dominant in French mines until the early 1820s. The situation could not remain as it was with the development of the depression phase. Prices were falling and labor costs, which were still 300 francs on average in 1820, rose rapidly and reached 700 francs in 1847. In addition, production continued to increase at a very high rate, which doubtless implied a more than proportional increase in haulage volume.

Underground haulage underwent four decisive changes during this period:

- Carrying and hauling on skids were replaced by wheel-mounted dogs or wagons.
- Rails were laid in the galleries.
- Man was progressively replaced by the horse as a driving force.
- Finally, self-acting inclines were installed on steep downward slopes.

Wheeled wagons were introduced in 1820 in the Aveyron department, in 1821 in Loire, in 1822 in Nord, in 1836 in Tarn, and in 1843 in Hérault. The Huelgoat mines in Brittany were influenced by the British, and were equipped with dogs in 1806.

Wheels were first made of wood with iron rims, then cast iron, and finally iron. Car bodies were made of white wood, which was light but fragile, or oak. Wood was later replaced by lighter and stronger sheet steel.

The oldest mine dogs held 150 kilograms of coal, which was the most that would be hauled under the best conditions with a car on skids. However, loads increased steadily. At Bessèges in 1858 the cars carried 1,100 kilograms, but this was an exception. In general, loads averaged between 250 and 500 kilograms.

In France, adoption of the wheel often coincided with the laying of metal rails. There was thus a replacement of skid-mounted cars, sometimes guided by longitudinals, with wheel-mounted cars running on rails. In contrast, wheel-mounted dogs that ran on wooden rails were used in Germany before 1800.

The first rails used by Compagnie d'Anzin in 1822 were grooved and made of cast iron. Aniche became equipped with iron-flat rails in 1825. Headed rails were used later.

Horses were first used only in mines where the seams were thick enough to support large galleries without having to mine the rock. This change was first introduced in the Loire coalfields in 1824. The first horses were taken down the mines in the Nord département in 1847 and galleries had to be specially fitted.

Naturally, different types of haulage were present in mines simultaneously for many years. The main roads were equipped first. In 1836, the engineer Manès, described mine haulage in the Carmaux mines as follows:

> Haulage inside is by sliding wagons fitted with iron skids and with a capacity of 1.25 hectoliters over beams laid transversely along the galleries. The haulers take these wagons along horizontal galleries and lower them in inclined galleries using a winch. . . . This method will soon be replaced everywhere by haulage on rails as at the Ravin mine.
>
> These railways consist of two lines of rolled steel bars 0.80 m apart laid on wooden sleepers. In the horizontal parts there is a single track with stages of 40 to 50 fathoms (about 80 to 100 meters). There are two metal tracks with 3 or 4 rails in the inclined parts.
>
> The cars or wagons with four wheels and wooden bodies and a door for unloading contain 5 to 6 hectoliters and run through the main galleries. They are loaded by wagons from the face. They are hauled manually to an incline, lowered using a braking hoist and then taken by trammers to the shaft where they are emptied. The coal is shovelled into hoisting-buckets of the same capacity and taken up to the surface either by means of a pit-head pulley or steam engines.[4]

When this was written in 1836, coal was still discharged and loaded twice: first in the main galleries receiving coal from the face and then at the bottom of the pit.

The improvements made in management of the workings made it possible for trammers to take the coal from the face. Unification of the haulage system was complete when haulage cars were put into hoisting-cages in 1848. Coal could be taken from the face to the sorting area in a single operation.

Manès described haulage in steeply sloping galleries starting from the use of a braking-hoist. This is in fact a development that uses gravity as a source of energy. A hoist fitted with a brake is used to lower loaded cars in 30° to 45°

slopes. A pulley with a brake is also used in which the descending loaded cars use the returning empty cars as counterweights. Of course, it appears paradoxical to lower the load when it subsequently has to be raised to take it out of the mine. In fact, the system enabled the best use of extraction mechanization in the mines created in the early nineteenth century. In the present case, the coal had to be taken as rapidly and as cheaply as possible to the foot of the shaft.

This new technique was important because it opened the way to a second reversal of the direction of movement of the loads with rising inclines. In this case the force of gravity was formed by the horse that hauled empty cars down, thus pulling full cars upward.

These improvements successively enabled considerable increase in the amounts hauled. Whereas haulage on skids enabled useful work of about 400 to 800 kilograms per day over a kilometer, haulage using wheeled cars enabled an increase to 2,000 or 3,000 kilograms under identical conditions. When the roads had sufficient slope, useful work could attain 9,000 kilograms for a single worker. Useful work over distances of 200 to 4,000 meters could reach 10,000 to 64,000 kilograms with the use of horses.

These new methods of haulage were most effective when the whole of the workings was organized appropriately, with sufficiently spacious main galleries to enable horses to be used and with slopes such as to enable them to move large tonnages effortlessly. This could only be achieved progressively and above all in new workings. The inertia factor here is certainly not related to the duration of fluctuations.

Another aspect of evolution was that nature as a production force was tamed, ordered, and regularized, paving the way for progress. Things had come a long way since the narrow tunnels with slopes of often up to 50° through which miners crawled at the end of the eighteenth century.

The first result of these changes was that it was now possible to cut coal at increasing distances from the shaft. The length of galleries increased but the number of pits decreased. Thus in 1810, Compagnie d'Anzin produced 280,000 metric tons from 25 pits. In 1876 only 13 pits producing 2,137,000 metric tons remained, but there was a total of 39,220 kilometers of galleries. The capital–product ratio fell, and the same capital could be used for more workers.

A second consequence was that the structure of intermediate consumption was deeply modified, with a strong increase in the consumption of cast iron and steel accompanied by a fall in the prices of these materials, whereas use of wood decreased in relative terms and the price rose.

Finally, the third result was that the change in haulage made it possible to reduce, at least comparatively, the number of workers involved in this unskilled work. At the end of the 1840s, no more than one hauler was needed for two hewers at the face whereas three had been necessary a century earlier. In addition, skilled workers were now needed to lay and maintain the roads and to repair the cars. The change of the production tool thus required labor with more skills. This trend certainly accounts in part for the rise in wages observed during this period.

The main changes in the haulage system were complete at the beginning of the prosperous phase from 1850 onward. They led to considerable increase in productivity and also made it possible to go beyond the limits that the development of productive forces had reached with the previous haulage method. With the exception of the installation of inclines and the fitting of axle-boxes to car wheels, we noted no important technical changes during the phase of prosperity from 1850 to 1873. Thus, all efforts were concentrated on generalizing the new methods to reach the most remote parts of the mines. Equipping main galleries at the beginning of the changes obviously had important repercussions on productivity by removing bottlenecks, and reducing costs for the very large quantities hauled. As the system spread, the marginal productivity of the new extensions tended to diminish. There was certainly an economy in live labor, but an increasing amount of past labor had to be expended to achieve this result.

We have not found any information indicating conclusively that technical progress might have lessened the quality of the coal during the prosperous phase. However, this type of information is rarely given in technical publications that ignore the improvements or simplifications that lead to employing unskilled labor instead of better-paid workers.

In fact, during the prosperous phase, the whole technical tool changed progressively to make better use of the new potential. Galleries were given a very slight slope in the direction of haulage, inclines were installed, tracks were widened, horses were used more and more as motive force instead of men, etc. While this occurred, the system was forced to its limits while further changes were getting ready: innovation requires a favorable environment.

10.4. Mechanization of Haulage after 1870

Wheeled haulage on rails enabled a considerable increase in the tonnages transported. However, evolution of the production tool, that of haulage itself, tended to make increasing amounts of coal converge toward the same point while the cutting faces tended to become further from this center. In mining, digging a shaft is a fixed cost while the driving of galleries to cut coal further away is a proportional cost. It is thus advantageous to extend galleries to the point at which the operation cost exceeds the depreciation cost that would be caused by the installation of a new pit. This shows the strategic position of haulage. It is essentially haulage and its cost that govern the point in which considerable sums must be invested in a new installation and a new risk is taken.

Mine head pits, equipped with powerful apparatus, were thus the destination of large quantities of coal that tended to saturate the main access galleries. Horse-drawn haulage became too slow to handle the quantities produced. Animal traction had reached its limits, and the question of mechanization was raised.

In 1870, the mechanization of draught could not go directly from horses to motor engines serving the same function. This would have implied a degree of miniaturization inside the mines that could not be attained by steam engines, not to mention the risks of fire or fire damp.

However, the principle of inclines, which was widely used in mines, made it possible to combine a fixed engine and transport mobility. It was enough to equip the main pulley with a mechanical engine or a transmission system connected to an engine

English mines were the first to use the new method in 1851, and were followed by Germany in 1862 and Belgium in 1865. The first mechanical traction systems were installed at Blanzy in the Loire département in 1872 on an incline. Anzin used the method in 1873 in a horizontal gallery, and the method was soon adapted by Aniche, Ferfay, and others.

Traction was by an endless chain or rope below the wagons or above the wagons. In the first category, the wagons were hitched to the low part and ran as complete trains drawn by a chain or a rope. The fore rope is called the main rope, and the rear one is the tail rope. Each cable is alternately hauled or hauler depending on whether the wagons are full and going to the shaft or empty and returning to the face.

In the second category, suitably spaced wagons ran alone at a speed of between 1.1 and 7.2 kilometers per hour. They were hitched above to a single endless chain or rope running round horizontal pulleys at the end of the run. One pulley was the driving pulley and operated continuously.

The decisive problem for this new type of haulage was that of putting the driving power down the pit. In mines with no fire damp, it was possible to install a steam engine at the bottom of the pit if there was adequate ventilation. Numerous examples of this type were to be found in England, but it was rarer in France. In 1888, 33 machines had been installed down mines, 31 of which were in the Bouches-du-Rhône and Gard départements.

Steam engines were sometimes installed at the mine head and a driving rope ran to the bottom of the shaft. The steam engine could also be installed at the bottom and the boiler at the surface, but the steam rapidly lost its energy as the temperature decreased. Thus, when the engine was separate from the boiler it was preferred to use compressed air to run the engines down the mine.

In Nord-Pas-de-Calais, practically all the pits were equipped with one or two 500 hp air compressors at the end of the nineteenth century.

The first mechanical installations in 1875 had a capacity of approximately 600 metric tons per day, but at the end of the century daily figures commonly reached 1,000 metric tons and were sometimes as high as 2,000 metric tons.

In terms of employment, mechanical draught eliminated unskilled labor once more. At Aniche in 1875, the mechanical system installed in the Saint-Marie pit replaced 39 horses. Thus, unskilled workers – generally children – were eliminated and operation of the new installations required highly skilled labor, particularly for maintenance. The new technique also stimulated use of skilled labor. In addition, although it resulted in a considerable increase in labor productivity, it tended to increase that of capital. The installation at Aniche cost 55,000 francs, which is distinctly less than the value of the horses.

Installation of mechanical traction involved the solving of complex problems, and the cost was high. It could therefore only be used in galleries with sufficient discharge. Nevertheless, here again, the system gradually spread to less profitable roads during the prosperous period. Thus in Westphalia, in 1905, there were 149 installations handling 372 metric tons per shift. In 1907 there were 195 installations but the average amount handled had fallen to 312 metric tons.

The disadvantage of rope traction is that it requires continuous supervision, made difficult by the fact that the parts are spread along the whole of the route. In addition, a simple breakdown brings the whole system to a halt and soon stops work upstream. A more flexible system was therefore needed to enable engines to be smaller, and this became possible with the use of electricity and internal-combustion engines. Compressed-air engines with independent tanks were also used.

The first electric locomotives appeared in Germany in 1889, and the first locomotives driven by petrol engines were used there in 1890. The first electric locomotives in France were used at Gardanne in 1898. These were then installed in Loire from 1900 onward. Mines in Nord-Pas-de-Calais began to use compressed-air locomotives in 1910.

This new type of haulage became widespread after World War I. In 1930, 450 locomotives were in use in mines and 178 of them were driven by compressed air.

The locomotive – electric or not – was not a decisive advance in relation to mechanical traction by rope, whose cost per metric ton was frequently lower. The main advantage lay in its flexibility, but in secondary galleries it competed with difficulty with horses. In 1930, it was estimated that threshold of profitability of a locomotive was 240 metric tons per shift per day, which was rarely the case in a secondary road. This explains why horses were still clearly dominant at that date.

In fact, at the end of the Great Depression of 1880–1890, haulage was no longer an obstacle to developing operations and the decisive advances that deeply modified mining. Whereas in the nineteenth century the most important technical changes were in haulage, the techniques used to break down coal were to be revolutionized in the twentieth century.

Although it is still incomplete, the study of technical changes in mining confirms the results obtained with quantitative measurements and also gives a different image of the innovation process by placing it in its technical, social, and economic context.

Notes

[1] Within the framework of this survey, the study of underground haulage was carried out by Anita Prigent and is the subject of a report entitled *Evolution technologique des transports à l'intérieur des mines en France de 1750 à 1930* [*Technical Development of transport inside the mines in France from 1750 to 1930*]. The information used in this article was taken from the report. In addition, five others have examined the problems of ventilation, explosives, control of fire damp, compressed air, and coal cutting.

[2] Rouff, M., 1922, Les mines de charbon en France au XVIIIe siècle, 1744–1791, Rieder, Paris, France.

[3] Grar, E., 1847, Histoire de la recherche, de la découverte et de l'exploitation de la houille dans le Hainaut francais, dans la Frandre francais et dans l'Artoid (1716–1791).

[4] Manès, 1836, *Annales des mines.*

References

Freeman, C., Clark, J., and Soete, L., 1982, *Unemployment and Technical Innovation*, Frances Pinter, London, UK.

Mensch, G., 1979, *Stalemate in Technology: Innovations Overcome the Depression*, Ballinger, New York, NY, USA.

Perez, C., 1983, Structural Change and Assimilation of the New Technologies in the Economic and Social Systems, *Futures*, October:357–375.

CHAPTER 11

The Diffusion of Process Innovations as a Factor Shaping Industrial Structures: The Case of Shuttle-less Looms

Cristiano Antonelli

11.1. Introduction

Diffusion of innovation has long been analyzed within a static context where the technology itself and the features of potential adopters were not allowed to change over time.

The epidemic approach, elaborated by Griliches (1957) and Mansfield (1961) modeled the diffusion of innovation as a logistic or S-shaped pattern governed by the spread of information concerning the advantages of the innovation among potential adopters, brought on by the cumulated epidemic effect of adoptions. The epidemic approach considers the diffusion process as the outcome of a contagion process in which, within a given population, at each point in time the chances of new adoption are governed by the ratio of old adoptions to total potential adopters. The relationship between the new adoptions and the old ones is such that at the beginning of the diffusion new adoptions increase very fast with respect to the small levels of shares of old adopters to total ones. Later on however, after the share of old adopters to total ones has passed a threshold laying between 37% and 50%, the rate of increase of new adoptions tends to decline.[1] It will eventually stop in the proximity of a share of old adopters to total close to unity.

In the epidemic model early adopters are the ones able to outguess the higher profitability of the new product, while late adopters rely on the knowledge of the higher total productivity of the innovation spread by the experience of early adopters. In the epidemic models of diffusion, the profitability of the innovation as well as the ability of adopters to appreciate it remain (or can remain) stable over time. The diffusion, in fact, is governed by the spreading of the information concerning the innovation within a static population of differentiated agents. Internal forces engendered by the demonstration effect of cumulated adoptions are the one dynamic motor in the process.

It is thus clear that the basic assumption of the epidemic model is in the uneven and static distribution of innovative skills or outguessing capacity among potential adopters.[2]

This rigid and static assumption seems more and more difficult to apply when all the changes associated with the diffusion of an innovation – such as the decline of costs, the learning effects, the role played by supply forces, the growth or entry of early adopters, and the exit or decline of late ones – are considered.

A first attempt to introduce a proper consideration of some of these dynamic factors into the analysis of diffusion process has been made by a recent wave of so-called equilibrium models. Equilibrium models stress the change over time in the ability of agents to appreciate the advantages of the innovation as well as the changes in the innovation itself brought on by post-invention improvements.

More precisely equilibrium models see the diffusion process as the result of the rational decision process of differentiated categories of adopters that *face continuate changes* of the features of the innovation to be adopted (Antonelli, 1987) and that *are supposed to be able to take advantage of processes of learning by using* (Davies, 1979).

At each point in time potential users decide to adopt the innovation when its marginal productivity equals the opportunity cost of its delay. The marginal productivity of the innovation, however, increases over time because of:

- Introduction of incremental technical change and consequent increase of its total factor productivity.
- Learning by doing of innovation producers and consequent reductions in costs to users.
- Learning by using among adopters and consequent increase of its total factor productivity.
- Erosion of monopolistic extra profits of innovation producers and consequent reduction in prices thus in costs to users.
- Secular decline of the relative cost of capital with respect to wages (for labor-saving process innovations) and consequent reduction in relative cost to users.

When the size distribution of a firm is considered, the aggregation of these results of the microeconomic analysis of the adoption process can yield an aggregate process of diffusion characterized by a strong sigmoid pattern. Current size distribution of a firm's potential adopters (i.e., normal and lognormal ones) proves in fact to yield sigmoid patterns of diffusion at the aggregate level, compatible with equilibrium models at the microeconomic one, given that large firms have higher chances of adoption than smaller ones.

Following an established literature with significant empirical evidence, large firms are in fact expected to have, at the beginning of the diffusion process, a higher level of anticipated payoff of the innovation and thus adopt it earlier than smaller ones.[3] Eventually though, smaller firms will also appreciate, thanks to the dynamic force already considered, the accrued marginal productivity of the innovation and will adopt it with some delay. Reasons for this difference between larger firms and smaller ones include the following:

- The longer period necessary for small firms to acquire relevant information about the real payoff of the innovation (David, 1969).
- The longer period in which large firms calculate the payoff with respect to small ones (Davies, 1979).

As a final result we have a diffusion process characterized by the early adoption of large firms and the eventual adoption of smaller ones fueled by dynamic forces increasing the advantage of adoption. When such a diffusion process takes place in a common industrial population, with a given and static normal or lognormal size distribution of firms, it will follow an S-shaped aggregate pattern.

The strength of these equilibrium models seems to be grounded in the explanation of a microeconomic behavior of adoption consistent with the empirical observation of aggregate S-shaped diffusion patterns. Such a result, however, is obtained from the static features of the population of potential adopters – such as the number and the size distribution of firms and the role of ultimate determinants of the diffusion process.

This preminent role of the size distribution of firms in determining the pattern of aggregate diffusion seems to be a limit. It precludes in fact the integration into the analysis of diffusion of possible changes in the population of potential adopters engendered by the features of the diffusion process itself.

This limit seems especially relevant when one considers that the diffusion of process innovation has major effects on the production costs of early adopters and their competitive advantage over late adopters.

The adoption of a process innovation is expected to make possible a downward shift of the average cost curve. Early adopters in such a context can be considered innovators with a transient competitive advantage in terms of an average cost curve lower than that of non-adopters. A variety of possible outcomes of such a change in the competitive advantage of firms brought on by the diffusion and the time lag in the adoption of the innovation is possible.

Early adopters can take advantage of lower production costs by expanding their long-term production levels and, for a given demand, their market shares.[4] In such a case, early adopters would not wait for the diffusion process to reach late adopters but would rather push it by means of accelerated rates of incremental adoption, which are a function of their rates of growth. The aggregate diffusion would then be fostered by a form of expanded intrafirm diffusion process. Moreover, the population of potential adopters would be affected by the diffusion process. Late adopters would abandon the industry because of their high production costs. In fact they would never adopt because of forced exit.[5]

An alternative outcome is possible when early adopters do not expand their long-term supply: entry of new competitors based on the early adoption of the new process innovation can then take place. In this case new entrants would take the place of late adopters and would force them to abandon the market: the new entry in fact would increase supply and reduce costs. Only modern firms equipped with the new capital equipment would be able to remain in the market.

In both cases the introduction of a new process innovation and the time lag in its adoption induce major changes in the competitive advantage of firms: the survival of late adopters is put at risk either by the growth of early adopters or by the entry of new adopters that are more able to take advantage of the new capital equipment available.

As a result of such a chain of events the number of firms and the size distribution can change along the diffusion process. The diffusion process in fact would be the ultimate cause of the size distribution of firms rather than a consequence of it. It seems now clear that a major limit of equilibrium models is the lack of appreciation of the effects of the differential ability of firms to adopt innovations on their growth and survival. The timing of adoption of process innovations becomes, in fact, a factor of competitive advantage and industrial selection.

11.2. Adoption Survival and Entry

11.2.1. The hypothesis

The features of the diffusion of shuttle-less looms in the Italian cotton-weaving industry in the period 1973–1982 provide excellent empirical evidence to test the role of diffusion of process innovation as a factor shaping industrial structure. Main features of the sector can be summarized as follows:

- The market is open to international competition.
- Labor costs account for 50% to 60% of total added value.
- Products are technically homogeneous with low levels of technical differentiation.
- Demand is highly price elastic.
- Markets are easily contestable, and the entry of new competitors takes place by means of reductions in prices.
- Equilibrium prices show secular declining trends because of the competitive pressure of new international producers characterized by low labor costs.
- New process innovations have been introduced in the entire production cycle from spinning (open end rotors) to finishing (electronic regulation of dying) through weaving (shuttle-less looms).
- The use of new synthetic materials and new product mix spread among final consumers.
- Major changes in the costs of inputs, namely, raw materials, energy, and labor, took place in the 1970s.

The industry adjusted to the turbulence by means of fast rates of diffusion of process and product innovation. The adjustment, however, was also the result of a major process of industrial reorganization. Heavy mortality rates and an inflow of new competitors changed the features of the structure of the Italian cotton-weaving industry (see *Table 11.1*).

Total number of firms declined from 405 in 1971 to 305 in 1985. Employment shrank from 42,335 to 22,010 with the loss of 20,025 jobs, i.e., 0.47% of the 1971 total. Overall average size of firms declined from 82 employees to 63. Production levels of the weaving sector, however, increased from 163,296 million tons of all products in 1973 to 226,401 in 1985. International competitiveness, as measured by the ratio of export–import to export+import, remained stable in relative terms at – 0.51 from 1971 to 1985.

In the same period the diffusion of shuttle-less looms spread quickly: the cumulated stock jumped from 734 installed in 1974 to 14,078 in 1985 [6] and the penetration index [7] passed from 1.05% to 20.18%.

Shuttle-less looms can be considered a major process innovation with major effects on the production cost of weaving firms. The shuttle loom had reached in the 1960s the maximum production capacity of 180–200 knots per minute, with a work load of one employee per four machines. Shuttle-less looms in the late-1970s had an average production capacity of 300 knots per minute with a work load of one employee per eight machines. In short, the introduction of shuttle-less looms made possible substantial increases in total factor productivity and lowered production costs by 30%

These elements seem to provide an excellent opportunity to test the hypothesis that the time of adoption of process innovation has major competitive effects on the ultimate result of changing the features of the population of potential adopters as it was before the diffusion process took place.

In such a context, characterized by falling prices, high levels of international competition, introduction of process innovations, price competition, and domestic market contestability, we expect to find that the time of adoption of a radical process innovation, such as shuttle-less looms, with major effects on production costs is associated with:

- The survival of early adopters.
- The entry of new firms as early adopters.
- The exit of late adopters.

We expect to find that existing firms that adopt the new production technique will be able to reduce production costs and to match the competition brought on by international producers localized in countries with low labor costs and by new domestic competitors. Moreover, we expect that in a context characterized by the decline of prices, late adopters using less-efficient techniques will be forced to abandon the market because of their relative inefficiency and consequent growing losses. Finally, we expect that the new production process offered to new competitors will provide the opportunity to overcome the (low) barriers to enter into the industry and make possible a flow of entries.

Table 11.1. The evolution of the Italian cotton-weaving sector.

Year	*Weaving firms*	*Employment*	*Export*	*Import*	*Total no. of firms*	*Average size of employment*	*Average size of production*	*Total production tons*	*Shuttle-less looms*	*Total looms*
1971	405	42,335	13,026	40,750	514	82	317	163,296	734	69,765
1972	382	37,319	17,504	40,971	476	78	340	162,019	968	63,935
1973	364	40,443	14,640	58,585	458	88	371	170,930	1,337	62,652
1974	353	38,425	15,465	56,851	450	85	402	181,920	2,658	60,960
1975	346	36,270	18,781	55,481	435	83	367	160,330	3,196	56,990
1976	333	35,630	27,401	81,049	420	84	430	181,190	3,680	54,100
1977	327	34,680	32,229	66,234	410	89	417	170,990	4,098	52,120
1978	333	32,953	39,932	69,335	403	81	434	175,019	4,700	49,171
1979	337	32,129	43,371	99,554	404	79	487	197,682	5,732	47,200
1980	344	30,310	33,167	94,260	407	74	501	204,240	7,220	45,304
1981	334	29,121	33,177	77,669	390	74	551	215,665	8,122	43,307
1982	330	27,559	33,579	92,116	386	71	572	221,297	8,602	40,765
1983	322	25,475	38,614	92,685	372	68	551	205,958	12,226	34,080
1984	318	23,283	43,005	96,168	363	64	636	231,391	14,241	32,110
1985	305	22,010	37,382	116,621	348	63	650	226,401	14,078	29,505

11.2.2. The test

The Archives of Associazione Cotoniera Italiana contains a large data set with information on the time of first adoption of shuttle-less looms, total employment, total production, and total looms installed in the period 1973–1982 for all the members of the Associazione, i.e., 116 firms of which 99 were active in 1973 in the weaving sector and 17 were established later. The data make it possible to analyze the features of the adoption process of shuttle-less looms, the relationship between adoption and survival, and the change in the population of potential adopters along the diffusion process.

In 1982 at the end of the period, a total of 55 firms out of the 99 active in 1973 were still in operation and had adopted shuttle-less looms. In 1973, 20 had already installed shuttle-less looms; the sequence of the adoptions in the following years (see *Table 11.2*) shows a peak of adoptions in the period 1974–1976 with 24 new adoptions. The aggregate diffusion pattern of shuttle-less looms within the 99 firms considered (see *Table 11.2*) closely follows the path of diffusion (see *Table 11.1*) of all the industry.

Table 11.2. Evolution of the stock of shuttle-less looms installed in the period 1973–1982 and flow of new adopters.

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
Stock of shuttle-less looms	661	974	1,409	1,805	1,999	1,628	1,686	2,002	2,852	2,937
Flow of new adopters	20[a]	7	8	7	3	1	0	4	3	2

[a]Firms with shuttle-less looms already in place during the first year observation.

Econometric estimates of the logistic and log logistic are given in equation below:

$$\log \frac{y}{(N-y)} = a + b_1 t$$

$$\log \frac{y}{(N-y)} = a + b_1 \log t,$$

where y is the cumulated number of shuttle-less looms installed, N is the number of total automatic looms in place in 1973, and t varies from 6 to 17 on the aggregated data of the 99 firms. The data on all the industry provides an excellent estimate of the diffusion pattern.

The following results, respectively, for industry one in equations (11.1) and (11.1a) and industry two in equations (11.2) and (11.2a) provide a strong and significant fit of the diffusion process:

$$\log \frac{y}{N-y} = \underset{(22.097)}{-5.042} + \underset{(11.032)}{0.196} \text{ TIME} \tag{11.1}$$

$R^2 = 0.936 \quad F = 121.705$

$$\log \frac{y}{N-y} = \underset{(15.368)}{-4.709} + \underset{(5.947)}{0.142} \text{ TIME} \tag{11.1a}$$

$R^2 = 0.815 \quad F = 35.371$

$$\log \frac{y}{N-y} = \underset{(20.290)}{-8.559} + \underset{(14.218)}{2.390} \log \text{TIME} \tag{11.2}$$

$R^2 = 0.961 \quad F = 202.157$

$$\log \frac{y}{N-y} = \underset{(11.199)}{-7.291} + \underset{(6.724)}{1.744} \log \text{TIME} \tag{11.2a}$$

$R^2 = 0.849 \quad F = 45.215.$

The results confirm the following:

- The diffusion process at the aggregate level has a strong sigmoid character that seems better approximated by the log logistic interpolation. After an early start the pace of the adoption process, in fact, slowed down.
- The diffusion process within the sample of firms has a pattern close to that of the universe but for the marked slowdown of the growth of installed shuttle-less looms in the years 1978 and 1979.

Following the equilibrium approach to the analysis of diffusion we expect such an aggregate result to be the consequence of the ability of the large firms to adopt the process innovation earlier than the small firms: Large firms should, in fact, have easier access to technical information, larger flows of investments, and longer periods of payoff.

The evolution of the size of new adopters (see *Table 11.3*) confirms the hypothesis. In the period 1974–1982, 35 firms already active in 1973 installed the new process innovation. The evolution of the average size of new adopters shows, in fact, a declining trend in terms of total employment. Moreover, the evolution of the average size of the first installment of shuttle-less looms has a stronger negative trend.

A further test of the positive role that the size of a firm has in determining the chance of eventual adoption of shuttle-less looms is provided by the econometric estimate of the probit equation (11.3) tested on the universe of 99 firms active in 1974:

$$\text{ADOPTION} = a + b_1 \text{ SIZE}, \tag{11.3}$$

Table 11.3. Evolution of size of new adoptors.

	Number of adopting firms	*Average size of employment*	*Average number of shuttle-less looms installed*
1974	7	828	39
1975	8	450	33
1976	7	244	15
1977	3	594	18
1978	1	22	12
1979	0	0	0
1980	4	281	18
1981	3	250	29
1982	2	376	18
Total	35		

where ADOPTION is a dichotomous variable with value 1 for firms that adopted shuttle-less looms in the period 1973–1982 and 0 otherwise; SIZE is a continuous variable measuring the employment of firms.

The results of the test confirm that the size of a firm had a strong and significant role in explaining the chances of adoption:

$$\text{ADOPTION} = \underset{(2.53)}{-3.158} + \underset{(4.74)}{0.283}\,\text{SIZE} \tag{11.3a}$$

$$R = 0.324 \qquad P = 0.006 \qquad X^2 = 19.35$$

(partial chi squared statistics between parentheses).

According to traditional equilibrium models, after large firms opened the adoption process, fueling the first rapid phase of diffusion, smaller firms should start their adoption, completing at a lower rate the second phase of the aggregate diffusion process.

In the weaving sector a much more complex process seems, however, to take place. Late adopters, in fact, are kicked out of the industry because of their higher production costs in a context of declining market prices and aggressive entry of new competitors.

In the same years 1973–1982 (see *Table 11.4*) 48 firms, out of 99 existing in 1973, died with a total mortality rate of 0.48. Exits peak in 1980 with 10 cases; the period 1977–1978 with 10 is also noticeable. The size of firms forced to exit is not significantly different from the average size of firms.[8]

The chances of a death of a firm seem to be associated with the timing of the adoption of shuttle-less looms. The large majority of exits was non-adopters. Only 19 firms, which eventually left the industry, had already adopted shuttle-less looms, while the remaining 29 were non-adoptors. In 1982 only 51 firms out of the original 99 were still active; of these 31 had already adopted shuttle-less looms and 20 had not.

In absolute terms at the end of the period, the probability of the death of adopters, 0.38 (19/50), was significantly lower than the probability of non-adopters, 0.59 (29/49). This significant relationship is confirmed by the results

Table 11.4. The features of exits.

	Number of firms	Average size of employment in weaving	Average size of total employment	Average size of production	Number of adoptors	Total employment in weaving	Total production	Total employment
1974	6	206	989	40,664	1	1,236	243,985	5,934
1975	6	47	50	15,455	3	282	92,731	300
1976	7	168	259	56,196	1	1,176	393,375	1,813
1977	8	296	795	73,028	4	2,368	584,230	6,360
1978	2	22	30	31,612	2	44	94,837	60
1979	1	58	93	36,300	1	58	36,000	93
1980	10	196	412	43,789	1	1,950	438,000	4,120
1981	4	62	87	19,178	4	4	76,000	348
1982	4	41	50		1	200	71,000	200
Total	48				19	7,572	2,030	19,228

Table 11.5. The features of entrants.

	N	Size of employment in weaving	Size of total employment	Average production in tons	Total employment in weaving	Total production	Total employment
1974	2	102	102	49,310	204	98,620	204
1975	0	0	0				
1976	2	119	222	63,622	238	127,245	444
1977	0	0	0				
1978	1	13	13	6,250	13	6,250	13
1979	2	182	182	336,984	364	1,010,943	364
1980	4	144	291	148,475	576	593,000	1,164
1981	6	292	339	90,293	1,752	541,000	2,034
1982	0	0	0	0			
Total	17				3,147	2,376	4,223

of the test of the probit equation (11.4) where the chances of exit depend upon the time lag in the adoption:

$$\text{EXIT} = a + b_1 \text{ ADOPTIONLAG.} \tag{11.4}$$

The dependent dichotomous variable EXIT equals 0 if the firm is still active in 1982 and 1 if it exits earlier. The independent variable ADOPTIONLAG measures the time lag in adoption.[9]

The test of equation (11.4) yields a strong and significant result:

$$\begin{aligned} \text{EXIT} = {} & \underset{(11.97)}{1.316} + \underset{(14.91)}{0.688} \text{ ADOPTIONLAG} \\ & \text{MODEL CHI-SQUARE} = 16.25 \quad \text{PROBABILITY} = 0.0001 \end{aligned} \tag{11.4a}$$

(partial chi-squared statistics between parantheses). Chances of death are strongly and positively associated with the size of adoption lag: thus the longer the delay in adopting the process innovation the greater the chances of exit.

The exit of 48 firms left considerable room in the industry. Early adopters, however, did not take advantage of this chance to increase their size and market share.[10] New firms entered the market taking advantage of the new process innovation: the analysis of the entry process (see *Table 11.5*) suggests, in fact, that to a large extent the room left by firms forced to exit has been occupied by new entrants. In the period 1974–1982, 17 new firms entered the cotton-weaving industry; 1 left in 1977; and 16 survived.

Features of the newcomers seem worth noting (see *Table 11.5*):

(1) New large firms entered the industry and filled the gap left by many small firms. The average production level from new entrants was over 139,000 tons a year; the average production level from exits was 46,000 tons a year.

(2) The entry made it possible to expand the production capacity of the industry. Newcomers added to total production of the cotton-weaving industry a capacity of 2,376 million tons (calculated by the sum of first year productions) compared with a total loss caused by exits of 2,030 million tons (calculated as the sum of last year productions). The entry process, thus, added a production capacity of 0.356 million tons a year.

(3) The newcomers, however, had a much lower level of total employment with 3,147 units in the year of entry, while firms forced to leave the industry had a total employment of 7,572 units. In fact, average physical productivity of workers in the new firms, measured in the year of entry, was 755 tons per head, against the 268 tons per head of firms forced to exit, measured in the last year of production.

(4) The entry was favored by the technological opportunity offered by the shuttle-less looms. Out of 17 entrants, by 1982, 14 had already adopted the shuttle-less looms, and 13 of them actually entered with shuttle-less looms already installed or adopted shuttle-less looms within one year.

(5) The contribution of newcomers to the diffusion process of shuttle-less looms has been impressive (see *Table 11.6*): the share of shuttle-less looms installed by newcomers grew from 4.7% in 1976 to 28.4% in 1982.

As clearly shown by the data in *Table 11.7* the aggregate growth of the stock of shuttle-less looms is the result of three different forces:

- Interfirms diffusion, i.e., adoptions by existing firms.
- Entry, i.e., adoption by newcomers.
- Intrafirm diffusion, i.e., new adoptions by early adopters.[11]

Analysis of the evolution in the time span considered of the contribution of the three forces suggests that the actual interfirm diffusion process among existing firms played an important role in the years 1974–1977. In the following years the entry of new firms based on modern capital equipment such as on shuttle-less looms gave the strongest contribution to the aggregate diffusion process. The intrafirm diffusion contributed with an irregular pace.

11.3. Conclusions

In recent years with the introduction of equilibrium models, the economic analysis of diffusion has made important progress and now has better tools to account for the dynamic effect of incremental technical change, learning by doing, learning by using, and interaction between demand and supply on the microeconomic rationale of the adoption of process innovations.

According to equilibrium models the diffusion of process innovation takes place because of the growing profitability of innovations – engendered by a variety of dynamic forces – and the consequent increase in the scope of application and opportunity for adoption by new categories of potential users.

Population of potential users, however, is considered to remain at a given and static level: such an assumption seems unrealistic and unnecessary.

The timely adoption of radical process innovation can become a factor of change on potential adopters; the survival of late adopters can be challenged when production costs made possible by the adoption of the new technique are sensibly lower than production costs with the old technique.

The empirical evidence on the diffusion of shuttle-less looms in the Italian cotton-weaving industry confirms that the introduction of the new weaving technology gave a major competitive advantage to early adopters.

More specifically the diffusion of shuttle-less looms in the Italian cotton-weaving industry has paralleled and fostered a major selection process. Early adopters have been able to survive but not necessarily to increase their market shares. Late adopters never had the chance to adopt: they were forced to exit. Their place has been taken by new large firms that entered the industry with shuttle-less looms in operation.

Table 11.6. The dynamics of entry.

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
Stock of shuttle-less looms installed by old firms	661	974	1,409	1,719	1,902	1,524	1,453	1,746	2,137	2,130
Number of firms entered	–	2	0	2	0	1	2	4	6	–
Stock of shuttle-less looms installed by new firms	–	0	0	86	97	104	233	256	715	807
Total stock of shuttle-less looms	661	974	1,409	1,805	1,999	1,628	1,686	2,002	2,852	2,937
Share of shuttle-less looms installed by new firms	–	0	0	4.7	4.8	6.4	13.8	12.8	25.0	28.4

Table 11.7. The contribution of new adoptions, newcomers, and intrafirm diffusion to the aggregate diffusion of shuttle-less looms.

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*
Stock of shuttle-less looms	661	974	1,409	1,805	1,999	1,628	1,686	2,002	2,852	2,937
Increase of shuttle-less looms	–	313	435	396	194	−371	58	316	850	85
Number of shuttle-less looms installed by new adopters	NA	276	263	109	56	12	0	74	87	ND
share on total increase (A)[a]	–	88%	60%	27%	28%	–	0	23%	10%	–
Number of shuttle-less looms installed by new firms/share	NA	0	0	86	13	0	27	160	460	ND
in total increase (B)[b]	–	0	0	21%	6%	–	46%	50%	54%	–
Contribution of the intrafirm diffusion process $[100 - (A + B)]$		12	40	52	66	–	54	37	36	

[a] Contribution of the adoption by new adopters.

[b] Contribution of the adoption by newcomers.

The aggregate diffusion pattern does follow a sigmoid pattern; such a path, however, is not explained by the sequence of adoption of large firms followed by small ones, as supposed by equilibrium models, but rather by a sequence characterized by the early adoption of large, efficient firms; the exit of late adopters; and the entry of new firms based on modern capital equipment.

In sum, the evidence analyzed suggests that the diffusion of process innovations can be the result of a selection process in the population of potential adopters, which shapes the industrial structures according to the differential ability of firms to adopt at the right time to available process innovations.

Such a case of diffusion of a process innovation that parallels and shapes a process of industrial reorganization stresses the need to analyze the diffusion of innovation in a dynamic context.

Notes

[1] The level of the threshold depends upon the cumulative lognormal and, respectively, cumulative normal path of diffusion which in turn is determined by the role of supply forces, learning by doing, and learning by using (Davies, 1979).

[2] In other words, in an epidemic context the diffusion of innovations should be instantaneous given that all potential adopters have the same amount of outguessing capacity. No diffusion would take place if early adopters have no chance of communicating their experience and satisfaction to potential adopters.

[3] The role of large firms as early adopters has been confirmed also by the results of the empirical analysis of diffusion based on the disequilibrium approach. See Mansfield (1963 and 1966), Metcalfe (1971), Globerman (1975), Romeo (1975). Consistent results of the analysis of the diffusion process of oxygen furnaces in the US steel industry, however, suggest that larger firms can be late adopters rather than early ones. See Adam and Dirlam (1966), Oster (1982), and Karlson (1986).

[4] Antonelli (1986) provides an empirical test of the positive effects of higher rates of international diffusion of a process innovation such as shuttle-less looms on international competitiveness as measured by the distribution of international market share in the cotton fabric products.

[5] It should be noted that, *ceteris paribus* of market forms and product differentiation, the effect of the reduction in costs engendered by the adoption of a process innovation, on the increase in supply of adopters, ultimately depends on the changes in the shape of the new total and marginal cost curve. Only when a reduction in variable costs *flattens* the total marginal cost curve can a relevant positive increase in supply take place.

[6] International comparison of rates of diffusion of shuttle-less looms, calculated with interpolation of a logistic curve, for the period 1976–1983 places Italy as 10 out of 28. See Antonelli (1986).

[7] Calculated as the percent share of shuttle-less looms in 1971 and 1985 to total looms installed in 1971.

[8] This is confirmed by the insignificant result of the estimate of the probit equation tested on the 99 firms active in 1973

$$\begin{aligned} \text{EXIT} = \ & 3.160 \ - \ 0.0003 \\ & (2.821) \quad (0.602) \end{aligned}$$

$$R^2 = 0.0048 \quad F = 1.112$$

(partial chi square between parentheses), where EXIT is a dichotomous variable with value 1 for the firms that died during the period 1973–1982 and 0 otherwise; SIZE is measured by total employment in 1973

[9] ADOPTIONLAG = (YEAR OF ADOPTION – 1971) where for non-adopters 1990 has been set as the year of eventual adoption.

[10] The results of the test of equation (11.5) where GROWTH measures the increase in production in the period 1982–1973 for firms still active in 1982

$$\underset{}{\mathrm{GROWTH}} = \underset{(13.404)}{136.19} + \underset{(0.043)}{0.017}\ \mathrm{ADOPTIONLAG} \qquad (11.7)$$

$$R^2 = 0.010 \quad F = 0.907$$

(t statistics between parentheses) suggest that the hypothesis of a causal relationship between adoption and growth can be rejected.

[11] Lack of data on scrap and resale of used shuttle-less looms by ailing firms reduce the consistency of estimates on the role of the intrafirm diffusion process.

References

Adam, W. and Dirlam, J., 1966, Big Steel Invention and Innovation, *Quarterly Journal of Economics* 80:167–189.

Antonelli, C., 1986, The International Diffusion of Process Innovations and the Neotechnology Theory of International Trade, *Economic Notes* 1(15):60–82.

Antonelli, C., 1987, The International Diffusion of Process Innovations with Incremental Technical Change, Politecnico di Milano.

David, P.A., 1969, A Contribution to the Theory of Diffusion, Standford Center for Economic Growth Memorandum, n. 71.

Davies, S., 1979, *The Diffusion of Process Innovations*, Cambridge University Press, Cambridge, UK.

Globerman, S., 1975, Technological Diffusion in the Canadian Tool and Dye Industry, *Review of Economics and Statistics* November (42):428–434.

Griliches, Z., 1957, Hybrid Case: An Exploration in the Economics of Technical Change, *Econometrica* October (25):501–522.

Karlson, S.H., 1986, Adoption of Competing Inventions by United States Steel Producers, *Review of Economics and Statistics* August (53):415–422.

Mansfield, E, 1961, Technical Change and the Rate of Imitation, *Econometrica* October (29):741–766.

Mansfield, E, 1963, The Speed of Response of Firms to New Techniques, *Quarterly Journal of Economics* May (77):290–311.

Metcalfe, J.S., 1981, Impulse and Diffusion in the Study of Technical Change, *Futures* October (13):347–359.

Oster, S., 1982, The Diffusion of Innovation among Steel Firms: The Basic Oxygen Process, *Bell Journal of Economics* Spring (13):45–56.

Romeo, A.A., 1975, Interindustry and Interfirm Differences in the Rate of Diffusion of an Innovation, *Review of Economics and Statistics* August (57):311–319.

Rosseger, G., 1979, Diffusion and Technological Specificity, *Journal of Industrial Economics* September (28):39–53.

Stoneman, P., 1983, *The Economic Analysis of Technological Change*, Oxford University Press, Oxford, UK.

CHAPTER 12

Determinants of Energy Technology Transitions: A Cost-Benefit Analysis Approach

Oliver S. Yu and Shu-Dong He

12.1. Introduction

The broad pattern of energy transition from muscle power to wood to coal and then to oil, gas, and electricity has been widely observed in the histories of many countries. The driving forces of this transition process seem complex because they involve a multitude of technological innovations and a wide variety of socioeconomic factors for which data are scarce and difficult to obtain. As a result, researchers, such as Marchetti and Nakicenovic (1979), have often concentrated on empirical modeling and analysis of the energy market share evolution over time. Although highly useful and intellectually provocative, such models and analyses generally do not provide an in-depth understanding of the underlying energy transition process.

An alternative to the empirical analysis is the simulation approach. This approach models technology transitions as the result of the economic and management decisions of the users to substitute one technology for another to maximize the utilities derived from technology use. However, most such studies have analyzed technology substitution by examining important technology attributes under a fixed set of user preferences (e.g., Stern *et al.*, 1975). This study attempts to broaden the approach by analyzing technology substitution under *changing* user preferences. Specifically, it models energy transitions as the aggregated result of shifting technology choices by energy users in the process of optimizing the perceived costs and benefits of producing or consuming goods and services under *evolving* social and economic conditions. In other words, evolving socioeconomic factors systematically alter the relative importance of technology attributes and cause the successive rises and falls in energy technology.

The purpose of this study is to show that the simple cost-benefit analysis approach can be used to simulate the energy technology decision process, and approximate the energy transition process. Ultimately, understanding these determinants and transitions will help us answer key questions on future energy trends, such as will electricity intensity continue to increase in the US economy, and address important energy policy issues, such as the changes that are needed and possible if we wish to revive the US nuclear power industry.

To carry out this study, we first identify the major direct and indirect costs and benefits of energy technology use and the socioeconomic and technology factors affecting these costs and benefits. We then use specific examples – home heating and industrial motive power – to illustrate how these factors affect technology choices and cause technology transitions.

As an initial analysis, the study deliberately ignores the effects of energy technology on economic development and the origin of technology innovation. Essentially, it assumes that various energy technologies have always been available, and waiting to be selected by the users under the optimal conditions. This assumption is reasonable if the results of the study are used mainly for understanding the potential energy technology transitions in the near future or the likely technology choice for a developing country, as the technology candidates are dominated by those currently available.

12.2. Perceived Costs and Benefits of Energy Technology Users

There are two types of energy technology users: the business users who use energy technologies in the production of goods and services (including energy services) and the individual users who use energy technologies in the consumption of goods and services. Because the choice of energy technologies in both cases eventually rests in human hands, the human perception of costs of and benefits from energy technology use is fundamental to the decision process.

The major perceived costs and benefits of energy technology use are listed in *Table 12.1*. In this table, we have grouped the more uniformly observable and readily quantifiable costs and benefits in the direct category. Costs and benefits in the indirect category are generally more perception-dependent and difficult to quantify, yet they can be influential to energy technology choice. Even those in the direct category, such as the cost of meeting regulatory requirements and the benefit of energy supply reliability, can be difficult to determine. For example, the purpose of meeting regulatory requirements, in principle, is to reduce the perceived public risk resulting from specific technology use to a socially acceptable level. It is difficult, however, to integrate the perception of risk over an entire society. Furthermore, this perception evolves not only with the degree of exposure, knowledge, and understanding, but also with many other socioeconomic factors, such as personal income.

Table 12.1. Major perceived costs and benefits of energy technology use.

	Business user	*Individual user*
Costs		
Direct	Production	Consumption or use
	(capital, labor, energy,	(capital, labor, energy,
	material, time, and space)	material, time, and space)
	Meeting regulatory requirements	Health and safety hazards
Indirect	Additional environmental impacts	Environmental and aesthetic effects
	Labor and organizational dislocations	Psychological effects
Benefits		
Direct	Production cost savings	Consumption cost savings
	Production reliability	Use reliability
Indirect	Avoid labor and time	Avoid labor and time
	(controllability and flexibility)	(convenience)
	Overall productivity improvement	Overall quality of life improvement
	Image and prestige	Image and prestige

12.3. Major Factors Affecting the Perceived Primary Costs and Benefits of Energy Technology Use

The major socioeconomic and technology factors affecting the perceived costs and benefits of energy technology use are summarized in *Table 12.2.* The first and probably most important factor is personal income. In aggregate, the personal income in an economy affects the availability of capital and, hence, the cost of capital both to businesses and to individuals. The costs of time and labor also rise with average personal income. More importantly, aggregate consumer preferences represented by consumption patterns (*Figure 12.1*) change with income (or per capita GNP), which in turn change production patterns of the economy (*Figure 12.2*). These changes affect the relative costs and benefits of energy technologies. For example, expanding proportions of discretionary expenditures for recreation reflect the increasing value of electricity for electronic entertainment and oil for transportation use. Another example is that as health expenditures rise, both in quantity and in proportion to total expenditures with personal income, public concerns about the environmental effects of energy technology are also likely to heighten. Moreover, the "worth of life," perceived either through total expected life earnings or through "the willingness to pay," increases with personal income. As a result, the perceived costs of health and safety hazards for individual users increase with personal income, which in turn result in added regulatory requirements and costs for energy technologies and their use.

The second factor is population concentration or the rate of urbanization. This factor directly affects the cost of land for energy facilities and the space for fuel storage. It generally raises the costs of accessing wood, as urban centers tend to locate away from forestland. Population concentration owing to urbanization, however, provides the economies of scale for building the infrastructure for energy delivery. On the other hand, crowding in urban centers is prone to drive up regulatory requirements.

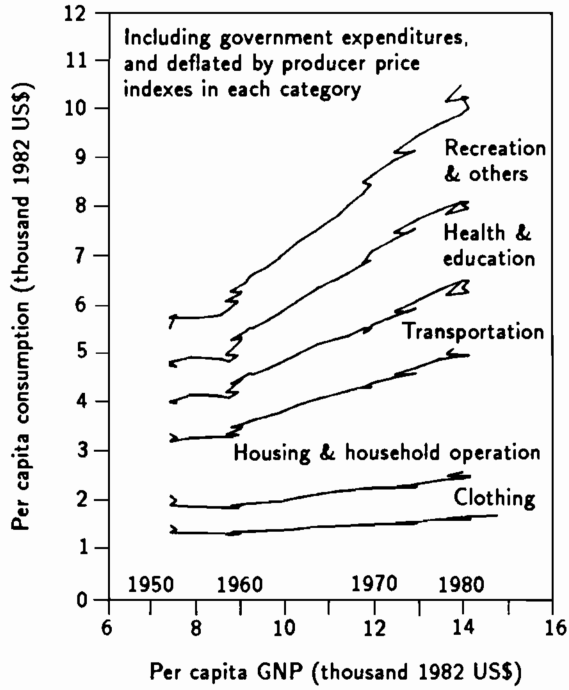


Figure 12.1. US consumption pattern changes by expenditure category.

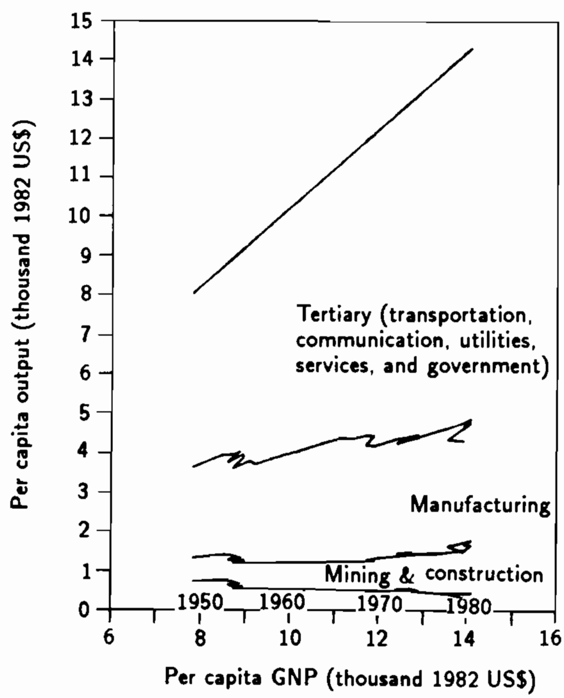


Figure 12.2. US production pattern changes by originating sector.

Table 12.2. Major factors affecting costs and benefits of energy technology use.

Factor	*Costs and benefits affected*
Personal income	Production costs particularly cost of capital Consumption costs particularly labor and time All other costs and benefits due to impact on consumption and production patterns
Urbanization scale	Production and consumption costs due to the effects of concentration and economy Cost of meeting regulatory requirements
Industrialization	Production costs due to the effects of high utilization and scale economy Cost of meeting regulatory requirements
Energy availability	Production and consumption costs particularly energy costs Cost of meeting regulatory requirements due to national security concerns Image and prestige
Energy form	Avoided labor and time in production and consumption Overall productivity and quality of life improvement

The third factor is the degree of industrialization. The high levels and concentration of energy in industries provide scale economies for energy production and delivery, but they are also likely to induce heavy regulatory requirements.

These three factors, income growth, urbanization, and industrialization, are broadly related to the process of economic development (Rostow, 1962). However, this relation is not necessary; for example, an OPEC nation can have high personal income without the attendant urbanization or industrialization.

The fourth factor is the availability of energy resources. Low availability results in the high cost of these resources. Furthermore, dependence on foreign energy resources often creates national security concerns and prestige issues and gives rise to regulatory requirements or public policies in favor of more reliable alternatives.

The last factor is the intrinsic characteristics of the energy form and its ability to facilitate innovative uses, which is difficult to specify. Nevertheless, we do see, for example, that oil- and gas-based plastic technologies and electricity-based information technologies have given oil, gas, and electricity distinct advantages over the direct use of wood and coal. The industrial motive power example below will provide a further illustration of the value of energy form.

12.4. Application of the Cost-Benefit Analysis Approach to Understanding of Energy Technology Transitions

We will apply the cost-benefit analysis approach to simulate technology transitions in two specific cases of energy use: home heating and industrial motive power.

12.4.1. Home heating

Based on the cost-benefit analysis approach, the choice of home heating energy technology by individual users may be modeled as a minimization of the total annualized cost of providing a given level of heating comfort. This cost includes the cost of heating appliance investment and operation, and fuel purchase and storage. We will first analyze the historical energy transitions in US home heating, and then examine the residential energy technology choices in different economies of the world.

In this analysis, we will use per capita real GNP as the proxy for economic development, which has pervasive effects on both direct and indirect costs of energy technology use. *Figure 12.3* shows the energy transition in US home heating from wood to coal and then to oil, gas, and electricity as per capita GNP increases. The direct costs of wood, coal, and oil (as a stand-in also for gas and electricity) are shown in *Figure 12.4*. Based on these direct fuel costs, assuming equal capital costs and lifetimes for the heating appliances, the historical transition seems to defy economic common sense because it moved from the lowest-cost fuel (wood) to increasingly expensive fuels. However, if we consider the total cost to the individual user, including the indirect costs of fuel storage and labor for fuel handling and appliance operation, this transition becomes sensible. As shown by the simplified analysis in *Table 12.3*, as storage and labor costs increase respectively with urbanization and income, total cost of home heating, also show in *Figure 12.4*, becomes highest for wood and lowest for oil, gas, and electricity. This simple analysis also simulates recent energy transitions in the US home heating. As oil costs rose sharply in the 1970s, coal and wood heating became once again competitive and enjoyed a brief revival until the oil price dropped again in the 1980s.

The same logic applies to historical transitions in US residential energy use shown in *Figure 12.5*. It is interesting to see in *Figure 12.6* that, in spite of the large data noises owing to currency conversions and differences in energy resource availability, the technology choice for residential energy use of 44 countries as a function of per capita GNP forms a pattern similar to the US historical experience. This similarity indicates that the simple cost-benefit analysis approach can be used to understand the major driving forces in energy transitions.

This case study demonstrates that energy transitions in home heating were the aggregated results of shifting technology choices by individual users in the process of minimizing the direct and indirect costs of heating energy consumption. Personal income, energy cost and availability, and the space- and labor-saving characteristics of oil, gas, and electricity were the key determinants of the shift from wood to coal and to oil, gas, and electricity.

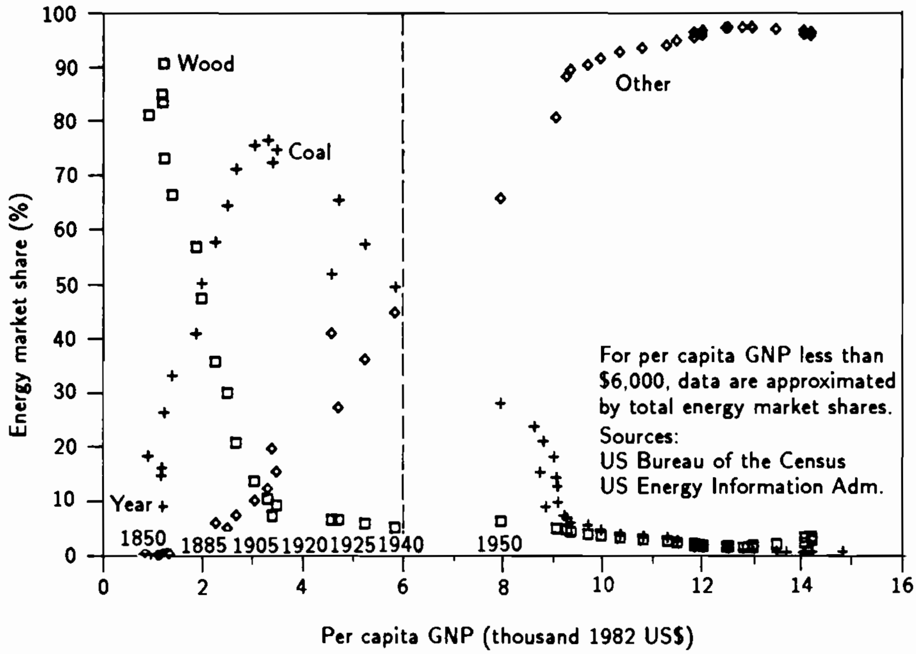


Figure 12.3. Energy transition patterns in US home heating.

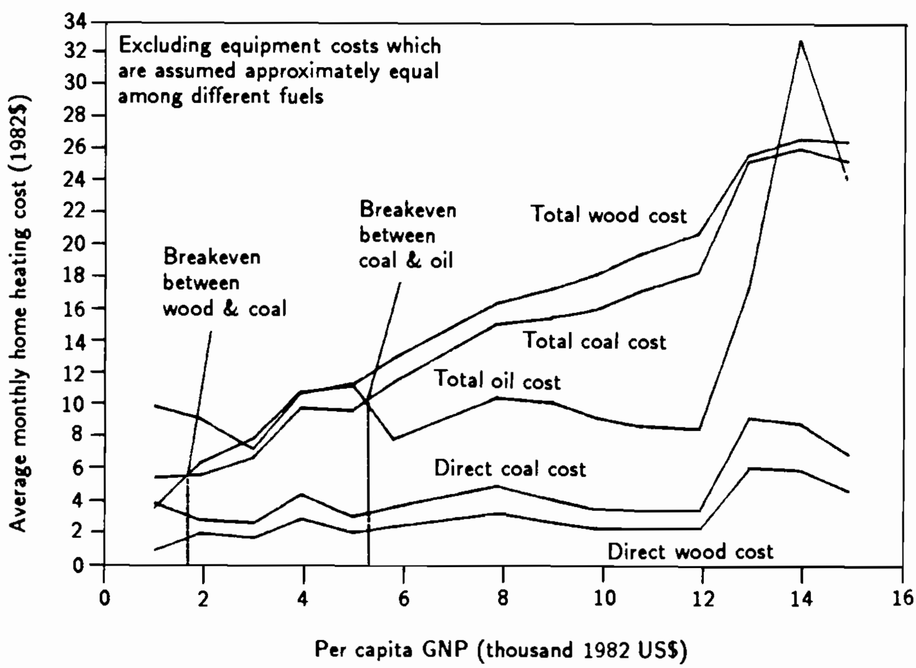


Figure 12.4. Comparison of US home heating costs.

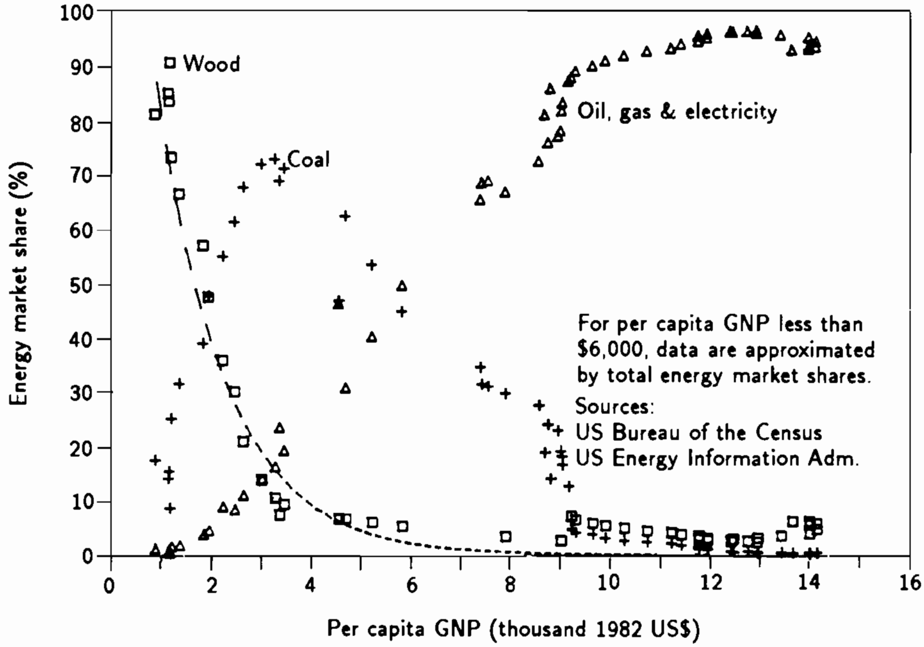


Figure 12.5. **Energy transition patterns in US residential sector.**

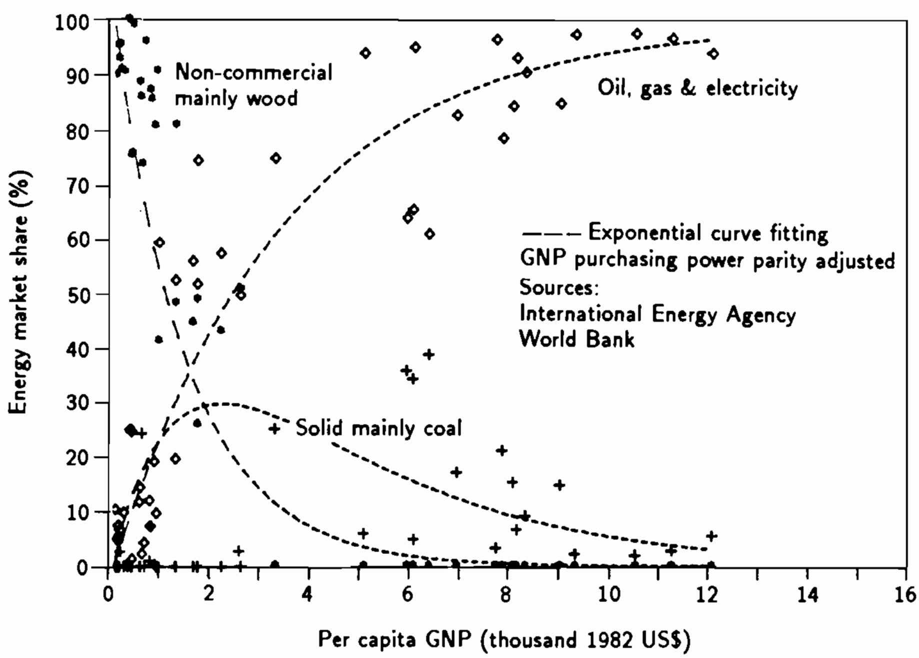


Figure 12.6. **Energy technology choices in 1972 in the residential sector of 44 countries.**

Table 12.3. Comparison of US home heating costs (1982 $).

			Average monthly heating cost per household[a]								
			Fuelwood				*Coal*				*Oil*
Year	*GNP per capita*	*Urban rate %*	*Direct*	*Stor*[c]	*Hand*[c]	*Total*[d]	*Direct*	*Stor*[c]	*Hand*[c]	*Total*[d]	*Total*
1867	1,000	24	1.0	1.5	1.1	3.6	4.0	0.4	1.1	5.5	10.0
1885	2,000	33	2.1	2.1	2.3	6.5	2.9	0.5	2.3	5.7	9.2
1905	3,000	43	1.8[b]	2.8	3.4	8.0	2.7	0.7	3.4	6.8	7.3
1921	4,000	51	3.0[b]	3.3	4.6	10.9	4.5	0.8	4.6	9.9	10.8
1926	5,000	54	2.1[b]	3.5	5.7	11.3	3.1	0.9	5.7	9.7	11.4
1940	6,000	57	2.5[b]	3.6	6.9	13.0	3.7	0.9	6.9	11.5	7.9
1950	8,000	60	3.3[b]	4.0	9.1	16.4	5.0	1.0	9.1	15.1	10.5
1955	9,000	63	2.7[b]	4.3	10.3	17.3	4.1	1.1	10.3	15.5	10.2
1963	10,000	67	2.3[b]	4.5	11.4	18.2	3.5	1.1	11.4	16.0	9.2
1965	11,000	71	2.3[b]	4.5	12.6	19.4	3.4	1.1	12.6	17.1	8.7
1969	12,000	73	2.3[b]	4.7	13.7	20.7	3.4	1.2	13.7	18.3	8.5
1976	13,000	73	6.1[b]	4.7	14.8	25.6	9.2	1.2	14.8	25.2	17.3
1980	14,000	73	5.9[b]	4.7	16.0	26.6	8.8	1.2	16.0	26.0	32.8
1985	15,000	73	4.6[b]	4.7	17.1	26.4	6.9	1.2	17.1	25.2	24.0

	Average monthly home heating requirements				*Space cost*
	Quantity	*Space*	*Handling*		*per ft^2*
Wood	0.50 cord	16 ft^2	10 man-hours	Rural	Negligible
Coal	0.25 ton	4 ft^2	10 man-hours	Urban	0.4[e]
Oil	1.00 bbl	Negligible	Negligible		

[a]Excluding equipment costs, which are assumed approximately equal among different fuels.
[b]Estimated at 2/3 of coal direct costs.
[c]Stor = Storage = urban space cost × urbanization rate. Hand = Handling = man-hours × per capital GNP/(8,760 hours per year).
[d]Total = Direct + Storage + Handling.
[e]1,200 ft^2 home rented at $500/mo (1982$).
(Sources: Direct costs prior to 1955 and home heating quality estimates based on Schurr and Netschert, 1960; direct costs after 1955 and space cost based on US EIA Annual Energy Review and US Statistical Abstract.)

12.4.2. Industrial motive power

As shown in *Figure 12.7*, between 1870 and 1940, the production and distribution of motive power in US industries shifted from waterwheels to steam-driven prime movers and finally to electric drive. To document the first transition, Atack (1979) used a simulation approach to show that manufacturers adopted steam power over water primarily for reasons of locational and seasonal availability and of direct cost savings.

The shift from steam engines to electric motors was more complex. Prior to the introduction of electric motors in the late nineteenth century, production machines in a factory were connected to a centrally located prime mover, such as a waterwheel or steam engine, by line shafts, belts, and pulleys. Electric motors first competed with steam engines in providing these "line drives." The ability of electric motors to be miniaturized soon replaced the line drives by unit drives

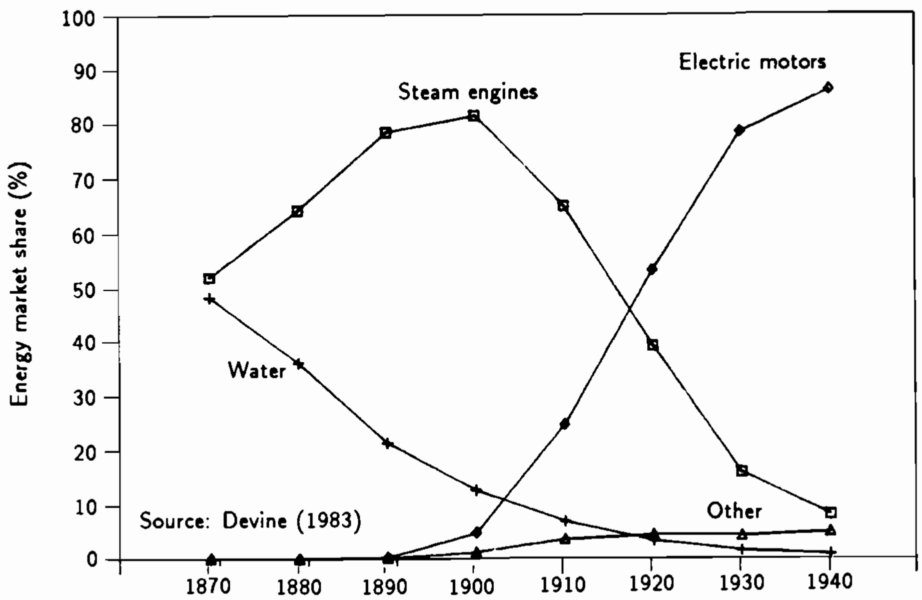


Figure 12.7. **Energy transition patterns in US industrial motive power.**

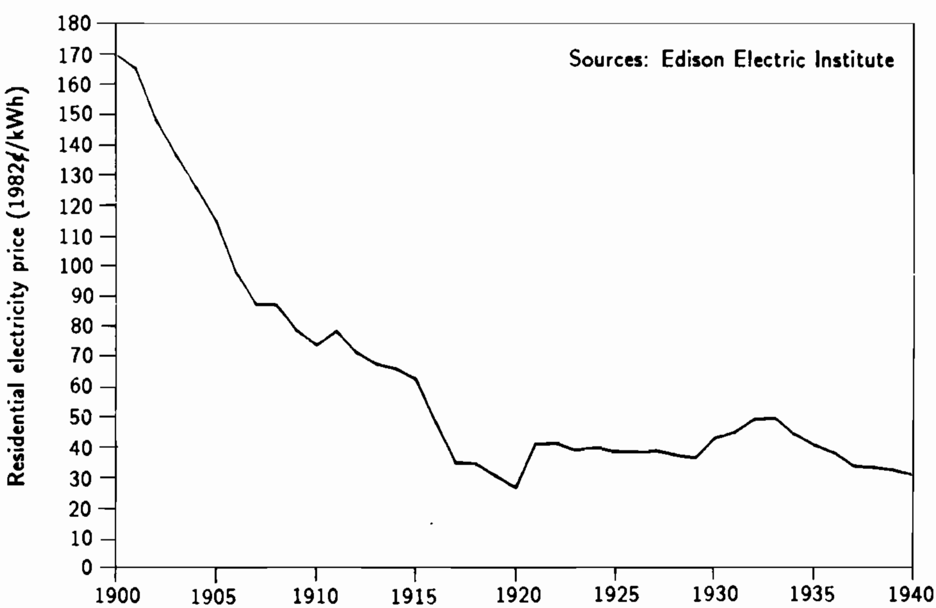


Figure 12.8. US electricity price, 1900–1940.

located closely to the production machines. Direct energy cost savings played a role in this substitution. Rapid economic growth during this period provided economies of scale to electricity production that, together with thermal efficiency improvements, drastically reduced electricity costs (***Figure 12.8***). In addition, the energy efficiency of electric drive made it even more directly cost-competitive against the steam engines (***Table 12.4***). However, the shift from steam to electric motive power, as observed by Devine (1983), was motivated primarily by the pervasive productivity improvement potential of the unit motor drive. Unfortunately, these important indirect benefits from overall manufacturing productivity improvement are difficult to quantify. The exhaustive research by Devine lists the following major indirect benefits from unit drive:

- Increased the flow of production in factories. Unit drive gave manufacturers flexibility in the design of factory buildings and in the arrangement of machinery to maximize production throughput and minimize material handling. Furthermore, production machines could be easily rearranged to match changing production requirements.
- Improved the working environment. Absence of overhead shaft drives led to improvements in illumination, ventilation, and cleanliness, and contributed to the quantity and quality of work.
- Improved machine control. The speed of an electric unit motor drive could be precisely and individually controlled to respond to specific production needs.
- Facilitated plant expansion. The locational flexibility of electric unit drive removed the constraints on plant expansion imposed by the rigid line drive. In a case study of the Scoville Manufacturing Company of Waterbury, Connecticut, Kapstein (1981) argues that the removal of constraints on expansion of production was the primary reasons for the company's switch to electric unit drive.

As Devine observes:

> Although unit drive used less energy and sometimes cost less than other methods of driving machinery, manufacturers came to find these savings to be far less important than their gains from increased production. With unit drive, electricity was used with its greatest economic advantage: processes could be arranged within factories to maximize throughput; plants could be more readily expanded; and a better working environment and improved machine control increased both quantity and quality of output. In essence, electric unit drive offered opportunity – through innovation in processes and procedures – to obtain greater output of goods per unit of capital, labor, energy, and materials employed. Electricity was now viewed as a factor in improving overall productive efficiency.

This observation appears to be supported by the productivity improvement in US manufacturing following the penetration of electric motors shown in ***Figure 12.9***.

Table 12.4. Representative full-load efficiencies of power production, transmissions, and distribution in manufacturing, 1900–1920.

	Boiler steam turbine gen.	*Elect. power trans.*	*Boiler & steam eng.*	*Elect. gen.*	*Elect. power dist.*	*Motors*	*Mech. power dist.*	*Overall*
Direct drive			8–10%[a]				33–75%	3–8%
Self-gen. & group drive			8–10%	92%	92%	77–90%	50–75%	3–6%
Self-gen. & unit drive			8–10%	92%	92%	77–90%		5–8%
Utility gen. & group drive	16%[b]	90%			92%	77–90%	50–75%	5–9%
Utility gen. & unit drive	16%	90%			92%	77–90%		10–12%

[a]Corresponds to 2–2.5 pounds coal per horsepower-hour.
[b]Corresponds to 1.7 pounds coal per kilowatt-hour, typical of the Detroit Edison Company in 1922. This efficiency was somewhat higher than the national average.

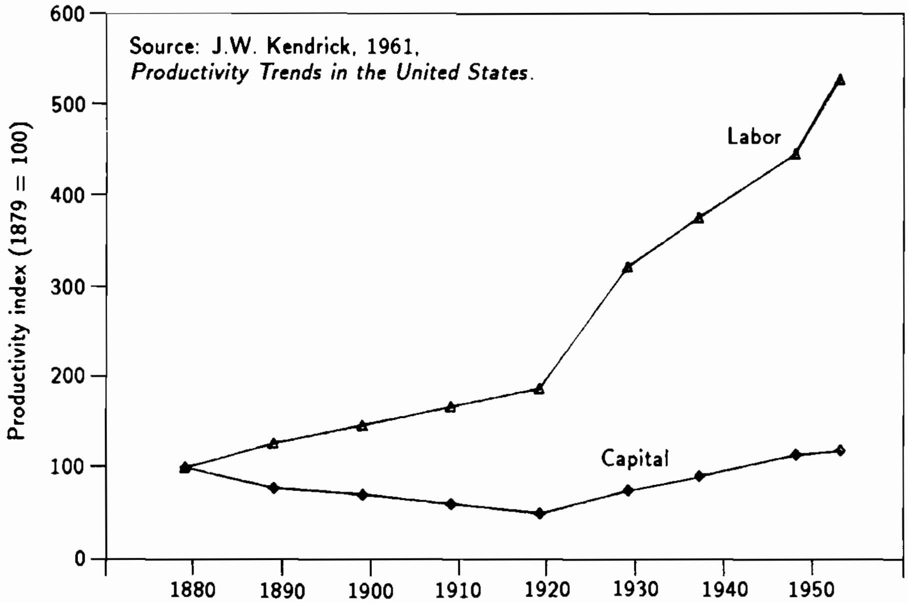


Figure 12.9. Capital and labor productivity in US manufacturing, 1879–1953.

This case study illustrates that energy transitions in industrial motive drive were the aggregated results of shifting technology choices by business users' in the process of optimizing the direct and indirect costs and benefits of production. Economic growth and the form value of electricity had major effects on these costs and benefits and were the key determinants of the shift from steam engines to electric motors.

12.5. Summary and Conclusions

Understanding the driving forces in energy transition and technology choice is of vital importance to planning both technology R&D and national policy. The simple analyses presented in this paper demonstrate that the cost-benefit analysis approach can be used to approximate the aggregate energy technology decision-making process under changing socioeconomic conditions. Using this approach, it identifies economic development, energy availability, and the energy form as the major factors affecting the perceived costs and benefits of energy technology use by businesses and individuals. These factors are, thus, the key determinants of energy technology transitions.

Owing to data paucity and definitional complexity, the effects of these determinants on perceived costs and benefits are far from being fully developed. This problem is particularly acute for many indirect costs and benefits that have pivotal importance in energy technology decisions. Continued data collection, analysis, and refined modeling of these effects are essential to developing an understanding of future technology transitions.

Appendix: The empirical estimates by Ersson and Lane (1985)

The regression model is

$$\text{GROWTH} = a+b * \text{GNP} + c * \text{MODERNIZATION},$$

where economic growth 1960–1977 is used as the GROWTH variable and MODERNIZATION is measured by the time of modernization of the country according to Taylor and Hudson (1972). The following selection of countries have been used:

N = 78 OECD countries, LDCs, and socialist countries.
N = 70 without socialist countries.
N = 46 without OECD countries, i.e., mostly LDCs.
N = 24 OECD countries.

Estimates of the parameter c (t-value in parenthesis).

$N = 78$	$N = 70$	$N = 46$	$N = 24$
−.36	−.31	−.14	+24
(−2.1)	(−1.6)	(−.90)	(−1.3)

References

Atack, J., 1979, Fact in Fiction? The Relative Costs of Steam and Water Power: A Simulation Approach, *Explorations in Economic History* **16**:409–437.

Devine, W.D., Jr., 1983, From Shafts to Wires: Historical Perspective on Electrification, *The Journal of Economic History* June:347–372,

Kapstein, E.B., 1981, Industrial Power at Scoville Manufacturing Company: A Research Note, unpublished manuscript, Kennedy School of Government, Harvard University, Cambridge, MA, USA.

Marchetti, C. and Nakicenovic, N., 1979, *The Dynamics of Energy Systems and the Logistic Substitution Model*, RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rostow, W.W., 1962, *The Process of Economic Growth*, 2nd Ed., W.W. Norton, New York, NY, USA.

Schurr, S.H. and Netschert, B.C., 1960, *Energy in the American Economy, 1850–1975*, Johns Hopkins University Press, Baltimore, MD, USA.

Stern, M.O., Ayres, R.U., and Shapanka, A., 1975, A Model for Forecasting the Substitution of One Technology for Another, *Technological Forecasting and Social Change* **7**:57–79.

CHAPTER 13

Scientific–Technological and Educational Cycles: Interconnection and Planned Implementation

Y. V. Yakovets

The most characteristic feature of the last quarter of the twentieth century is a profound upheaval in science, technology, education, ecology, and organizational forms of production and management. These upheavals, taken together, form the content of the contemporary revolution in productive forces as the basis for the next stage in the progressive development of society. It is necessary to investigate the structure and interrelations of revolutionary leaps according to the principal elements of these productive forces.

The structure of scientific–technological cycles is considered in the following ways:

- The five phases of cycles (pre-birth, mastering, diffusion, stable development, and aging – change of the next cycle).
- The kinds of cycles – scientific, invention, innovation, and technological cycles in certain succession.
- The depth of revolutionary leaps, changes of generations of technology, scientific–technological directions, and scientific–technological revolutions, when transitions to new technological principles embrace leading branches of industry.

Now a transition to new generations of technology is being carried out nearly once every 8–10 years; to new directions, once every 40–50 years. The life cycle is longer – 15–20 and 70–90 years.

Historically, nearly 17 comprehensive upheavals in the technical basis of society and 7 revolutions in science have been observed. I regard the contemporary upheaval as the second scientific–technological revolution. The third revolution may be expected between 2010–2020.

Every revolution has its own structure. In the structure of the second scientific–technological revolution, one may single out a nucleus (microelectronics, biotechnology, and informatics); a radical transformation of production technology (complex flexible automation of production, new, highly effective technologies, new generations of materials, mastering of nontraditional energy sources, new kinds of transportation and communication, industrial mastering of outer space and the ocean); a qualitative leap in the service sphere (medicine, education, science, personal consumption); and a new military-technical revolution.

The modern scientific–technological revolution begins in a number of developed countries, fully prepared to accept it, then it spreads from the epicenter embracing more and more countries. The countries formerly lagging behind may skip several stages in their technological development, on the basis of international division and cooperation of labor in this particular field.

The consequences of scientific–technological revolutions are a leaplike increase of labor productivity, a sharp renovation and considerable expansion of the assortment of the goods produced, solutions to ecological crises, and affordable machinery.

But a revolutionary upheaval engenders new contradictions – a reduced demand for man power, a growth in unemployment in developed capitalist countries, and the creation of new weapons of mass destruction.

Scientific–technological cycles are interconnected with the educational cycles. The structure of educational cycles includes:

- Changes from generation to generation in terms of knowledge, skills, and qualification levels, as well as by forms of labor organization.
- Periodical revolutions in education, which are accompanied by qualitative leaps in content and technical methods of teaching and in the organizational forms of education.

Changes between generations are likely to occur every 30 to 40 years. Revolutions in education occur less often than upheavals in technology and are connected with scientific revolutions.

The peculiar features of the current revolution in education (the seventh in the history of mankind) are the restructuring of comprehensive and professional education as applied to the structure of modern scientific–technological revolution; the introduction of active methods and effective technical means of teaching on the basis of computer instruction, videotechnology, and TV; and the creation of the system of unbroken education, systematic reviewing, and increasing knowledge throughout the period of active labor performance of man.

It is necessary to elaborate the effective mechanism of planned use of scientific–technological and educational cycles. In the USSR this mechanism includes a transition to the through-planning of the scientific–technological progress on the basis of economic methods and reconstruction of the system of personnel training and retraining.

The essence of through-planning of scientific and technological progress is a strategic breakthrough on the basis of mastering and diffusing new highly efficient generations of technology. The through-planning embraces all levels of management: national economy, branch, territorial, and enterprise level. It is necessary to combine state orders with business contracts, financial and credit advantages, differentiated flexible prices, and economic incentives for the enterprises.

Various forms of integrating research and production spheres are being developed: inter-branch scientific–technological complexes, temporary creative teams, scientific–industrial associations, and engineering centers.

The reconstruction of an educational system includes the reform of the comprehensive school and vocational training, profound restructuring of higher and secondary education, and qualification improvement. Conditions are provided to activate the human factor of the scientific–technological progress; technical creativity is encouraged on a wide scale.

Scientific–technological and educational revolutions are by nature global. Even greater now are the possibilities of international cooperation in research and development for implementing global scientific and technological programs and for assisting developing countries to master the new technologies to carry out the educational revolution.

It is advisable to unite the efforts of scientists from different countries to study the regularities of the scientific, technological, and educational development, to organize forms of production and management, and to elaborate effective methods of implementing these regularities.

CHAPTER 14

Long-Term Movement of the Economy: Terminology and Theoretical Options

J.-L. Escudier

14.1. Introduction

Since the beginning of the twentieth century, analysis of possible long movements in the economy has resulted in a large volume of literature written by economists but also – more recently – by certain historians. However, from one author to another the words used differ according to the language used (which is increasingly not the mother tongue of the writers); the doctrinal and theoretical references also differ, in relation to a frequent lack of lexical and conceptual rigor, features that characterize the economic discipline of analysis of long swings. The same words, or the same expressions, are sometimes used to define different concepts and, conversely, a single concept may be defined by several expressions that the author considers to be synonymous. The use of different languages and translations obviously accentuates these disparities and tends to imbue this field of research with a "soft focus" effect. Thus, certain American economists tend to use the term "long cycle" for economic movements that are shorter than the long movement and closer to Kuznets's "secondary movements," Arthur Burns's "trend cycle," or Walter Isard's "transport building" cycle. Confusion with regard to terms increases when American economists frequently refer to the cycle revealed by Juglar as the "major cycle" and that Kondratieff himself sometimes used the same expression for his long cycles.

Questioning the opportuneness of favoring a particular expression to support a theory is comparatively rare, as most economists do not refer to the intellectual process that led them to using one expression rather than another. When a single expression is used, has it been chosen with reference to the author's theoretical options or is it the result of a desire to be understood by the whole of the scientific community that is likely to approach this discipline? When several expressions are used in the same text to describe and analyze the process of the long movement in economics, can it be concluded that the author considers them to be equivalent and that he has chosen the solution of "not sticking to the words" in order to devote himself more thoroughly to the study of the phenomenon?

Attempting to analyze the denominations used by different authors to qualify the concept of the long movement immediately brings up the problem of choice of the expression to be used throughout this paper. By definition there is no "neutral" expression that would bear neither a positive nor a negative connotation; choice is all the more delicate because there is an impressively long list of expressions. Being aware of this difficulty but wishing to overcome it, I have opted for the expression "long movement of the economy" ["mouvement long de l'économie"], which does not seem to have been the subject of much debate, perhaps because generally it is not translated literally. Such a choice is purely contingent to the argument and is not the affirmation of any theoretical option whatsoever.

I try first to show the emergence of the ambiguities to which economic analysis of long cycles leads. An attempt is then made to provide a more coherent framework for the expressions used in the study of long cycles, and the suitability of the name "cycle" for long movements of the economy is discussed.

14.2. Confusion and Dispersion of Economic Thinking With Regard to the Long Cycle

Pre-Kondriatieff analysis of long-term movements of the economy can be described as embryonic insofar as discussion was more intuitive than theoretical or statistical. In 1896, Parvus wrote in terms of "long waves with expansive tendency" and "long waves of economic depression,"[1] and in 1913 Van Gelderen referred to the two phases of the long wave as "flux" and "reflux."[2] There was more concern with describing the two phases of expansion and depression than the movement as a whole. Designations of the long movement remained evasive. In this context, Kondratieff's contribution was considerable from both the methodological and theoretical points of view.

14.2.1. The beginnings of incomprehension: The controversy concerning Kondratieff

No doubt a number of the present inaccuracies and points of incomprehension or disagreement concerning suitable terminology for describing the long movement already existed when Kondratieff was writing. It also seems obvious that the thinking – perhaps insufficiently expressed on this point – of the man who was to be posthumously consecrated as the pioneer in this field of investigation was subjected to semantic slide at the mercy of translators.

The conflicts between Kondratieff and other writers is therefore examined before discussion of the inherent difficulty in problems of translation. Finally, an attempt is made to explain the terminology chosen by Kondratieff.

The Controversy Concerning Kondratieff

We shall not reanalyze the controversies between Kondratieff and Soviet economists concerning the causes of long economic movements. Several authors such as George Garvy have discussed the theories of Oparin, Gerstein, and Trotsky.[3] We are concerned here with the role terminology played in these quarrels.

Trotsky was perhaps the most energetic critic of the expressions used by Kondratieff, refuting the latter's use of the expression "long cycles" by analogy with "short cycles." Trotsky saw in this a uniquely formal analogy insofar as the periodic repetition of short cycles is conditioned by the internal dynamics of capitalism, but

> As for those large phases of the trend of the capitalist evolution (of 50 years) for which Professor Kondratieff incautiously suggests use of the term "cycles," we must stress that their character and duration are determined not by the internal dynamics of the capitalist economy but by the external conditions which constitute the framework of capitalist evolution.[4]

Ernest Mandel took up this polemic to back Trotsky's position:

> When Trotsky correctly rejected Kondratieff's use of the term "long-term cycle" in analogy with the normal industrial cycle, it was essentially because the sudden upward turning points of the long waves cannot be explained primarily by internal economic causes. For that same reason, there can be no mechanical symmetry between the length of the industrial cycle and the length of the long wave.[5]

As Kondriateff's theoretical positions were criticized by most orthodox Soviet economists, particular attention should be paid to the support that he received from one of his contemporaries, W. E. Motyleff, on the highly controversial point of whether long movements are cyclic or not: "If there are Marxists who deny the cyclical character of the long cycles, this can only be imputed to their insufficiently profound analysis of the evolution of the capitalist society during the 19th century."[6]

Did Kondratieff's Thinking Suffer From a Perverse Effect of the Translations?

Kondratieff's thinking was very rapidly deformed by translations from the first distribution of his work. For example, one of his articles illustrates the way in which the changing of terms led to a conceptual shift.

In the spring of 1925, Kondratieff published an article that he entitled "Bol' schije cycly Konjonktury" in the collection of the Conjuncture Institute in Moscow. In 1926 this article was translated into German and entitled "Die languen Wellen der Konjunktur."[7] The term cycle used unambiguously by Kondratieff had already been replaced by the term waves. In 1935, *The Review of Economic Statistics* decided to publish a summary of this paper by Kondratieff and the editorial committee entrusted the work of translation to W.F. Stolper. It is obvious that the translation and the summary were derived from the German version and not from the original in Russian. Stolper or the editors of the journal took the liberty of entitling the translation "The Long Waves in Economic Life" [8] whereas a literal translation would have been "The Long Cycles of the Conjuncture." However, this summary-translation of one of Kondratieff's most famous papers was very widely distributed in the United States and in the world scientific community. In 1950 it was even incorporated in one of the American Economic Association publications *Readings in Business Cycle Theory.*[9]

The confusion – if such confusion exists – between the terms "cycles" and "waves" with regard to Kondratieff's article still exists today. In 1979, the Fernand Braudel Center for the Study of Economies, Historical Systems, and Civilizations published the whole of Kondratieff's article under the same title, using the chapters translated by Stolper in 1935 and translating the chapters that had been summarized at the time.[10] An Italian translation was published in a book in 1981 and entitled *Le onde lunghe della congiuntura.*[11] Within this context, Patrick Verley's greater rigor should be acknowledged since in 1981 he published several translations of extracts of Kondratieff's work, and entitled the 1925 article "Les grands cycle de la conjoncture."[12]

Inaccuracies, first in the German translation and then in the English translation, are all the more regrettable since the scientific community will, henceforth, refer in a general manner to these translations and not to the text published in Russian.

Inadequate Lexical Accuracy

In the famous article of 1943 in which he relates the controversies between Kondratieff and orthodox Russian authors, G. Garvy stressed that Kondratieff always used the expression "long cycles" and not "long waves."[13] In contrast, Pierre Dockes and Bernard Rosier affirmed that "Kondratieff himself stressed the idea of 'movement' or 'waves' and refused the term 'cycle'."[14] Ernest Mandel adopted an intermediary position, considering that the terminology used by Kondratieff evolved over the years: "Probably under the influence of criticism by Trotsky and other Russian Marxists, in 1926 Kondratieff replaced the concept of "long cycles" by that of "long waves."[15] However, Mandel was quick to add that "in the light of their contents, his 'waves' are identical to the 'cycles'."[16] This statement is clear at the least and, from Mandel's point of view, such as to reject Kondratieff's analysis.

These quotations alone express the measure of the difficulty in reaching an agreement today on the meaning of Kondratieff's work and even on the terminology that he used.

A process of fluctuation of each economic variable. Probably sensitive to the criticisms made concerning the cyclic nature of long movements, in 1925 Kondratieff attempted to determine what he meant by "wave-like" or "fluctuating" processes:

> By "wave-like" or "fluctuating" processes are meant processes of variation which are changing their direction in the course of time and are subject to repetition and reversion. Such are changes in prices, in the rate of interest, in the percentage of unemployed. These elements are subject to change in various directions. Considered as continuous, the processes of change may be represented by curves whose directions and slopes vary, exhibiting a series of recurring maxima and minima. None of the points of these curves is identical with any other, since it represents a different moment of time and a different combination of economic factors in production, distribution, etc. Hence the statement that the process is subject to reversion and repetition is not to be applied in an absolute sense, but is to be used only to distinguish this class of change from the other, which admits of no repetition or reversion.
>
> The conceptions of reversible and non-reversible processes, as well as those of statics and dynamics, belong, strictly speaking, to the domain of natural science in the narrower sense of the word, such as physics, chemistry, and biology; and their importance in those sciences is very great. But if the necessary caution is exercised in making use of the conceptions in economics, there would appear to be no obstacles to their application in this field as well; and the use of the conceptions of reversible and non-reversible processes in economics may be looked upon as an application of a general idea to a specific class of cases.[17]

These "wave-like processes" to which Kondratieff refers are often not considered by him to be cyclic since reversible processes are either regular or irregular: only the first are cyclic. Kondratieff sometimes used the expression "wave-like movements." He developed the idea in which he went from a "wave-like movement" to a "long cycle." In "The Long Wave in Economic Life" he wrote: "The long cycles appear as a wave-like movement about the average level."[18] It could be argued from such a definition that the definitions used by Kondratieff were not semantically based. Indeed, he first referred to "fluctuating" or "wave-like" as a process that was not necessarily cyclic, and then defined the long cycle in a way that came down to reusing a characteristic of fluctuating phenomena: movement around a level or a mean point.

Long waves constituting a long cycle. In 1925, Kondratieff used the expression "major cycles" to refer to the long movement that later bore his name.[19] His use of such an expression could only stimulate terminological confusion, since the Americans already used the term "major cycles" to refer to 7-to-11-year cycles. However, this label also implies a hierarchy of cycles going beyond their respective durations: whatever it is, the major cycle is the one that accounts for and/or has more importance than minor cycles. Thus, very logically, since their analyses of cyclic phenomena are different, Mitchell, on the one hand, and Kondratieff, on the other, do not place the major cycle at the same level. It clearly appears that Kondratieff preferred the expression "long cycles" to "major cycles." Nevertheless, this double denomination did not lead to any ambiguity in the interpretation of Kondratieff's thinking.

In contrast, simultaneous use of the expressions "long cycle" and "long waves" or "undulating movement" [in German "Wellenbewegungen"] is not such as to render Kondratieff's thinking in the clearest way possible. Close imbrication of waves and cycles occurs frequently in his work, particularly when he considers that "the objection to the regular cyclical character of the long waves, therefore, seems to us to be unconvincing."[20] One has the feeling that Kondratieff preferred to use the expression "long waves" when he described movement of the different economic variables taken individually or as a whole. However, he reserved the term "long cycle" for his interpretation of these movements. In fact, this distribution between the time of observation and the time of analysis often appears in Kondratieff's work.

14.2.2. The period of lexical anarchy

After the intellectual and physical elimination of Kondratieff, his work began to be very widely distributed from 1935 onward with the publication in *The Review of Economic Statistics* of a summary of one of his major articles.

Thus, from 1930 onward, two attitudes can be detected in the analysis of long movements, depending on whether the authors were acquainted with Kondratieff's work or not. After World War II, Kondratieff's articles were the subject of exegeses in the best universities. However, the terminology used by supporters of the various schools to qualify long movements in economics did not become any more accurate – quite the opposite.

The Isolation of French Economists

Kondratieff's work was not only a posthumous discovery of the American scientific community. In February 1925, he published an article in *The Quarterly Journal of Economics* describing the work of the Institute of Conjuncture in Moscow, and in August 1925 the same journal published one of his articles entitled *The Static and the Dynamic View of Economics*.[21] In 1929, Jan Tinbergen referred to Kondratieff's work in an article published in Dutch.[22] In 1933, Léon Dupriez quoted Kondratieff's 1926 article.[23]

In spite of this, in the 1930s the majority of French economists was totally unaware of the work of Kondratieff. Neither Simiand [24] nor Lescure [25] referred to Kondratieff's work or, hence, to the terminology that he used. Simiand described long movements in economy as "long phase economic fluctuations" ["fluctuations économiques à longue période"]. He was the first to call upward and downward phases "phase A" and "phase B." After the World War II, numerous authors talked in terms of phase or period A and B of the Kondratieff, thus associating two economists who knew nothing of each other's work. Lescure moved even further from the notion of "long cycle" since he did not talk in terms of long movement but only of its two components that he called the "long upswing" ["période longue d'essor"] and the "downswing period" ["période de dépression"]. This procedure was the same as that used by pre-Kondratieff economists. There are also no references to Kondratieff in Albert Aftalion [26] or in Louis Baudin; these authors used the expression "long duration movements" ["mouvements de longue durée"], which Baudin separated explicitly from cycles since "long duration price movements cannot be considered as obeying a law."[27] The same silence is observed in Bertrand Nogaro and Charles Rist, although it is true that they are much more circumspect with regard to hypotheses of long movements of the economy. No doubt some economists fully acquainted with the work of Joseph Schumpeter heard of Kondratieff's publications through the writings of the Austrian scholar. We have no hesitation in placing Francois Perroux and G. Bousquet in this category, although they were occupied with other fields of research.

In this context, which shows the relative isolation suffered by analysis of long movements in France during the inter-war period, Robert Marjolin figures as a precursor among French economists for several reasons. In 1938, in an article for an American journal, he presented the work of Francois Simiand and referred to Kondratieff.[28] In his thesis, published in 1941, Marjolin gave the first analysis of the work of Kondratieff to be published by a French economist.[29] In this work he did not use the expression "le Kondratieff" that Schumpeter had invented several years before but remained faithful to the expression "long duration movements." However, Marjolin's terminology evolved rapidly since in the same year he published an article in which he used only the term "long cycle."[30]

It was only after World War II that references to Kondratieff appeared in French economics textbooks, in particular under the influence of Henri Guitton who mentioned Kondratieff's work in his *Fluctuations Économiques* in 1951 and in his contribution to the *Traité d'Economie Politique* edited by Louis Baudin in 1953. It was Henri Guitton who wrote the article on the Kondratieff cycle in the *Dictionnaire des Sciences Economiques* edited by Jean Romeuf and published in 1956.

Schumpeter's Skill

When in the 1930s Schumpeter took up analysis of long-term movement of the economy, he obviously found himself faced with the difficulties that Kondratieff, Van Gelderen, and a number of others had faced before him in qualifying a process whose complexity turned any denomination into a simplification or a fallacy. He was thus skillful enough to use an expression that turned out to be easily accepted by the media.

> Historical knowledge of what actually happened at any time in the industrial organism, and of the way in which it happened, reveals first the existence of what is often referred to as the "Long Wave" of a period of between 54 and 60 years. Occasionally recognized and even measured before, especially by Spiethoff, it has been worked out in more detail by Kondratieff, and may therefore be called the Kondratieff Cycle.[31]

Thus, the "Kondratieff" cycle was born and Schumpeter concentrated on defining the framework within which he would use it:

> The term Kondratieff Cycle is for us but a name for a certain set of facts (a certain long-time behavior of the price level, the interest rate, employment, and so on), none of which is open to doubt. It is true that the term also implies an interpretation to the effect that this behavior of our series is amenable to interpretation on the same lines as their behavior in shorter cycles.[32]

However, it is explained below that the expression "Kondratieff cycle," or "Kondratieff" alone as is sometimes used by Schumpeter in *Business Cycles,* has been considerably misused since its invention.

Once he had decided to name the long movement in economics after one of the first to study it, Schumpeter did not worry about whether it was a long cycle, long waves, or oscillation phenomena. He frequently used the expression "long waves" but sometimes talked in terms of "cyclic process." He also used – but more rarely – the expression "long swing of the Kondratieff" [33] and makes at least one sacrifice to lyricism with "the long and gentle sweep of the Kondratieff."[34] Regarding the period from 1782 to 1842, Schumpeter declared that, "Those years . . . cover what according to our tentative schema – it is very tentative – we call a Long Cycle or a Kondratieff." He adds immediately that, "We have seen reasons to believe that this long wave was not the first of its kind."[35]

Should it be considered that in Schumpeter as in Kondratieff there is a division between the long wave that would be derived from the observed phenomenon and the long cycle that would correspond to a name resulting from the analytical procedure applied to this observation by the author? We do not think so, precisely because use of the term Kondratieff "released" Schumpeter from these preoccupations.

Whereas Schumpeter – as we have seen – was ready to use the expression "long cycle" and since his explicit reference to Kondratieff might bring him closer to the term, these theoretical options appear to us to be more appropriate to the expression "long waves" insofar as the innovation process that he describes appears to be largely exogenous in relation to overall economic movement. His very definition of the Kondratieff cycle seems to be marked with this trend. It could also be added that all the neo-Schumpeterians have adopted the term "long waves."

The "Kondratieff" After Schumpeter

The expression "Kondratieff" used incidentally by most authors since Schumpeter adopted the term appears to us to be as ambiguous as the expression "long waves." Talking in terms of just the "Kondratieff" avoids saying whether one considers that there is a cycle, wave, or turbulence during the long period. To talk of the "Kondratieff" also avoids facing the always delicate problems of translation. However, doing this means using what is perhaps the most ambiguous and hazy expression.

Some authors have profited from the indeterminant character of the term and have chosen it as a neutral expression. It is surprising to note that Gaston Imbert used this procedure in his very serious work of synthesis. Admitting that he was frequently faced with lack of precision in the terminology used since the beginning of this field of research, he noted in his introduction that "It is necessary to avoid any confusion by using terminology which is free of any ambiguity." For him,

> the term "Kondratieff" that we use to designate the 1733–1743 through 1787–1789 movement does not contain in itself any hypotheses with regard to the nature of the movement. It signifies only a long duration movement. It is a simple question of facility of terminology which leads us, concerning 1733–1743 onwards with the increase in the capitalist price trend, to talk in terms of Kondratieff movement 1, 2, etc., without making prior judgment concerning their nature and simply to indicate long movements.[37]

It appears to us that a neutral expression is difficult to personalize, particularly when this expression must be unique so as not to leave as subjacent a debate that one claims to clarify. Gaston Imbert specifies that

> In order to avoid any *a priori* judgment on the nature of this movement, we shall avoid the terms "long cycles" and "long waves" used by Schumpeter and Kondratieff. We shall use the expressions "long duration movement" ["mouvement de long durée"], "long movement" ["mouvement long"], and "Kondratieff movement" ["mouvement Kondratieff"] and we shall talk of "long upward movement" ["hausse the longue durée"], "long downward movement" ["baisse de longue durée"], "hausse longue," and "baisse longue."[38]

It could be added that it is out of place to talk in terms of "Kondratieff movement" when the latter described the phenomena that he revealed as being cycles or waves.

Indeed, the most logical reasoning would require use of the expression "the Kondratieff" to be made by authors who are descendants of Kondratieff, that is, by those who consider that the movement of investments in infrastructures – or at least productive investment in general – is the main axis around which the long movement of the economy revolves. If one does not claim as mich rigor as above (so as not to be accused of rigorism), and if one integrates the evolution of economic thinking, it can be accepted that talking in terms of "the Kondratieff" means making a reference to the theoretical opinions of the person who popularized the expression, i.e., Schumpeter. But this is not so. The expression is certainly found in the work of the neo-Schumpeterians but has also been used by monetary theorists like Léon Dupriez ("Le Kondratieff à partir de 1974") and by Marxists like Louis Fontvieille.

If one continues a little on scope that the expression "Kondratieff" should have, it seems obvious that it is a long cycle and not long waves. There are two reasons for this. Firstly, as has already been mentioned, Kondratieff preferred to call the long movement that he analyzed a "long cycle." Secondly, Schumpeter referred to the long movement as a "Kondratieff" by analogy with the "Kitchin" and the "Juglar," which nobody – particularly Schumpeter – would deny are "cycles."

Schumpeter's cleverness in naming long movements "Kondratieffs," thus providing a loophole, served as an example concerning movements lasting 20 to 25 years and revealed mainly in the United States. O'Leary and Lewis considered that

> It is convenient to have a name to distinguish a secular swing from the other cycles in economic activity. Following the precedent of naming each cycle after the person who first focused attention upon it . . . we suggest calling this the "Kuznets" cycle, hoping that Professor Kuznets will not take amiss this tribute to his pioneering work.[39]

However, the "Kuznets," which was named with reference to the personalization procedure generated by the "Kondratieff" and which is frequently compared with, if not opposed to, the latter has the same problems of interpretation, as is shown by this desire for clarification expressed by Moses Abramovitz:

> W.A. Lewis has suggested that we recognize this long record of work by referring to the 15–25 year general waves in economic changes as *Kuznets cycles*, and I propose to accept this suggestion with the understanding that the use of the term "cycles" is not meant to prejudge the question whether these waves are significantly self-generating, a matter about which Kuznets is dubious and which the present writer regards as unsettled. Following Kuznets' own practice in recent writings, it may be better to substitute "swings" or "waves" for "cycles."[40]

14.3. Long-Term Movement Terminology in Contemporary Economic Thinking

Although terminological inaccuracy concerning the study of long movements in economics forms part of the history of economic thinking, it is even more marked today insofar as the increased amount of work in the field operates using numerous theoretical procedures. A critical presentation of the terminology is followed by discussion of the theoretical validity of the term "long cycle."

14.3.1. Different expressions in different languages

We attempted to determine whether theoretical logic prevailed in the use of four expressions in the terminology specific to the analysis of long movement in economics. It appears that the expressions concerned are governed to a considerable extent by the author's language, and to date the use of one term rather than another is not for reasons of methodological or theoretical clarification.

"Long Swings": An English Language Specialty

English-speakers seem to have clearly established a distinction between "long waves" and "long swings"; "long waves" is a fairly general term for the concept that the French describe as "mouvement de longue durée," "mouvement long de l'économie," or "fluctuations de longue durée." For English-speakers, talking in terms of "long swings" means placing oneself outside the field of investigation and above all outside the hypotheses of accepting a movement with an average length of 50 years. However, the term "swing" is widely used by English-speaking writers when they wish to refer to the theories of Kuznets and Burns. Thus, an author like S. Solomou, for whom the reality of the long movement is not obvious and who considers "that Kuznets swings are a much more generalized growth pattern than previously thought,"[41] is at odds with those who use the expression "Kuznets swings" with "Kondratieff waves."

This distinction appears to be less well established for German and French economists. Thus, Thomas Kuczinski, referring to a passage in Book III of *Das Capital* in which Marx may have suggested the problematics of the long movement of the economy, translated "sehr lange Zeitraûme sich erstreckender Schwingungen" by "fluctuations extending over very long periods" and deduced that Marx was talking about "long waves."[43] However, "Schwingungen" means "oscillation" and the suitable term in English would be "long swing."[44]

Likewise, French-speaking economists now rarely use the expressions "oscillation longue" or "oscillation de longue durée," which would be literal translations of "long swing." The reason for this attitude is probably that the cycle shown by Kuznets for the United States does not seem to be established for European economies. There is quite a clear reference in fairly old work by Dupriez:

> We have intentionally used here the expression "long movements" rather than "long swing." Indeed, interpretation is more neutral and does not make an *a priori* judgment concerning the character of the movement. . . . It should be accepted that historical experience is still too short to enable a final judgment to be made . . . , if the movement really is a swing, it now concerns exactly three cycles; if it is not a swing there have been five or six known inversions. It will be agreed that this is a fragile basis for generalization, particularly without the backing of logical arguments, using certain statistical regularities.[45]

Subsequently, Dupriez used the expression "le Kondratieff" much more frequently without any further details.

When talking of "swings" to describe the long movement it should be understood that there is a central point or "pivot" at the source of the movement. What could this central point be in a process integrated in the broader framework of the secular trend and which, in essence, undergoes continuous transformation (transformations of salary relations, technological conditions, socially necessary requirements, qualification, organic composition of capital, etc.)?

A French Expression: Fluctuations Longues [Long Fluctuations]

The expression "fluctuations longues" or "fluctuations de longue durée" is often used by French-speaking authors but is rarely translated into English and even more rarely used by English-speaking authors. The very fact that it is widely used results in its appearing to have no particular connotation; it does not refer to any precise option in French economic thinking. However, etymologically the word "fluctuations" refers more to the irregularity of a movement than to its periodicity.

Some authors consider that fluctuations may contain the long cycle; this can be seen already in Kondratieff if it is allowed that "long wave" is the translation of "fluctuations longues" and also in the work of several contemporary authors such as Louis Fontvieille, who discusses long cycles in an article entitled "Fluctuations longues et rapports de production."[46]

"Long waves": An International Expression

The expression "long waves" is very dominant in the literature in English and perhaps even more so in English translations of French, Dutch, Italian, and German articles. The expression "long waves" is used extremely frequently because it is convenient. The framework of one's research is situated by the international scientific community with no apparent difficulty when one refers to "long waves"; economists who have not concentrated specifically on this discipline place others easily (perhaps too easily) with Kondratieff and Schumpeter when they hear or read this expression. More than the term "Kondratieff," the expression "long waves" also appears to be the largest common denominator for researchers of different schools and different nationalities to be able to communicate with each other in this discipline without directly encountering incomprehension or risking rejection. It can be seen that the meaning of this expression has been considerably devalued and that this semantic inaccuracy appears to indicate theoretical inaccuracies.

J. Tinbergen is one of the rare authors to have tried to define a framework for the use of the term "waves" for the description of long movements. For Tinbergen,

> The alternation of rising and declining movements can be regarded as waves, meaning that successive phases are connected by a causal mechanism or a logical interrelationship. Those who adhere to this interpretation prefer to speak of "waves" or more precisely "long waves," because the existence of two or three shorter cycles has been established, including the so-called business cycle, also called a "Juglar" after the French economist who discovered it. Those who are not yet convinced of the existence of a causal relationship between the various phases are inclined to avoid the use of the term "waves," and prefer to speak instead of "so-called long waves."[47]

This attempt at clarification does not seem satisfactory for several reasons. Firstly, Tinbergen does not adopt a definite position on the term "Kondratieff cycles," which he nevertheless uses in the title of his article. Secondly, he suggests that a distinction should be drawn between the expression "long waves," which is effectively much-used, and the term "so-called long waves," which is hardly ever found in the literature on long movements. Finally, and most important, he considers that the defenders of an interpretation based on a "causal mechanism" or logical interrelation prefer to use the expression "long waves." However, many authors plainly use the term "long waves" to distinguish "long cycles" from "Kondratieff."

For French-speakers, the term "long waves" can refer to two different concepts depending on whether it is understood as "ondes longues" or "longues vagues." The word "ondes" includes the idea of recurrence and not just that of repetition, which is not the case with the word "vagues." In French at least, the latter term describes a concept that leaves more room for the irregularity of the phenomenon.

"Long cycle": A Forbidden Expression

After the polemic concerning Kondratieff's work, the debate on whether the long movement was cyclic or not quieted somewhat, in particular because of the popularization of the expression "the Kondratieff" by Schumpeter. The controversy resumed in the 1950s. The method of someone like Henry Guitton is revealing on this point:

> The Kondratieff cycle is an intermediate being and as such is difficult to classify. Since it is called a cycle, it should be studied with the other cycles. Nevertheless, since it is said to be long it is generally discussed with growth phenomena. Is it a conjunctural event or a structural feature? Is it caused from inside (endogenous) or from outside (exogenous)? . . . It is true that one should be certain of the existence of this long cycle before answering these questions.[48]

However, these questions received less attention since, during the years of great prosperity, the problems of the long movement were not topical. In recent years, the expression "long cycle" has been used in passing by numerous authors, but there has rarely been exclusive use of the term. In addition, it appears that there is no significant link between the use of this expression and a theoretical procedure centered on a cyclic analysis of the long movement of the economy.

Ernesto Screpenti, who made several attempts at modeling long movements of the economy, is one of the few authors to make systematic and exclusive use of the expression "long cycle" ["cycli lunghi"] in Italian. Screpenti is, nevertheless, far from having a deterministic view of the long cycle, as is expressed clearly in his own words:

> My personal view is that it is neither possible nor reasonable to explain deterministically Kondratieff cycle periodicity on the ground of a unique, regular and general process. This is a phenomenon in which random shocks may play a major role. . . . The 50–60 year, quasi-regular periodicity that does appear

> to have occurred in the four Kondratieff cycles historically observed may well be the consequence of a fortunate synergy of some different phenomena.[49]

It can even be considered that such a theoretical position would be more likely to refute the long-cycle notion. Insofar as one considers that periodicity is random, it becomes difficult to accept the endogenous nature of the process. However, it appears to us that periodicity and endogeneity are necessary for it to be possible to call an economic movement a "cycle." Thus the expressions "long waves" or, better still, "long-term movements" would be better suited to Screpenti's theoretical position.

One might also express astonishment at the position of Angelo Reati who, although close to Mandel's theories, uses the expression "long cycles" while mentioning that "the defenders of this theory prefer to talk in terms of long waves to show clearly that it is not a phenomenon which recurs automatically as the word 'cycle' would suggest." Reati added to the confusion when he specified that "the expression 'long cycles' used here is therefore equivalent to long waves."

In contrast, certain authors whose theoretical options could lead to use of the term "long cycle" do not use it, preferring the equivocal expression "long waves." One sometimes has the feeling that, faced with accusations of "determinism" and "mechanistic analysis" from their detractors, certain authors prefer to censor their own work rather than deepen their terminological choices in the light of their theoretical options. Thus, not so long ago, Bernard Rosier and Pierre Dockes wrote the following:

> As for the designation of long rhythm, recognizing its cyclic character does not appear to us to be synonymous with proposing a mechanistic analysis. On the contrary, we shall see that the inversion of the long conjuncture is always problematic. We therefore retain the clearer term of *long cycles*.[51]

Since taking this position, which appeared to us to be judicious, Dockes and Rosier chose to stop using the term "cycle long" and henceforth used "rythme long," which nevertheless implies just as much as "cycle" the repetitive nature of the long movement.[52]

14.3.2. Can the term "cycle" be used to describe the long movement of the economy?

Today, the term "long cycle" is still the most criticized of the expressions used to define the long movement of the economy. It is true that the period of large-scale polemic on the theme is over, but the few adepts of the expression are still subjected to comments or even calls to order from detractors of the long cycle. Renewed activity in research on the long movement of the economy has, nevertheless, far from resulted in clarification of terminology use. We shall therefore attempt to determine the reasons for the acceptance or refusal of the expression "long cycle" before proposing a semantic and theoretical framework for the long cycle of the economy.

Detractors of the Long Cycle

Given the context of semantic imprecision that characterizes the discipline, it is understandable that the authors who keep away from the "long cycle" label do not do this in the name of similar semantic analysis and theoretical procedure.

Gaston Imbert gave an extremely restrictive definition of the term "cyclique," which led him to refuse the attribution of this term to all fluctuations in economic activity:

> The regularity of a movement in space is expressed by the term "cyclic." In this case, as in a swing movement, fluctuation runs through the same or nearby points. The movement closes in on itself and thus has the geometrical appearance of circle, the etymological source of the word "cycle." [53]

Imbert bases his argument on the fact (which remains to be demonstrated) that "the more fundamental the movement, the more real evolution closely follows the positions of equilibrium; amplitude and shifts become relatively smaller for long movements than for economic cycles," [54] and he deduced that

> within this narrow limit of long-term shifts it is not possible to talk in terms of long duration movements since even in the most perfect case of fluctuation around an equilibrium level through nearby zones, the circle does not appear to us to be a satisfactory geometrical representation. The idea of duration must be reintroduced, resulting in a spiral image.[55]

Imbert follows his argument through to the end and finishes as follows: "We are thus led to rejecting the spatial image of "cyclical" fluctuations in favor of a 'spiroidal shape'." This opinion is made all the more surprising by the fact that Imbert describes the cycle demonstrated by Juglar as an "economic cycle" throughout his book.

Meanwhile, Jean-Claude Asselain wrote:

> The cyclic representation of price movements has the disadvantage of wrongly suggesting symmetry between rise and fall phases and of ignoring the specific nature of industrial prices; for the latter, there was a dominant downward trend throughout the 19th century. . . . The regularity of alternance cannot in any case be attested with as much certitude as for the Juglar cycle. In addition, before being able to affirm the existence of the cycle in the strict sense of the term, it is necessary to demonstrate the mechanisms of cyclic return: this leads back to the problem of explaining the long movement of prices.

Reference to the so-called mechanistic or mechanical nature that could be included in the notion of "long cycle" appeared in 1975 when Paul Hanappe wondered about the possible inversion of the long movement of the economy. It should be noted that at the same time Hanappe used the term "Kondratieff" as a neutral expression; he refuted the "long cycle" in favor of "long fluctuations":

> It is strange that the Kondratieff is felt to be a movement of a mechanical or magic nature. Of all the large general economic movements, this effect appears to be the one most closely related to the expression of the liberty of men, the one whose occurrence results most directly from major confrontations between social groups, in generation choices, in victories, in the great strategic struggles whose range is neither episodic nor definitive. . . . The Kondratieff thus appears as the mark of profound coherence at generation scale and which derives from historical choices.[58]

This observation – which is in fact very general and likely to be accepted by the majority – leads to Hanappe's conclusion:

> As a result of this global nature of the Kondratieff and the non-determinist character of the underlying options, the expression "long cycle" which is used in connection with it is not very appropriate. Indeed, the word "cycle" suggest recurrence and regularities of the type expressed in an econometric model. . . . This type of "theoretical" approach is even less suited to the Kondratieff than to other economic movements, and particularly shorter conjunctural movements. The expression "fluctuations de longue période" (long fluctuations) is thus much more appropriate for all the phenomena concerned.[59]

Alfred Kleinknecht's position is original insofar as he is one of the few authors to consider explicitly as similar and equally erroneous the terms "long waves" and "long cycles." He prefers to draw up a theory of long-term change in trends in economic growth rather than attempting to build a new model of long economic cycles. Kleinknecht thus specifies:

> Conceivably it may be better in this connection to avoid the admittedly popular, but misleading, notion of "long waves" or "long cycles," and instead to speak of "periods of accelerated or decelerated growth" (Mandel). Terms synonymous to this could be Dupriez's "mouvements de fond" or Spiethoff's "Wechsellagen." [60]

Curiously, Kleinknecht refers to an expression used by Mandel to define the two phases of the long movement. However, far from rejecting the term "long waves," since he generally described the long movement in this way, Mandel very clearly contrasted the expressions "long waves" and "long cycles." In a recent work, Kleinknecht develops a more traditional critique of the "long cycle":

> No evidence has been given for the existence of Kondratieff long waves as *true cycles* . . . a concept of long cycles can only attain credibility if long cycle theorists develop convincing *endogenous* models of the long cycle – i.e., it has to be demonstrated that A periods necessarily develop into B periods, and vice versa.[61]

Robert Boyer's position is similar to Kleinknecht's even if the former appears to keep away much more clearly than the latter from the problematics of the long movement. To refute the "problématique des ondes longues," Boyer put forward an argument that has frequently been expressed since Kondratieff:

> As much as the inversion from expansion to crisis is mainly endogenous, recovery appears to be strongly dependent on the political, social and technological context, . . . which tends to confirm the strong variability of the duration of the downward phases of the Kondratieff.[62]

However, it is clear that many economists use the terms "ondes longues," "long waves," or "ouvement long de l'économie," whereas they put forward exogenous interpretations of this movement for both downward and upward reversals. This intransigence is all the more astonishing since for the period from the nineteenth-century Great Depression until today R. Boyer uses the same sequential chronology as most supporters of the long movement. His system of analysis refers to long phases of expansion followed by crisis phases during which regulation is thought not to function. Thus, one might consider that Boyer and Mistral glimpse a form of long movement of the economy when they write that "the current crisis . . . shows that an original growth mode has reached its limits; its successes since World War II have resulted mainly from a combination of accumulation conditions (of the 'Fordist' type to put it in a nutshell) and a form of overall regulation (of the 'administered type')."[63] It might be considered that they glimpsed a form of long movement of the economy. However, this is not so since they are quick to add that "the present crisis is not a transitory accident and even less the downward phase of a Kondratieff cycle foreseeable long ago."[64]

In the same work, these authors put long movement and determinism in the same class: "everything indicates that the solution to present imbalances does not result mechanically from previous economic determinism, unlike what is implied for example by conceptions in terms of Kondratieff cycles."[65] However, we feel that there is no more determinism in Kondratieff's procedure than in the very notion of economic cycle, whatever its length.

A Semantic and Theoretical Framework for the Long Economic Cycle

No movement in economics can lead to a return to its initial position for the single reason that duration is in itself a variable that is liable to modify economic relations. Even if growth is excluded, no identical reproduction can occur in economics since the time that passes modifies the variables. Thus, in the strict sense there should be no cyclic phenomena in economics. However, this is a limiting view that would deny the use in the social science and the humanities of terms that have been forged etymologically and conceptually in astronomy, physics, or biology. In fact, why should one require that an economic cycle should meet the conditions of an astronomical cycle? Use of the term "cycle" occurred by analogy and not by similitude. In addition, biological, climatic, or chemical cycles do not lead back to the initial position either. Why not accept, like H. Guitton, that although a "cycle" is etymologically a closed movement that would ensure a return to the precise point of departure, a cycle with growth would be an *open cycle*.[66] Once this has been stated, two essential factors appear to emerge for the determination of a cycle in economics: the endogenic nature of a process and the repetition of this process.

The endogenic nature of the cyclic process. We have seen that the exogenous or endogenous nature of the variables put forward to account for the trend reversals in the long movement has been for many authors the prime condition for the possible denomination "long cycle." The debate between Kondratieff and his detractors was based on this: Kondratieff talked in terms of a long cycle since he considered that savings and investment were at the root of changes in trend, and his detractors refused the expression "long cycle" since they considered that exogenous causes governed trend reversals. Ernest Mandel, as a faithful follower of Trotsky, wrote of the movements demonstrated by Kondratieff that even if the latter had replaced the expression "long cycle" by "long waves," "judging by their contents, these waves are identical to the cycles."[67] Beyond the debate on vocabulary, there is clearly a difference of opinion regarding theoretical options.

Today, for the supporters of the Marxist school of regulation, the procedure as regards the study of long movements is based on analysis of the contradictions that occur at the heart of the production process. Overaccumulation results from a process of lowering of the rate of profit itself as a result of a fall in the productivity of capital from the middle of the recession period onward. Recovery only occurs when there is a process of revalorization of the labor force and devalorization of part of the capital.

Repetition of an endogenous process. Since there is no question in economics of claiming to obtain for any phenomenon whatsoever a regularity as close to the absolute as for certain astronomical, physical, biological, or chemical phenomena, it must be decided within what limits one can talk of "regular repetition." Kondratieff himself replied to this important question:

> It has been objected that long waves lack the regularity which business cycles display. But this is wrong. If one defines "regularity" as repetition in regular time-intervals, then long waves possess this characteristic as much as the intermediate ones. A strict periodicity in social and economic phenomena does not exist at all – neither in the long nor in the intermediate waves. The length of the latter fluctuates at least between 7 and 11 years, i.e., 57 percent. The length of the long cycles fluctuates between 48 and 60 years, i.e., 25 percent only.
>
> If regularity is understood to be the similarity and simultaneity of the fluctuations of different series, then it is present to the same degree in the long as in the intermediate waves.
>
> If, finally, regularity is understood to consist in the fact that the intermediate waves are an international phenomeon, the long waves do not differ from the latter in this respect either.
>
> Consequently, there is not less regularity in the long waves than in the intermediate ones, and if we want to designate the latter as cyclical, we are bound not to deny this characterization to the former.[68]

This regularity of the cycle cannot be interpreted as determinism. There is no determinism in the long movement of the economy insofar as an upward phase is not the inexorable result of the downward phase. There is no objective need to return to the growth phase since three solutions can occur during a slow growth period: either a slide toward a "stationary economy" or a change to another form of production or a solution to internal contradictions enabling the capitalist mode of production to attain a fresh phase of expansion. Repetition of structures and economic behavior does not occur from one long cycle to another. On the contrary, since it corresponds to a phase in the evolution of the mode of capitalist production, each long cycle contains the transformations of the preceding cycle. Overaccumulation of capital and the necessary devalorization of this capital to recover the conditions for prosperity are part of repetition, as is also the revalorization of the labor force that occurs during the depression phase. Obviously, the devalorization of capital and the revalorization of labor display different, indeterminate forms during each long cycle. For P. Boccara, "the repetition makes it possible to make a more rigorous distinction between the very originality of each long cycle, without which no comparison is possible in the strict sense of the term between the series of capitalist long cycles."[69] Thus, from this viewpoint the long cycle does not signify "reproduction" but transformation of the capitalist mode of production through a nonlinear movement.

14.4. Conclusion

Kondratieff's insufficiently explained position with regard to use of words, Schumpeter's choice of "Kondratieff" as a term for a movement about which there had already been 30 years of controversy concerning its nature and its very existence, and the interpretations subsequent to the thinking of Kondratieff and Schumpeter have tended to create an extremely confused situation in which it is difficult to rely on words to communicate a concept and, beyond this, a theoretical position.

It is surprising that the progress in economic thinking that has led to the refining of the methods, theoretical procedures, and concepts used has hardly emerged in the field of the long movement of the economy. The terms used today are as vague and sometimes even more watered down than when Kondratieff went on the warpath against his Soviet colleagues in the 1920s. However, the lack of precision in the terminology concerning the notion of long-term movement is indubitably related to the small extent of knowledge of the phenomena that involve the long term to such a point that this field of economic investigation has been qualified as a "dark area."

To date, the terms used to analyze the long movement of the economy have too frequently been a source of confusion, misunderstanding, and incomprehension. Nevertheless, this diversity will become a factor in the clarification and deepening of the debate if an effort is made to enrich the statistical apparatus of quantitative history. The scientific bases of the various theoretical options will thus be strengthened, and the terminology used will be more precise. This detour by way of deepening the scientific material seems to us to be indispensable for the reconciliation of theory and terminology in the investigation of the long movement of the economy.

Notes

[1] This statement was published in Sächsischen Arbeiterzeitung and developed in Die Handelskrisis und Die Gewerkschaften, 1901.

[2] Van Gelderen F.J., 1913, Springvloed–Beschouwingen over Industrieele Ontwikkeling en Prijsbeweging, in *Die Neue Tijd*, 4, 5, 6, quoted by E. Mandel in *Le troisième âge du capitalisme*, p. 285, Vol. I.

[3] Garvy, G., 1943, Kondratieff's Theory of Long Cycles, *The Review of Economic Statistics* **25**(4):203–220.

[4] Trotsky, L., 1923, On the Curve of the Capitalistic Development, an article dated 21 April and published in the 4th Book (April–July) of *Vestnik Sotsilistitscheskoi Akadermii*, p. 9, quoted by Garvey, G., op. cit. p. 213

[5] Mandel, E., 1977, *Long Waves of Capitalist Development: The Marxist Interpretation*, Cambridge University Press, p. 28.

[6] Quoted by G. Garvy, op. cit., p. 211, note 46.

[7] Kondratieff, N.D., 1926, Die langen Wellen der Konjunktur, in Archiv für Sozialwissenschaft und soziäl politik, J.C.B. Möhr, Tübingen, pp. 573–609.

[8] Kondratieff, N.D., 1935, The Long Waves in Economic Life, *The Review of Economic Statistics* **17**(6):105–115.

[9] Readings in Business Cycle Theory, A.E.A., 1950, pp. 20–4

[10] Kondratieff, N.D., 1979, The Long Waves in Economic Life *Review* **2**(4):519–562.

[11] Kondratieff, N.D., 1981, I cicli economici maggiori, A cura di Giorgio Gattei, Biblioteca Cappelli.

[12] Verley, P., 1981, Quelques extraits d'articles de Kondratieff (traduit et présenté par) Recherche et travaux de l'Institut d'Histoire Economique et Sociale de l'Université de Paris I, Bulletin No. 10, November, pp. 23–54.

[13] Garvy, G., 1943, Kondratieff's Theory of Long Cycles, *The Review of Economic Statistics* **25**(4):203.

[14] Dockes, P. and Rosier, B., 1984, La crise contemporaine réactualise l'oeuvre de N.D. Kondratieff éliminé par Staline il y a cinquante ans, *Economie et humanisme*, No. 280, November–December p. 65.

[15] Mandel, E., 1976, Le troisieme âge du capitalisme, Union générale d'Editeurs, Collection 10–18, Paris translated from German by Bernard Keiser, p. 285, note 35.

[16] Mandel, E., op. cit., p. 285.

[17] Kondratieff, N.D., 1925, The Static and Dynamic View of Economics, *The Quarterly Journal of Economics* August:580–581.

[18] The Long Wave in Economic Life, *Readings in Business Cycle Theory*, A.E.A., 1950, pp. 30–32.

[19] This expression is found in The Static and Dynamic View of Economics, *The Quarterly Journal of Economics*, August 1925, p. 581.

[20] The Long Waves in Economic Life, op. cit., 1950, p. 41.

[21] Kondratieff, N.D., 1925, The Static and the Dynamic View of Economics, *The Quarterly Journal of Economics* August, pp. 575–583.

[22] Tinbergen, J., 1929, Het Ekonomiese Getij, de Socialistische Gids, Vol. XIV, No. 10, October, quoted by A. Kleindnecht, in *Innovation Patterns in Crisis and Prosperity; Schumpeter's Long Cycle Reconsidered*, Macmillan Press, 1987, p. 215.

[23] Dupriez, L. and Fontigny, P., 1933, L'évolution économique de la Belgique en période de baisse fondamentale des prix (1871-1896), in *Bulletin de l'Institut des Sciences Economiques*, Institut des Sciences Economiques, Université Catholique de Louvain, 4e année, No. 4, August pp. 407–451.

[24] Simiand, F., 1931, *Le salaire, l'évolution sociale et la monnaie*, Essai de théorie expérimentale du salaire, Librairie Félix Alcan, Paris, 3 vols.

[25] Lescure, J., 1933, *Hausse et baisse des prix de longue durée*, Editions Domat-Montchrestien, Paris. Lescure, J., 1932, *Des crises générales et périodiques de surproduction*, revised, corrected and updated 4th edition, Editions Domat-Montchrestien, Paris.
[26] Aftalion, A., 1932, *L'or et sa distribution mondiale*, Librairie Dalloz, Paris.
[27] Baudin, L., 1938, *Quelques observations relatives à l'influence des variations de la production de l'or sur les mouvements de longue durée des prix*, in Revue d'Economie Politique, September–October, p. 1326.
[28] Marjolin, R., 1938, Francois Simiand's Theory of Economic Progress, in *The Review of Economic Studies* **5**(3):162.
[29] Marjolin, R., 1941, Prix, monnaie et production. Essai sur les mouvements économiques de longue durée, doctoral thesis in Law, Presses Universitaires de France, Paris.
[30] Long cycles in Capital Intensity in the French Coal Mining Industry, 1840 to 1914, *The Review of Economic Statistics* **23**:165–175.
[31] Schumpeter, J., 1935, The Analysis of Economic Change, *The Review of Economic Statistics* **17**(4):13.
[32] Schumpeter, J., 1935, The Analysis of Economic Change, *The Review of Economic Statistics* **17**(4):14–15.
[33] Schumpeter, J., 1939, *Business Cycle: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, 2 vol., McGraw-Hill Book Company, New York and London, Vol. II, p. 463.
[34] Schumpeter, J., *Business Cycle*, Vol. II, p. 501.
[35] Schumpeter, J., *Business Cycle*, Vol. II, p. 252.
[36] Imbert, G., 1959, *Des mouvements de longue durée Kondratieff*, Bibliothéque de la Faculte de Droit et des Sciences Economiques d'Aix-en-Provence, La Pensée Universelle, p. 2.
[37] Imbert, G., op. cit., p. 48, footnote.
[38] Imbert, G., op. cit., p.2.
[39] O'Leary, P.J. and Lewis, W. A., 1955, Secular Swings in Production and Trade, 1870–1913 in *The Manchester School*, Vol. XXIII, May 1955, p. 148 (reprinted in *Readings in Business Cycles*, A.E.A. 1965, pp. 547).
[40] Abramovitz, M., 1961, The Nature and Significance of Kuznets Cycles, in *Economic Development and Cultural Change*, reprinted in *Readings in Business Cycles*, A.E.A., 1965, p. 520.
[41] Solomou, S., 1987, *Phases of Economic Growth 1850-1973*, Cambridge University Press, Cambridge, and Kondratieff Long Waves in Economic Growth 1859–1913, in *Review* **10**(3):507–534.
[42] Kuczinski, T., 1985, Marx and Engels on Long waves, duplicated document, 1985 p. 6.
[43] Marx, K., 1894, *Das Kapital*, Kritik der Politischen Oekonomie, Buch III, Hamburg, p. 145.
[44] The French translators of this passage (Mme Cohen-Solal and Gilbert Badia, Editions Sociales, 1976, p. 170) used the expression "oscillations s'étendant sur des périodes très longues."
[45] Dupriez, L., 1947, Des mouvements économiques généraux, Institut de Recherches Economiques et Sociales de l'Université de Louvain, Tome II, p. 5.
[46] Fontvieille, L., 1986, Fluctuations longues et rapports de production, in *Issues*, No. 25, 2éme trimestre pp. 43–94.
[47] Tinbergen, J., 1981, Kondratieff Cycles and So-called Long Waves: The Early Research, in *Futures*, August pp. 258–263.
[48] Guitton, H., 1954, Oscillation et croissance, *Economie Appliquée*, Vol. VII, Nos. 1 and 2, January–June p. 189.

[49] Screpenti, E., 1986, *Some Demographic and Social Processes*, paper presented at the Colloquy Technological and Social Factors in Long Term Fluctuation, Sienna, December, p. 164.
[50] Reati, A., 1986, Le taux de profit et la composition organique du capital dans le cycle long de l'aprés-guerre: le cas de l'industrie francaise de 1959 à 1981, in *Recherches Economiques de Louvain*, Vol. 52, No. 2, June, p. 129.
[51] Dockes, P. and Rosier, B., 1983, Rhythmes économiques, crises et changement social: Une perspective historique, *Economie Critique*, La découverte/Maspéro, p. 118.
[52] Dockes, P. and Rosier, B., 1987, Approche en termes de régulation et dialectique, Innovations-conflits, duplicated document.
[53] Imbert, G., op. cit., p. 369.
[54] Imbert, G., op. cit., p. 370.
[55] Imbert, G., op. cit., p. 370.
[56] Imbert, G., op. cit., p. 371.
[57] Asselain, J.C., 1985, *Histoire économique de la révolution industrielle à la première guerre mondiale*, Presse de la Fondation Nationale des Sciences Politiques et Economiques, Dalloz, pp. 302 and 303.
[58] Hanappe, P., 1975, Les crises contemporaines: Vivons-nous un retournement du Kondratieff?, *Metra*, Vol. XIV, No. 4, p. 709.
[59] Hanappe, P., op. cit., p. 709.
[60] Kleinknecht, A., 1981, Innovation, Accumulation and Crisis: Waves in Economic Development?, *Review*, 4, Spring, p. 706.
[61] Kleinknecht, A., 1987, Innovation Patterns in Crisis and Prosperity: Schumpeter's Long Cycle Reconsidered, Macmillan Press, p. 33.
[62] Boyer, R., 1986, *La théorie de la régulation: Une analyse critique*, La découverte, p. 84.
[63] Boyer, R. and Mistral, J., 1983, Le temps présent: la crise, (II); *Annales Economie, Société, Civilisation* 4 (July–August):772.
[64] Boyer, R. and Mistral, J., op. cit., p. 772.
[65] Boyer, R. and Mistral, J., 1983, Le temps présent: la crise (I); *Annales Economie, Société, Civilisation* 3 (May-June).
[66] Guitton, H., 1954, Oscillation et Croissance, in *Economie Appliquée* 7(1,2):192–193.
[67] Mandel, E., Le troisiéme âge du capitalisme, op. cit. , p. 285.
[68] Kondratieff, N.D., 1950, *The Long Waves in Economic Life*, p. 35.
[69] Boccara, P., 1980, L'analyse des cycles longs: au-delà de l'empirisme et de la spéculation, l'élaboration théorique, *Issues*, Nos. 7–8.

(Translated by Simon Barnard)

CHAPTER 15

Generational Factors in an Evolutionary Theory of the Long Wave

Andrew Tylecote

15.1. Introduction

If we choose to define the long wave as a cycle, of some 40 to 60 years duration, in rates of economic growth of the world economy and/or its components, it must be clear that we are not dealing with something regular or constant. We should then not seek much regularity or constancy in the elements of a long-wave theory: my theory, accordingly, is an evolutionary one, which seeks to reflect the evolution of the political, social, and economic factors affecting the world economy. One element of regularity is the fairly regular appearance of new ***technological styles*** (Perez, 1983) that have the potential to produce a quantum jump in productivity, but only once the socio-institutional framework has adjusted sufficiently; initially, this framework is mismatched with the style, tending to cause economic deceleration and a socio-institutional crisis that are likely to accentuate each other, combining in a depression crisis which in time leads to the necessary socio-institutional adjustments (***Figure 15.1***). Alongside this system, and interacting with it, is a system of various long and short feedback loops, which for a more-or-less regular wave should behave as shown in the diagram in ***Figure 15.2***.

The short feedback loops are positive and weak, tending to accentuate downswings and upswings, while the long feedback loops – of the order of 20 or 30 years in length – are negative and strong, tending, at the appropriate stage, to turn the long wave. It is not difficult to integrate the two systems, at least partially, by making technological change at least partly endogenous: we need only assume, plausibly, that the economic upswing (involving fast diffusion of the established style) hastens the appearance of the next style, which (as already

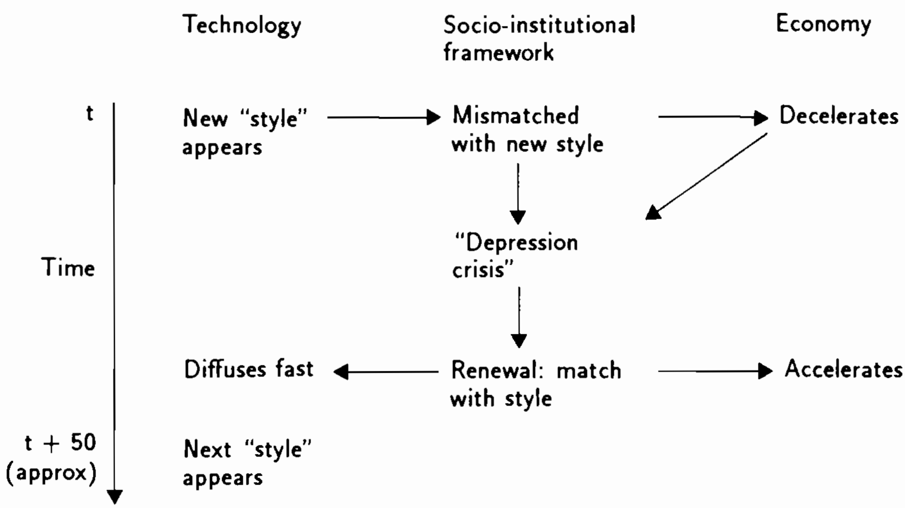


Figure 15.1. The socio-institutional framework.

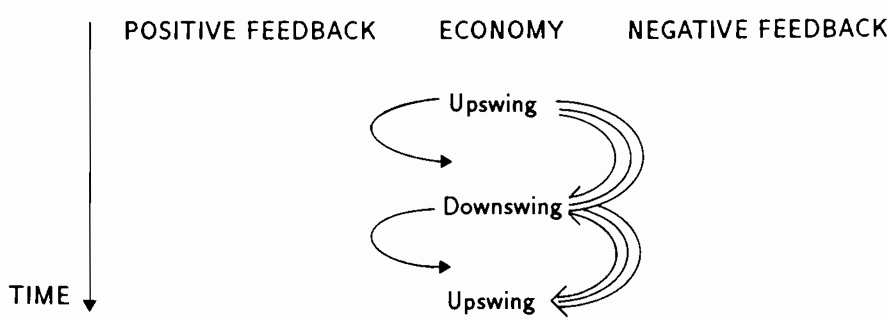


Figure 15.2. Feedback loops for a regular wave.

argued) interacts with the mismatched socio-institutional framework to help bring about a downswing and a "depression crisis." This is then a powerful negative feedback loop from upswing to downswing; a corresponding loop from downswing to upswing is provided by the socio-institutional renewal brought about by the depression crisis, for the renewal releases the potential of the new style.

Now let us spoil the neatness and regularity of our combined system. Let us suppose, for example, that the socio-institutional renewal is for some reason insufficient to release fully the potential of the new style, which then continues to diffuse slowly, without any general upswing in economic growth. [I argue that this was true of the French Revolution renewal and the "water transport style" whose appearance in the UK had provoked it (Tylecote, 1987)]. Or we may find, at a particular period, negative feedback loops that are short. Such periods tend to cause what I have called "crises of the peak." The rapid expansion of the upswing and the fast diffusion of the new style that accompanies it put their own strains on society and polity, which find expression in crises of the peak of various kinds: waves of strikes, revolutions, wars, and civil wars. When and where the previous depression crisis had led to radical reform, the crises of the peak will not be severe, and their chief importance will be in the dysfunctional changes they bring about in polity and economy – changes that take time to have an effect. They can then be regarded as part of a negative feedback loop of the right length. (See below for an example.) However, where the strains of the upswing bear on an "unreformed" structure – the United States of the 1850s, Germany and Austria-Hungary of the 1900s – the crises of the peak thus provoked are severe and have an immediate effect on the rate of growth of the world economy: thus, the US Civil War depressed the world economy during the 1860s and World War I depressed it between 1914 and about 1924. This produces what I call a "dent" in the long-wave upswing, which makes it harder to establish statistically, and may make the upper turning point impossible to locate, since (as in the case of the 1910–1929 period) there is a "double top" before and after the dent.

I am concerned here primarily with feedback loops that have a definite generational character: that is, their length is defined by the rhythms of the life of man. I believe I have found three such loops, which I call, respectively, *population feedback, Easterlin-Modigliani feedback*, and *policymakers' feedback* (or the Thatcher-Chirac effect). The first is multifaceted, and can at some periods involve slow, positive feedback (counter-cyclical) while at other times it causes pro-cyclical slow negative feedback. The other two have only been pro-cyclical in character, with slow, negative feedback.

15.2. Loop 1: Population Feedback

The very length of the long wave hints at a link with population: just two generations. Perhaps the upswing has an effect on the birthrate, which (once the babies are adults) makes for a downswing, and the downswing has an opposite effect on the birthrate, which – in turn, a generation later – makes for an upswing? I shall argue in this section that population feedback can, in certain circumstances, work in this way, to accentuate the long wave – pro-cyclically; it can also, in other circumstances, work to weaken or mitigate the wave – counter-cyclically. It depends on economic and social arrangements, and international relationships, as we shall see. Let us begin by looking at the first link in the chain of causation – the effect of a long wave upswing on population.

15.2.1. The effect of the long wave on population

Two hundred years ago it was taken for granted – by Adam Smith, for example – that an upswing in economic growth would lead to a corresponding upswing in the growth of population. Birthrates would rise, death rates would fall, and there would be an inflow rather than an outflow of migrants. But since that time we have seen, in the countries that have grown richer, that the birthrate has fallen to an extraordinarily low level – so low that even similar falls in the death rate, and at least some net immigration, have left the populations in these richer countries virtually stagnant, while the populations in poorer nations increase at an unprecedented rate. Does this not cast doubt on the relation between income and population?

Certainly, the long-term effect of economic growth on the population is a complex one. Once it has had time to affect women's status in society, and people's attitude to children, then the way is open for the perverse effect of prosperity on fertility. But such shifts are not the immediate and automatic response to a few years' fast growth; on the contrary, as we shall see later in the argument, the effect may rather be to entrench the status quo, until a crisis sets the social currents moving again. When we look at the statistics, the old simplicities turn out to be not a bad guide to population growth over the long wave. In the upswing, the death rate does tend to fall, and the birthrate to rise; in the downswing, these changes are reversed, at least relatively (that is, if the death rate does not actually rise, it falls more slowly than before).

These fluctuations are usually greatest among the poorer strata in each country, and the poorer countries in the world – naturally enough, for these are the people who live closest to the margins of subsistence.

In the twentieth century, the core countries have enjoyed too high a standard of living for their death rates to be very sensitive to the long wave. But birthrates are another matter; right across the continent and across social structure (though as I said, most particularly among the poor), the upswing brings a "baby boom," like that of the late 1940s to 1960s in the core, and the downswing brings a "baby slump" like that of the 1930s (and the 1970s and 1980s). It seems clear that the prospective parents of each generation set standards for themselves and their children according to the circumstances in which they have grown up. If, when their childbearing age arrives, it looks as though it will be difficult to reach those standards, they economize on the number of their children in one way or another (perhaps by postponing marriage). If their economic situation seems unexpectedly favorable, on the other hand, they "treat themselves" to an extra child or two. This would explain the findings of d'Souza (1985) in a study of rural Indian families that it was above all the "downwardly mobile" who restricted their families; it certainly accounts for the way the "children of the downswing" have large families in the upswing, while the children of the upswing limit the size of their families, in the downswing. As Richard Easterlin, who gave his name to these long cycles in fertility, points out, these fluctuations are the more remarkable because the children of the downswing are relatively few in number; for them to produce a baby boom the increase in births per parent must more than compensate for the drop in numbers of possible parents. It does, comfortably; the number of parents in the USA in the 1950s was around 10% under trend, while birthrates were 20% over trend (Keyfitz, 1972). This overcompensation is encouraged precisely by the small numbers in the childbearing ages:

> Not only would their promotion be relatively rapid, but in any one position, in so far as there is age-complementarity in production, they would frequently have the advantage of meeting situations in which they were too few to do the necessary work, with resulting appreciation of their services. This would often be expressed in material terms and would result in high wages and good prospects relative to what people of their age would be paid in a different age configuration, and they would have a sense of security and well-being. Their confidence is well-founded, for they will spend their whole working lives in the same advantageous position. They translate their advantage into child-bearing, perhaps projecting their security into the next generation and feeling that their children will be in demand just as they are [Keyfitz, 1972].

The unfortunate children, of course, are not in demand: not only do they have to rear *their* families in the downswing, but they have to compete with a crowd of rivals of their own age for whatever job opportunities there are.

15.2.2. The effect of the long wave on migration

So far we have discussed only the natural growth of population – that is, fertility and mortality. For the population of the world as a whole, this is all that matters; but it may be important for the world economy just how the world's population is distributed among different countries, which may be very much affected by *migration*. Every international migration is of course an emigration and an immigration (except for those unfortunates who do not survive the journey), and its causes can be grouped into those relating to circumstances in the country of emigration and those in the country of immigration; in other words, migration *push* and migration *pull*.

For our purposes the relative strength of migration push and migration pull is of great importance. In the downswing we may take it that push generally increases, but pull decreases; in the upswing it will be the other way around. Since the 1920s it has been generally known that pull has predominated: for people in poorer countries have generally been quite glad to move to richer ones, even during the upswing, but the richer countries have only been prepared to let them in, in large numbers, when they were short of labor – that is, in the (late) upswing. So during the upswing migration has risen; during the downswing it has fallen (but see below).

In terms of population growth in the richer countries, this pattern of migration has been *pro-cyclical*: it has tended to quicken growth when it was relatively quick already. However, most migrants are young adults, of whom there was, as we have seen, a shortage in the upswing and a surplus in the downswing, so the effect on the population in this age group has been *counter-cyclical*. (Whether this has had much effect on the life chances of the natives is doubtful: immigrants generally take a lower place on the ladder than most natives, so their arrival in the upswing, by allowing employers to go on expanding their businesses, creates more senior/skilled openings for native workers.)

Before the 1920s – certainly between the early 1840s and World War I – migration pull was far more important, for during that period migration was not so much from poor to rich countries, as from overpopulated to underpopulated, from the old world to the new, from Europe to the Americas. Until the War the door to the Americas stood open – restrictions on immigration were insignificant (for the healthy, in any case), and there was usually a fair chance of finding work somewhere, although clearly it was much better during times of boom (which were not necessarily the same there as in Europe). All the would-be migrant needed was the fare; the importance of the early 1840s as the beginning of this period, was the occurrence of a very sharp fall in fares, with the development of the railways and of cheap steamship travel on the North Atlantic (Thomas, 1973). In this period, then, it was above all hard times and poor prospects in Europe that drove people across the Atlantic; if there was a boom in the Americas to go to, so much the better. In fact, as Brinley Thomas has shown, their arrival tended to *create* a boom, such was the shortage of labor (Thomas, 1973).

There has been a strange echo of this period, in the last decade and a half, *within* the Americas. Although rules on immigration into the United States have been quite restrictive, they have not been effectively enforced: it has been easy to get into the United States (and stay there) not only over land from Mexico, but also by air on tourist visas (Crewdson, 1983); once there, jobs, of a sort, have been available, with no questions asked. In consequence the flow has responded to the migrants' need, to migration push, which has of course increased as the situation among potential immigrants – notably in Latin America – grew more and more desperate; by the early 1980s the inflow was running, according to some estimates, at around two million a year. Whether this immigration could be expected to have the same economic effects as that of the nineteenth century is discussed below.

15.2.3. The effect of population on the long wave

Early in history high density of population made it practicable, as Ester Boserup (1981) has shown, to develop sophisticated methods of production for the market to replace primitive subsistence forms. Slowly, high population density, and above all high *growth* of population became neutral, then detrimental. For improvements in transport made it economical to carry on market relationships over greater and greater distances, and thus keep plants of a given size busy with a lower density of population. As that factor demanding population growth faded, another, with the opposite tendency, became more important: with scarce land and capital, rising population means less for each; and by the nineteenth century, in Europe and Asia, land as well as capital was decidely scarce. On the other hand, there was abundant land and work in the Americas, so a surge in population could vent itself to some extent in migration; if there were higher birthrates and/or lower infant mortality, the upturn in migration would follow some 20 years later.

The vent of migration was not perfect – although, as I said above, the gates to the United States were open, still the uprooting would be cruel, and of course the journey there would cost more than many could afford – particularly for those who most needed to go. So a surplus of labor would be reflected in poverty and hidden unemployment in the countryside, and open unemployment in the cities, with the driving down of wages in the occupations easiest to enter. Those years or decades, then, in which the labor force grew more than usual, would be marked, in the old lands, by more social distress; but the total national output would not be depressed. On the contrary: there is always *some* use for an extra hand on a farm – land can be worked more intensively, a hillside can be terraced, a marshy patch drained. In other areas of the economy too, there would be more investment – more housing and more capital being laid out by the more prosperous fathers to set their sons up in business. And if that was so in Europe, how much more true was it in labor-hungry North America, with its abundant natural resources – with more workers, whether native or immigrant, there would assuredly be more output, and roughly in proportion.

As we have already seen, the vent of unhindered immigration to the United States was closed, rather abruptly, in the 1920s (and to the rest of the Americas virtually closed in slump) and only reopened when it was less needed, after the Depression was over. Migration push could then do nothing to mitigate the downswing of the late 1920s and 1930s, while migration pull accentuated the upswing of the 1950s and 1960s. But in any case, by the 1920s it was beginning to be doubtful whether high immigration was likely to have much effect on the American economy. The Prairies and the Pampas and the other good land available for European settlement had been settled: the frontier had now been "closed," and the amount of new labor the Americas could absorb depended – much like other countries – on the rate of expansion of their industry.

A more gradual change to be reckoned with has continued to this day, and still has a long way to go in the periphery and semi-periphery: the divorce between the factors that affect the supply of labor and those that affect the demand for it. The peasant farmer if he has another son (supply of labor) will find some use for him (demand for labor). This is particularly easy when he is producing, at least partly, for the family's own consumption – for that, with an extra mouth to feed, will presumably increase. But even when he is producing for the market, he can count on selling the extra output, since agricultural markets are competitive – you sell what you choose at a price you cannot choose, and cannot affect. For the nineteenth century artisan or entrepreneur, the situation was not dissimilar. The employee, on the other hand, cannot call work into existence for any extra children he may have – and we are, in the core, most of us employees now. Even the self-employed are now rarely producing for competitive markets on which they could unload whatever extra output a larger family enabled them to produce; they are certainly not producing for their own subsistence.

The neoclassical economist would say that, nonetheless, the extra children can call the extra work into existence for themselves by bidding down the price of labor. Along with other post-Keynesian economists, most notable Joan Robinson – and indeed Keynes himself, in Chapter 19 of the *General Theory* – I reject that supposition. There is no room to rake over the controversy now, except for one post-Keynesian point that fits well into my reasoning here: in so far as excess labor *does* drive down real wages it increases inequality, and that, as I argue elsewhere, tends nowadays to depress the economy decidedly (Tylecote, 1985 and 1986).

This argument leads us to rather complex conclusions about the effect of population increase on the long wave. Everywhere in the nineteenth century, faster growth of the labor force led to faster growth of output, and this effect occurs in the periphery even now. It was particularly strong in the 1845–1914 period, when a large part of the extra population tended to be channeled to where it was most useful, in the Americas. Since, as we have seen, the birthrate would be fastest in the upswing, the increment to the labor force would be highest in the downswing (the lag is rather less than half a long wave, but near enough, particularly if we assume that the birthrate takes time to respond to the economic situation). So this was a factor that would tend to weaken the wave, and must have had a considerable effect in this direction, in the nineteenth century. The weakening would have been more marked in the 1845–1914 period, because the downswings of the wave – such as they were – would have generated their own migration push, and thus made their own contribution to growth in the Americas. In fact the most important factor in migration push may have been neither growth of the labor force, nor slower growth of output, but a surge in growth of productivity as a new technological style diffused. The migration push generated was all the stronger when this diffusion undermined cottage industries, as it did in Western Europe in the 1840s and in Southern and Eastern Europe around the turn of the century (Thomas, 1973). We might add a fourth factor, agricultural disasters of one sort or another – the 1840s potato blight, for example – but it may be that these are not unconnected to technological diffusion. [1]

It is highly unlikely that in the post-1914 period (as well as today) the counter-cyclical effects of population feedback have been of any importance. If the total output of peasant farmers in the periphery is higher than it otherwise would have been because their labor force is larger, then it is hardly likely to have a noticeable positive effect on the world economy. Indeed, if they put more output onto world markets and depress primary product prices thereby, then it will increase international inequality and (as I shall argue below) thus tend to *depress* the world economy. Such will be the effect, also, of any increase in intra-national inequality due to an oversupply of labor, particularly unskilled (again, see below). What of the current echo of nineteenth century migration push, in immigration to the USA? Following up the argument above, it seems likely that the USA can benefit greatly nowadays from an increase in its labor force; and given the qualifications of many of the immigrants, the countries of emigration must be losing.

15.3. Loop 2: Easterlin-Modigliani Feedback

While the counter-cyclical effect of demographic factors have been declining into insignificance, an interesting pro-cyclical effect has just appeared. (Here I would acknowledge my debt to neoclassical economics. It is decent, I suppose, having rejected one neoclassical idea, to make amends by borrowing another, albeit for a use for which it was not intended.) Modigliani's life-cycle theory of saving explains differences in the proportion of income saved among different groups of the population, by pointing to the effects of age differences (thus helping to undercut the unfashionable Keynesian notion that wealth and poverty are largely responsible). Put very simply, people of working age save for their retirement, then (having retired) *dissave*. This is clearly true for a large part of the population, but it used to be very much less so. In the last century, most working people were able to save very little and, if they happened to outlive their capacity to work, fell back on the charity of their children, other relatives, or (God help them) the workhouse. Those few who owned most of the wealth did so as landowners or entrepreneurs, and as such might well continue to accumulate till their death, even if they passed the responsibility for day-to-day management to their children or others. Even the old middle class of small entrepreneurs were mostly in a similar situation. The relatively recent rise of the new middle class, the salariat, and the secure blue-collar workers just below them began to make the life-cycle theory of saving relevant. However, two even newer developments helped to make an increasingly accurate description of part of the pattern of saving in the core economies: first, the rise of the occupational pension system as a compulsory feature of middle-class (and increasingly, working-class) life; second, the spread of the practice of purchasing a house with borrowed money. Previously, the feckless – like the poor still – might save little, live in a rented house, and take their chance of a poverty-stricken old age. Now, in most countries, a large (and ever-increasing) proportion of the working population is *forced* to save: their pension contributions are deducted from their pay at source,[2] and they are usually obliged to pay back the borrowings for their houses, before retirement. Once retired, the careful might once have conserved their capital – but now they usually have little choice (or need): the pension fund will dissave for them, by paying them a regular sum until death. So saving is increasingly carried out by the working population, dissaving by the retired (but note that the new system has yet to have its full impact on that age group). Clearly the rate of net saving, of subtraction from consumption expenditure, will depend on the proportion of the population in the working-age groups, and to a lesser extent on the proportion of retired people.

Let me give a specific example. Let us suppose the long wave lasts 50 years, 25 of upswing and 25 of downswing, and that the birthrate is decidedly higher during the upswing. Let us suppose that workers begin to save at age 20, and they save at a constant rate until retirement at 65. We see at once that five years before the beginning of the downswing the number of new savers increases, as those born during the baby boom grow up, and remains at the higher level for the next 25 years, that is, through most of the downswing. Taken by itself, that would provide just the effect we are looking for – a progressive decrease in the propensity to consume, and thus a deflation of demand in the economy, more or less at the same time as an increasing labor force needs more work and therefore higher demand.

Can this entry effect be taken by itself? Should we not consider at the same time the effect of *exit* from the saving labor force? As soon as we do, we confuse the picture somewhat. The baby boom of the *previous* upswing will start to reach retirement age after 15 years of the current one, and the last will leave the labor force 15 years into the downswing. This means an opposite effect, reflationary while reducing the labor force, over a period that begins and ends only five years earlier than the entry effect. However, there is good reason for neglecting this exit effect, for the past and present. It relates to an age group 50 years older than the new entrants, which means that the total size of the age group is smaller and there is a much smaller proportion of them in the mortgaged-and-pensioned salariat. On the other hand, although we may be dealing with a smaller absolute change in numbers than we see in the new entrants, they may count double in that they do not simply cease to save – they dissave.

In fact, for the present long wave, the issue is settled by the effect of World War II on the group of the men born between 1910 and 1925; from about 1975 onward the number of men retiring has been lower, not higher, than the long-term trend. Therefore, since that time, the exit effect, whatever its strength, has pushed in the same direction as the entry effect. In the future, when populations and salariats are more stable, we can expect the exit effect to come into its own. Having seen the reflectionary entry effect of the current baby slump start to come through around 1990, we shall have to wait until 2010, unless retirement ages are reduced, for the reflationary exit effect of the last baby boom to begin.

Thus we cannot expect demographic effects on saving to be *neatly* pro-cyclical in the future: but it happens that they are highly pro-cyclical just at present. We would expect them to be more pro-cyclical in countries where:

(1) There is a particularly high ratio of working age to retired population.
(2) The tendency of the working-age population to save, and retired to dissave, is not lessened by the impact of a generous pay-as-you-go pension system and health service.
(3) The birthrate since 1970 has been particularly low, so that the number of dependent children is small: this makes it easier to keep up a high rate of saving, as well as depressing demand for housing.

My research in this area is at an early stage, but it is notable that regarding item (1) Germany and Japan lost a particularly large proportion of (then) 18 to 35 year olds between 1939 and 1945, which is now depressing the proportion of retired people. This must be deflationary for them. On point (2), Italy has a particularly generous pay-as-you-go state pension system; France has private pay-as-you-go schemes in addition to the state scheme; and Sweden, in addition to a very generous state scheme (hybrid pay-as-you-go/funded), has an exceptionally generous health system that effectively makes all conceivable health care free for the old (see Castellino, 1982; Rosa, 1982; Stahl, 1982). These pay-as-you-go elements may well reduce the deflationary Easterlin-Modigliani effects in these countries. Finally, regarding (3), the West German birthrate has been notably low since 1970 – well below the replacement rate, again deflationary.

This would lead us to expect a particular deficiency of internal demand in West Germany and Japan; this appears to have existed at least since 1980, in view of the unusually high public sector deficits in both, the huge Japanese trade surplus and the large West German trade surplus combined with average to high unemployment.

15.4. Loop 3: Policymakers' Feedback

To anyone with some knowledge of modern history, the resemblances between the last five years and the late 1920s are striking – not so much in the state of the economy, although there are resemblances, but in the attitudes of policymakers and those they represent. In the 1920s monetary discipline was practiced and interest rates were accordingly kept high; so it is in our time. Policymakers were for fiscal discipline too, and managed to keep their budgets in surplus – our governments are in deficit, but (at least outside the USA) most unwillingly and as little as possible, given the state of their economies. (In fact, given that prices are now rising, which tends to erode the value of outstanding public debt, while then they were falling, and thus increasing their value, the fiscal stances, in terms of the rate of change of real public sector debt, are much the same.) The greatest similarity is in the attitude to make money by those already rich. There was to be, in the late 1920s, no impediment to this activity. In the USA, controls on mergers were left virtually unenforced by the Justice Department under successive Republican presidents; in the UK cartels were allowed to flourish. Trade unions, on the other hand, were regarded as obnoxious interference with the market mechanism, and their increasing weakness was welcomed. Money once made, should be kept: rates of taxation on the income of both corporations and rich individuals were low. For the 1980s, this description holds almost without modification – there is the single change of detail, and that is mergers rather than cartels now flourish in the UK.

The development of such attitudes and policies early in the downswing is, I am sure, connected with the complex process I have called *inequality feedback*, (see Tylecote, 1988), and to be fully understood should be analyzed in that context. Here, however, I propose to concentrate on what might be called the *generational factors* involved. Screpanti (1986) has usefully surveyed the literature on this issue, some of which argues that a generation is formed or influenced by its environment from birth to young adulthood, while others stress the much shorter period of transition to adulthood, from the late teens and early twenties. The latter view seems more plausible: childhood is a period of relative shelter from the wider environment, in which the child is exposed within the family to values that are likely to be traditional; even in school, the attitudes picked up from peers and teachers have been formed over a long period of individual and parental experience. It is on emergence from school, into higher education and/or employment (or military service), that a period of intensive engagement with the wider environment begins, which I take to be the most formative period, so far as environmental influences are concerned. That gives us an age for formation of attitudes of perhaps 20 to 25 (note that the people we are most concerned with here are opinion-formers and policymakers whose education is likely to be prolonged; working-class attitudes are likely to be formed earlier).

The next point that concerns us is at what time do our opinion-formers and policymakers acquire the power to form opinions and make policies. This we may put (there is more agreement here) at some point in their forties. We are dealing, then, with a lag of about a generation – perhaps 25 years – between, so to speak, input and output. Not, let me stress, that our opinion-formers' and policymakers' views are set in concrete from their twenties onward. On the contrary, they evolve and change in response to the new situations and intellectual currents to which they are exposed on their way up the greasy pole. But their choices are conditioned by the basic attitudes established during this formative period.

Let me now give some examples. The British Tory, Harold Macmillan, had served in trenches in World War I alongside his working-class contempories. When he and his generation came to power in the 1940s and 1950s, it would have been deeply repugnant to them to have engaged in any conflict with these men, and their trade union leaders, which could decently have been avoided. Their deepest instinct was for social harmony. (It may be that one element in the different attitudes in US politics and business at this time was the lack of men who had had such experiences.) A later generation, that of J.F. Kennedy in the USA and Edward Heath in Britain, came to adulthood during the Great Depression or just afterward, during World War II. For them, the conclusion drawn was likely to be that the market mechanism, so much beloved by their fathers, could not be left to rule the economy without extensive government intervention whenever and wherever it seemed necessary. Thus, in their period of power, in the late 1960s and early 1970s, they resorted freely to welfare programs, fiscal reflation, income policies, etc. In this generation the right-wingers are moderate and cautious.

The latter example relates to the adolescents of the downswing, who rule in the late upswing. We now move on 10 or 20 years, to those who grew up in the upswing. Those on the Left, who have roots in the working class, or at least close connections with it, will know of the Depression from bitter childhood experience or at least from much-discussed family experience. However, those on the Right will for the most part have suffered little more than fleeting inconveniences from the Depression; by the time they were ready to take a wider view of it, it was gone. This generation's Right Wing learned from the upswing, their upswing, One Big Thing: that the capitalist system, with its market mechanism, is wonderful, and that interference with it is a nuisance. They may at first accept their elders' cautious compromise policies intellectually, but their hearts are not in them and, when – in the peak and early downswing – those policies appear to run into trouble, this new generation of formers and makers is easily persuaded that the fault was in the fact, not the form, of the compromises: away with them, and back to pure doctrine! They associate themselves with a New, or radical, Right – not conservative, but reactionary – wishing to turn the clock back to old inequalities of wealth and power (and largely, though briefly, succeeding), promising (and inevitably failing) to restore old bonds of loyalty and old structures of security: thus Thatcher, Kohl, Chirac, and those around Reagan, thus also the men of the 1920s, like Coolidge and Hoover in the USA and Montagu Norman and Baldwin in Britain, whose ambition was to get back to the prices, the industrial relations, and the capitalist freedoms of the period before 1914.

The victories of the reactionary Right of the 1920s turned to dust and ashes a few years later, and those of their modern successors, such as they are, are not likely to be much more enduring. Nonetheless, we are stuck, for years to come, with this generation of the upswing – we may winkle the New Right out of government, but how do we dislodge it from business, the mass media, the bureaucracy, the judiciary...? (And if entrenched there, it is harder to beat in politics.) Its predecessors proved slow learners during the 1930s; those in today's New Right are also slow learners now. Here is an insight into the length of the long wave: long upswings breed long years of reactionaries, who beget long depressions.

Notes

[1] The diffusion of the "steam transport style" undermined the rural economy, reducing the demand for "outworkers" and thus forcing those with little land to concentrate on the crop – the potato – which has the highest yield per hectare. This much facilitated the movement of disease.

[2] This is true only of funded pensions systems. Where pay-as-you-go is the rule, our argument is weakened.

References

Boserup, E., 1981, *Population and Technology*, Blackwell, Oxford, UK.

Castellino, O., 1982, Italy, in J.J. Rosa, ed., *The World Crisis in Social Security*, Bonnel Editions, Paris, France.

Crewdson, J., 1983, *The Tarnished Door: The New Immigrants and the Transformation of America*, Times Books, New York, NY, USA.

Easterlin, R.A., 1961, The Baby Boom in Perspective, *American Economic Review* **51**: 869–911.

Easterlin, R.A., 1971, Does Human Fertility Adjust to the Environment?, *American Economic Review* **61**: 399–407.

Keyfitz, N., 1972, Population Waves, in T.N.E. Greville, ed., *Population Dynamics*, Academic Press, New York, NY, USA.

Perez, C., 1983, Structural Change and Assimilation of New Technologies in the Economic and Social Systems, *Futures*, October: 357–375.

Rosa, J.J., ed., 1982, France, *The World Crisis in Social Security*, Bonnel Editions, Paris, France.

Screpanti, E., 1986, Some Demographic and Social Processes and the Problem of Kondratieff Cycle Periodicity, paper presented to the Siena Long Wave Workshop, 15–17 December.

d'Souza, V.G., 1985, *Economic Development, Social Structure and Population Growth*, Sage, New Delhi, India.

Stahl, I., 1982, Sweden, in J.J. Rosa, ed., *The World Crisis in Social Security*, Bonnel Editions, Paris, France.

Thomas, B., 1973, *Migration and Economic Growth*, Cambridge University Press, Cambridge, UK.

Tylecote, A.B., 1985, Inequality in the Long Wave: Trend and Cycle in Core and Periphery, *Bulletin of the European Association of Development Institutes* 1: 1–23.

Tylecote, A.B., 1986, On Inequality and the Rate of Profit in the Long Wave, *Economic and Industrial Democracy* 7: 29–44.

Tylecote, A.B., 1987, Technological Styles and Their Relation to Fluctuations in Relative Growth Rates: Britain and France 1780–1914, mimeo.

Tylecote, A.B., forthcoming, *The Long Wave and the Present Crisis of the World Economy*, Routledge and Kegan Paul, London, UK.

CHAPTER 16

Reading the Economic Events of the Last 15 Years from the Long-Wave Theory Perspective

Tommaso Sinibaldi

16.1. Introduction

In the last 10 years there has been considerable resumption of interest in the study of long-wave theory. Since around the early 1970s several scholars in different countries (West Germany, the Netherlands, the United States, and the UK), simultaneously and independently have begun to review and to learn from the ideas of Kondratieff (1926) and Schumpeter (1939), who are responsible for the first developments of this theory, and to perform their own in-depth studies on the topic.

According to the long-wave theory, the economies of industrialized countries are subject to long-term cyclical fluctuations (about half a century). Different explanations have been proposed about the causes of these fluctuations. According to Mensch (1975), who has studied and has attempted to verify Schumpeter's intuitions empirically, such fluctuations are due to cyclicity of equal duration in the technological innovation process. The fundamental technological innovations are not distributed in a uniform way in time, but tend to be concentrated in certain periods, lasting 15 to 20 years, with about half a century between them. Of course this explanation leaves the question of why this happens unsolved.

Instead, Jay W. Forrester and his group at MIT have built a model that gives an "endogenous" explanation of the long waves:

> We believe the National Model now provides a theory of how the economic long wave is generated. The process involves overbuilding of the capital sectors in which they grow beyond the output rate needed for long-term equilibrium. In the process, capital plant throughout the economy is overbuilt beyond the level justified by the marginal productivity of capital. Finally, the overexpansion is ended by the hiatus of a great depression during which excess capital plant is physically worn out and financially depreciated on the account books until the stage has been cleared for a new era of rebuilding [Forrester, 1984].

Aside from the different theories on the wave generation mechanism, the experts studying the subject have done an enormous job in the last 10 years of collecting and classifying the data relative to the last two centuries of economic history. This work has certainly reinforced the empirical evidence of the long-wave theory.

An evaluation that finds unanimous agreement is that we have entered into a warning phase of the wave since the early 1970s, a phase that should last another 10 years or so.

In this paper I will take for granted that the long-wave theory is accepted and, moving from this foundation, I will attempt to tackle two issues. In the first place, I will attempt to show how the long-wave theory allows a reading and an explanation of the economic events of the last 15 years that seem rather clear and realistic. It seems to be an important issue owing to the fact that the economic events of the last 15 years have led scholars, experts, and the public to question the most accredited economic theories and the same credibility of economic science. This explanation will be formulated primarily in terms of inflation, interest rates, and public deficits.

In the second place, I will attempt to show with summary comparisons how the economic policies followed by governments of industrialized countries have been, at the outset of this wave's waning phase, considerably different from – and substantially better than – that which occurred during the previous similar phases, and the last in particular (1920s and early 1930s).

16.2. Companies at the Beginning of the Wave's Waning Phase

In the early 1970s the market economies of countries were (without realizing it) at the end of a period of strong economic growth, lasting more than 20 years. The business and company managers had no other experience in their professional life but that of a long, uninterrupted growth in production and in markets. It is natural and understandable that they believed that the trend corresponded to the natural order of things and was therefore destined to last indefinitely. On the other hand, the critical and doubting voices of those (intellectuals, scholars) who were far away from the daily imperative of business and would have been able to give different analyses and forecasts were very few and weak. Thus investments to increase capacity continued to be made. New factories were built to turn out new automobiles, household appliances, and the materials necessary for steel works, refineries, highways, and every kind of plant and infrastructure.

Then, in only a very few years, almost all the markets rapidly reached the saturation point. The investments made in previous years were excessive and useless to a great degree and therefore destined inevitably to become losses to be canceled from the accounting books.[1]

In 1975 the industrialized world went through the first real and strong recession after more than 20 years of uninterrupted growth: for the first time production and consumption diminished in absolute value. The activity sectors that had led the long previous development entered into a stage of long-term stagnation owing to the phenomena of demand saturation. Around 1973, in almost all the industrialized and market-economy countries, the maximum production was reached of automobiles, household appliances, steel, base chemical products, etc. Energy consumption as well, which is evidently a general indicator of both production and consumption, reached a maximum around 1973–1974: in the following years it stayed around the same value or decreased slightly.

In the business world the perception of a "change in climate" was fairly rapid, though at intuitive and not rational levels. Many understood that something was "broken" in the mechanism of development without end. Already in 1977–1978 the companies stopped investing to increase capacity and sought the most painless ways to cancel from the accounting books the losses relative to wrong investments.

Obviously, a relatively simple way to obtain the result is to operate in a context with high inflation and strong barriers to entrance. In the market economies of industrialized countries there are strong barriers to entrance, as we well know, because of the oligopolistic nature, accentuated primarily in the highest capital-intensive sectors, for which of course the problems of overdimensioning of investments posed themselves most urgently.

The inflationary context, too, was given historically, not only because the companies understood its utility and pushed in that direction, nor only because it was theorized and backed by authoritive economists, but because the governments and even public opinion instinctively perceived it as a lesser evil and the most flexible reply to the historical stage that was beginning.

16.3. Governments and Public Opinion: The Growing Public Deficits

The stagnation of production and consumption that started in 1974 had growing and considerable effects on employment. Two or three years later the companies realized that the size of the market was stationary by then and turned their efforts (and still continue to do so) toward goals of increasing productivity. Therefore, unemployment could become higher and aggregate demand could drop in a tightening up process such as that which occurred after 1929. The governments instinctively resorted largely and increasingly to deficit spending.

Rereading the newspaper articles and the comments on the political economy of the 1920s, one finds an impressive analogy with the 1970s.[2] The coal industry, for example, and railway construction, which were leading elements in development for approximately half a century, had reached a stalemate position at the end of the 1920s. Similarly, in the 1970s the following industries reached a stalemate position: automobile, oil, petrochemical, household appliances, i.e., the industries that were the fulcrum of the strong development in the previous 25 years. Around 1980 we could have had another 1929 (Marchetti, 1983).

On the other hand, 30 years of critical reflection, in-depth studies, and developments of Keynes's theory had put full employment at the top of economic policy targets in all the industrialized countries. Some modest experimenting in deficit-spending policies done in the 1950s and 1960s had made the politicians, central banks, and managers in the public sector much more confident and uninhibited in using this instrument.

These experiments were not actually real deficit-spending policies. When Keynes wrote the *General Theory* he had a vivid memory of the Great Depression in which there was, in the USA, for example, a 30% drop in income in four years. Therefore, he probably thought of public deficits on the order of magnitude of those that occurred in these years, if not higher, and not the 1–2 percentage points with respect to the GDP we had in the 1950s and 1960s.

Nevertheless, consequent to these modest experiments, and above all to the economic policy debate that accompanied them, in the early 1970s the Keynesian theories and analyses became part, even though superficially, of the cultural and social context. In this context hard deflationary policies such as those effected in the early 1930s (see *Table 16.1*) were clearly out of question as was the pursuit of the mythical goal of balancing the budget.

Table 16.1. Annual variations in consumer prices in the USA and in Italy between 1925 and 1935 (in %).

Year	*USA*	*Italy*
1925	+ 2.5	+12.3
1926	+ 0.9	+ 7.8
1927	− 1.9	− 8.6
1928	− 1.4	− 7.3
1929	+ 0.0	+ 1.6
1930	− 2.3	− 3.1
1931	− 8.8	− 9.6
1932	−10.3	− 2.6
1933	− 5.2	− 5.9
1934	+ 3.3	− 5.1
1935	+ 2.4	+ 1.4

(Sources: *Historical Statistics of the United States, Colonial Times to 1970*, US Department of Commerce, 1975; *Sommario di statistiche storiche dell' Italia 1861–1975*, ISTAT, Rome, 1976.

In the early 1970s, therefore, the industrialized countries were ready for the first large and real experiment of deficit-spending policy. This policy was implemented on a large scale and with profound similarities in all the market economies of the industrialized countries. Starting from those years there is in all these countries a clear trend to the growth of deficits. This growth shows important differences from country to country; it is higher in countries such as USA, Italy, Belgium, and Sweden, and smaller in France, the UK, and mainly

West Germany. Nevertheless, the trend to a strong and general growth is evident.

Probably the future historian will give the politicians credit for having been much more far-sighted in this circumstance than the economists. The majority of the latter in the early 1970s called themselves Keynesian (though with different marks and attributes), but very few suggested or supported a large and growing recourse to deficit spending over long periods. Most preferred the role of latter-day Cassandras and predicted catastrophes linked with the growth of public deficits.

16.4. Long-Term Keynesian Policies?

All this also stems from a certain interpretation of the Keynesian theories that was and still is prevalent: i.e., that the analyses of Keynes, and therefore the instruments of economic policy he suggested (deficit spending in particular), refer substantially to the short cycle.

Obviously, Keynes could not tackle all the problems in the *General Theory*. Moreover, the ideas he advocated were already so innovative for the cultural context in which he operated that he certainly must have given himself – despite the provocative approach that pervades the whole work – the exigency of gradualism. To have the idea accepted that deficit spending in the short term could be a fundamental instrument for controlling demand, and therefore income and employment, must have already seemed to him – as it was – a very remarkable result.

An exegesis of his analysis on deficit spending as referred to the short term, therefore, is certainly not mistaken in literal terms, but in the deepest spirit of the work probably it is.

In the chapters in which Keynes lets himself make more general and "millenarian" considerations (for example, in Chapter 8, 10, and 14) it seems to be undoubtedly so:

> When involuntary unemployment exists, the marginal disutility of labor is necessarily less than the utility of the marginal product. Indeed it may be much less. For a man who has been long unemployed some measure of labour, instead of involving disutility, may have a positive utility. If this is accepted, the above reasoning shows how "wasteful" loan expenditure [3] may nevertheless enrich the community balance. Pyramid-building, earthquakes, even wars may serve to increase wealth, if the education of our statesman on the principles of classical economics stands in the way of anything better [Keynes, 1936].

It seems difficult to understand that the building of pyramids and the reconstruction after an earthquake or a war as short-term anti-cyclic "operations."

Keynes's view on the subject was likely the following: that in certain periods of time (one or two generations, he says in Chapter 24) humanity, or better still the "advanced countries," would have completed the grant of capital goods allowed by the state of technology and capable of "more than reproducing itself" (i.e., those that showed a marginal capital efficiency higher than zero).

With this an era of relative scarcity of capital would have ended and the "useless" investments would again be necessary (the pyramids of Egypt or the cathedrals of the Middle Ages) – investments, we may add (but certainly Keynes would have agreed), that have contributed so much to make the world we live in beautiful and enjoyable.

The "useless" investments, in fact, have the precious nature of losing value with abundance, and therefore they supply unlimited horizons to human activity (and to policies backing the investments and the demand):

> Ancient Egypt was doubly fortunate, and doubtless owed to this its fabled wealth, in that it possessed two activities, namely, pyramid-building as well as the search for the precious metals, the fruits of which, since they could not serve the needs of man by being consumed, did not stale with abundance. The Middle Ages built cathedrals and sang dirges. Two pyramids, two masses for the dead, are twice as good as one; but not so two railways from London to York [Keynes, 1936].

According to Keynes the alternative to the "useless" pyramids of Egypt or the Medieval cathedrals was, let us remember, war or earthquakes.

We shall not dwell further on the question of what Keynes really wanted to say: intricate and of course open to opinion. The *General Theory*, as a few other great works, is an inexhaustible source of new ideas and interpretations. What seems to be much less a matter of opinion, instead, is the fact that in the last 10 to 12 years we have had a historical verification of the practicality of deficit-spending policies over long periods of time. We can well say today that the capacity and the solidity of the financial structures were largely underestimated by the economists. Fifteen years ago, for example, when the ratio between public debt and GDP in Italy was around 35% (see *Figure 16.1*), no one would have sworn that it was possible that this ratio would exceed 100% without reaching bankruptcy of the state and financial chaos well before it did. These limits have been exceeded today, and the Italian situation is judged precarious on more than one side, but the collapse did not occur, and life goes on. In other countries, on the other hand, it seems that the historical memory of the not-so-far-off past has been lost. In the United States, after these last few years of "savage" growth of deficit, the ratio between public debt and GDP reached 35%, i.e., the same as in 1966, and much lower than that of all of the 1950s, a time in which no one worried much about the size of the public debt.

Therefore, the economies of the industrialized countries showed that they could withstand deficit-spending policies that lasted over a time period of 10 to 12 years. The question now is how long can this continue. Let us see what reply can be given with the long-wave theory.

16.5. The Long Wave of the Public Debt

Figure 16.1 suggests with good evidence that one can identify, for at least 30 years, a long wave of public debt. This wave seems, as is logical, to be counter to the long wave of the economy. Around 1973 we would have reached the peak of the economy's long wave and the public debt's lowest value. In the following

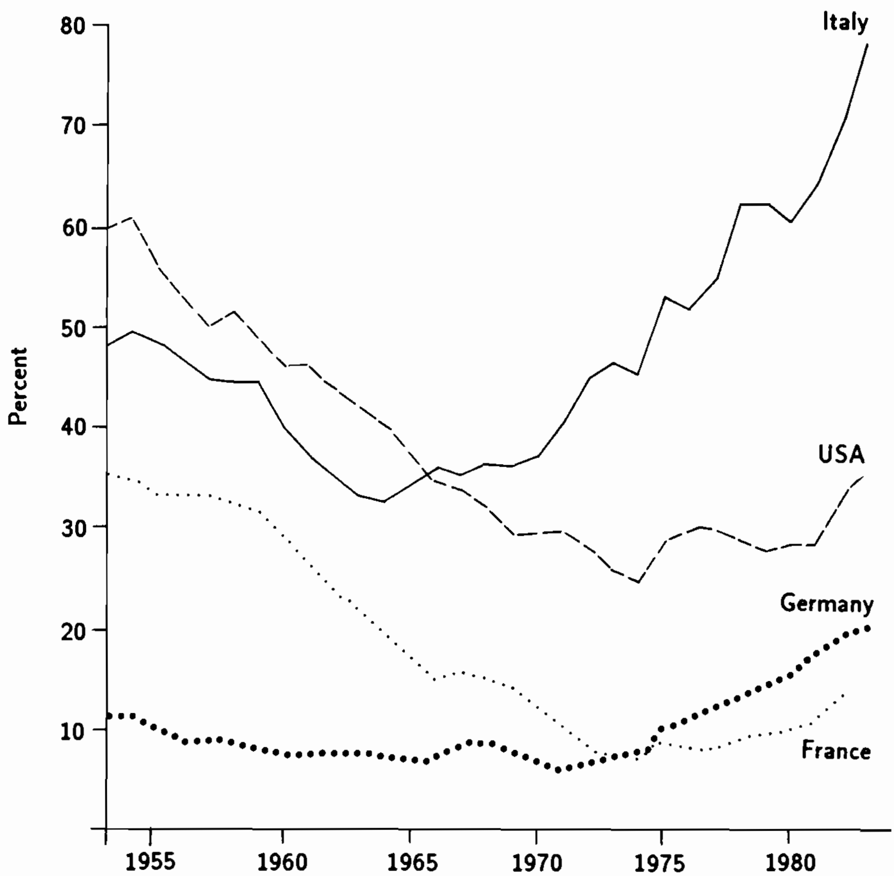


Figure 16.1. Ratio between public deficit and GDP for Italy, USA, West Germany, and France.

years the economy's wave cycle is waning and that of public debt is growing. Since the entire long wave should have a duration of approximately 50 years, an ascending or descending phase should last about 25 years. If all this is true the waning phase of the economy wave (and the increasing phase of the public debt) should last about until the end of the 1990s. Around 1985–1986, however, we should be at about the middle of this phase. If we make a further supposition that the wave has a more or less sinusoidal trend, it is in approximately these years that the low point of the public debt wave should be placed. The latter would thus continue to grow until the end of the 1990s, but starting from 1985–1986 the speed of this growth should begin to diminish, i.e., the derivative should reach its highest value. Since the derivative of the debt are the annual deficits [4], from 1985–1986 the deficits (in the form considered of their ratio with GDP) should begin to diminish slowly. This latter conclusion seems to be likely enough in line with current forecasts.

At this point, we may ask ourselves if and to what extent the economies of the industrialized countries may withstand a prospect of public debt growth (even though in decreasing rhythms) for another 10 to 12 years. This prospect seems realistically tolerable for countries like the USA in which the ratio between public debt and GDP is currently around 35%, or even more for West Germany and France, countries in which this ratio is much lower. On the other hand, it seems much riskier for countries like Italy (or Sweden or Belgium), countries in which this ratio has or is starting to go over 100%. But it is very hard to make reliable judgments on the subject: the way in which the industrialized countries have tackled this waning phase of the wave is substantially new and the limits of compatibility are not really known to anyone.

The whole subject has rather fragile empirical confirmation for the time being. However, it is a rather conclusive consequence and a logical outgrowth of the long-wave theory and of the way in which the waning phase of the wave was actually met in these years by the economies of the industrial countries. If one accepts the long-wave theory, these conclusions and forecasts may seem reasonable and well founded.

16.6. The Inflation Wave of 1973–1980

Starting from 1973 and at least until 1980, there was a strong wave of inflation in all the major industrial countries. From 1981 on this wave seems to have turned into a waning phase, but some countries (Italy, for example) are still experiencing the aftereffects.

A phenomenon of this kind – so broad and generalized – has no match for as far back as 30 years in history: one must go back to the war to find anything like it.

Rivers of ink have been spilled on the causes of this phenomenon and it does not seem that this paper is the appropriate place to discuss this problem further. What seems important to point out here, instead, is that on the basis of the supposition of overall explanation that is proposed herein, the wave of inflation of 1973–1980 was not an accident due to the two oil "shocks" or to some other historical event, but was an organic element of the economic policies that, almost "instinctively," the industrial countries put into effect to meet the waning phase of the wave. The strong inflation rise of 1973–1980, in fact, allowed the companies to put their accounts back in order, reabsorbing in a relatively painless way the effects of mistaken (excessive) investments made in the previous years. Between 1973 and 1980 almost all of the leading industrial countries had negative real interest rates (see *Table 16.2*).

The debtors (the companies) were rewarded: it was a "soft" way to discharge onto the holders of financial activities (families, but also, for example, the petro-dollar sheiks) the cost of this mistaken investments.

Of course, there may be different opinions on how "fair" all this was from the ethical viewpoint. It is hard, however, to deny that this solution was likely the most concrete and the least painful for canceling those errors. The road taken in this circumstance is, among other things, diametrically opposed to that

Table 16.2. Real interest rates[a] in some countries from 1953 to 1984 (in %).

	USA	*France*	*Germany, F.R.*	*Italy*	*UK*	*Sweden*	*Japan*
1952–1962	+2.1	+1.7	+4.4[b]	+3.6	+2.5	+1.1	–
1963–1972	+2.0	+1.8	+4.0	+2.8	+2.6	+1.7	+1.5[c]
1973–1980	−0.3	−0.8	+2.9	−3.9	−1.9	−0.3	−1.9
1981–1984	+6.5	+3.7	+4.9	+3.8	+4.4	+2.3	+4.9

[a] Real *ex post* interest rates calculated as simple algebraic difference between current interest rates of medium long-term securities and actual price variation rates. The data relative to each period are the arithmetical average of annual data.
[b] 1956–1962.
[c] 1966–1972.

Table 16.3. Real interest rates[a] in the USA and in Italy from 1925 and 1935 (in %).

Year	*USA*	*Italy*
1925	+ 2.0	− 9.1
1926	+ 3.4	− 4.6
1927	+ 6.1	+11.9
1928	+ 5.6	+10.7
1929	+ 4.4	+ 1.9
1930	+ 6.6	+ 6.3
1931	+12.9	+12.8
1932	+14.9	+ 5.5
1933	+ 9.4	+ 8.6
1934	+ 0.5	+ 7.6
1935	+ 1.0	+ 0.9

[a] Real *ex post* interest rates calculated as simple algebraic difference between current interest rates and retail price variation rates.

followed in the previous similar circumstance (according to long-wave theory), i.e., in the early 1930s. In those years, very violent deflationary policies were practiced (see *Table 16.3*), benefiting the creditors (and this did little to help economic recovery) and strangling the debtors (and this was fatal).

Prices fell violently: in the presence of very slow-moving nominal interest rates and in a context of financial markets with very little flexibility and sophistication, this meant a sudden and dizzy rise in real interest rates. In the USA they went from 4.4% in 1929 to 12.9% in 1931 and 14.9% in 1932. In Italy from 1.9% in 1929 to 12.8% in 1931.

16.7. Real Interest Rates

In contrast, currently, after approximately seven years of negative real interest rates (between 1973 and 1980), we have positive (and rather high) real interest rates.

A simple and intuitive explanation of the latter phenomenon is that there is a certain hysteresis of the nominal interest rate compared with the inflation rate.

In the ascendant phase of the inflation process there are thus – above all if this is rapid and considerable – nominal interest rates that grow with some delay compared with inflation: this delay is translated in low and also negative real interest rates, such as exactly what happened between 1973 and 1980. In the descendant phase of the inflationary wave, instead, this delay in the adjustment of the interest rate with respect to those of prices may be translated into positive and rather high real interest rates. Of course, an important objection to this rather simple reasoning is that a delay in the adjustment of the rates that by now amounts to five years or more seems excessive.

It is necessary to consider at least two elements that have a considerable role in explaining this delay. The first and obvious is that the strong and rising public deficits of the last 10 to 12 years have created strong pressure on the rates to rise. But if in the next few years a return, even though gradual, for the public deficits should begin, the pressure on rates would be attenuated quite a bit. Let us consider the case of a country (more or less the case of Italy) in which at present the public deficit is around 15% of the GDP and total domestic credit around 20% or so. The latter then is absorbed for about three-fourths to finance the public deficit and only for one-fourth to finance the economy. A reduction of one point of the deficit with respect to the GDP (from 15% to 14%), if the total domestic credit stays more or less constant in percentage terms compared with GDP, means that the "space" for the financing of the economy grows naturally by one point compared with GDP (from 5% to 6%), i.e., it increase 20%. Of course, the case of Italy is an extreme case. But in all the industrialized countries the percentage of total domestic credit absorbed by the financing of the public deficit is by now close to 50% or above, therefore we have this "multiplier" effect of available space for credit to the economy.

A second factor of delay in the lowering of rates is perhaps due to the ambiguous role that the banking system plays today in all the industrial countries. In fact, if on the one hand it has kept the traditional role of financer of the sectors with structurally negative financial deficits, then on the other it has also taken on the new role of opponent of the public sector as buyer itself or main investor of the securities it issued. Thus, on both sides it has clear interest in keeping the rates up and a high contractual power to do it.

To dwell further on this analysis would be out of place here. The few elements mentioned can give an explanation, even though brief, of why the real interest rates have been persisting at high levels for more than five years now. This does not mean that this must last much longer: the forces that are pushing in the other direction would prevail and this would happen relatively soon.

Notes

[1] This summary interpretation of the facts is very like the analysis of Forrester and his group at MIT. For details, see, for example, J.D. Sterman, An Integrated Theory of the Economic Long Wave, *Futures*, April, 1985.

[2] See, for example, J.M. Keynes, 1931, *Essays in Persuasion*, Macmillan, London, UK.

[3] In a note Keynes clarifies that "loan expenditure is a convenient expression for the net borrowings of public authorities on all accounts, whether on capital account or to meet a budgetary deficit. The one form of loan expenditure operates by increasing investment and the other by increasing the propensity to consume."

[4] Of course, this is not quite exact: the annual deficits can also be financed with instruments other than borrowing (monetary instruments).

References

Forrester, J.W., 1984, Innovation and Economic Change, in C. Freeman, ed., *Long Waves in the World Economy*, F. Pinter, London, UK.

Keynes, J.M., 1936, *The General Theory of Employment, Interest, and Money*, Macmillan, London, UK.

Kondratieff, N.D., 1926, Die langen Wellen der Konjunktur, *Archiv für Sozialwissenschaft und Sozialpolitik*, December **56**: 573–609.

Marchetti, C., 1983, Recession: Ten More Years to Go?, Paper presented at the Convention Long Waves, Depression and Innovation, Siena, 26–29 October 1983.

Mensch, G., 1975, *Das Technologische Patt*, Umschau Verlag, Frankfurt, Germany, F.R.

Schumpeter, J.A., 1939, *Business Cycles*, McGraw-Hill, New York, NY, USA.

Summary Points

Harvey Brooks

1. First of all I must confess to be both an amateur and an agnostic with respect to long waves. I am an amateur in that I am not an economist and I have not studied the vast outpouring of literature on the subject. I am an agnostic in that I do not believe the long-wave hypothesis has been unequivocally established or disproved. The contemporary debate over long waves in the "tide of human affairs" is quite reminiscent of the debates that once used to be popular in my youth on climatic cycles, a field which, 60 years later, is about as fuzzy in the 1980s as it was in the 1920s. Cyclical phenomena in inherently noisy systems are hard to establish, especially when one is uncertain as to which variables are the appropriate ones to focus on. Moreover, people are rightly skeptical of empirically established cyclical phenomena in the absence of convincing mechanisms that could plausibly produce such phenomena.
2. In my view the existence of a technological cycle or an S-curve is better established than long waves, although as I said earlier, the question of what really constitutes the unit of analysis or the appropriate ordinate for the S-curve remains controversial. The most convincing examples seem to be the long-term substitution of energy technologies and transportation modes, which the IIASA group have shown us, and it is certainly suggestive that the intervals between maximum, abstract penetration of those technologies is close to the magic 50 years of the Kondratieff waves.
3. But a fundamental issue that I think has not been very much addressed in this workshop is just how technology cycles could lead to long waves in overall economic activity, and indeed in many other social variables as we have had proposed at this meeting. I think there is indeed one possible mechanism, and I think I can illustrate and explain it with two examples that have been discussed in these sessions – that of transportation modes and of the penetration of electric motors (and the consequent demand for electricity). In the consolidation or takeoff phase of the technology cycle the performance of the technology and the efficiency of its production is

undergoing rapid improvements through many simultaneous mechanisms: through economies of scale in production, through descent of the learning curve in design and production, through learning by users (including the constant discovery of new uses and the adaptation of the technology to these uses). Even more important may be the indirect benefits in terms of convenience, reliability, public image, cleaner environment, safety, etc., which are harder to quantify but were emphasized in Vasko's presentation of the Yu and He paper. The steady accumulation of benefits produces a virtuous cycle of declining cost and expanding demand through price elasticity and through diffusion of information to potential adopters or consumers of the technology. This virtuous cycle erects almost insurmountable barriers against potentially superior technologies and designs while the market is being penetrated in accordance with the detailed mechanisms proposed by Brian Arthur and referred to by Professor Rosegger. This barrier to innovation in the same domain persists until the market begins to saturate and the socioeconomic benefits of the next increment of development or improvement of the technology in relation to costs begins to decline. At the same time many *dis-benefits* such as pollution, safety, and resource consumption problems begin to become apparent, often increasing nonlinearly with the scale of application of the technology. It is in this maturing state that the technology becomes once again susceptible to competition from a new technology in the emergent phase. This competition is at first rather uncertain; it receives some stimulus from the turbulence in prices, costs, and political attitudes that, as Nakicenovic and Fontvieille pointed out, often accompany the mature phase of the technology. It is at this point that a new technology life cycle may start, launching a new long-wave upswing.

4. Several conditions would be necessary for this scenario that I have stretched really to constitute an explanation of long waves.

    a. First the technology would affect a very large segment of the economy through its upstream and downstream linkages to other parts of the economy. Automobiles and electric motors satisfy this criteria. Each is coupled to many other simultaneous technology life cycles, such as oil refining, suburbanization, growth of national markets, that are different but tightly locked in with the auto cycle. The electricity case also satisfies this criteria – productivity in manufacturing (as we have seen), efficiency and declining cost, electricity generation, etc.
    b. There must be a lot of other technology life cycles (TLC) that proceed independently. In other words the economy must be truly dominated by one technology cycle, which may be and usually is, composed of many other technology cycles tightly linked to it. To me that is the most uncertain and least easily established part of the hypothetical picture I have described.
    c. The technology must probably also be tightly linked to many other social variables so that many social transformations go on simultaneously, which act to reinforce the technology cycle somewhat in the

manner described in Dr. Weidlich's equations. (Incidently, the negative reinforcements described by Weidlich may also be important in the mature phase.)

d. The technology cycle must have about a 50-year interval between emergent and maturity phases. This is hard to explain, but seems to be true of most of the technological trajectories explored by Nakicenovic and Marchetti. This needs to be better understood. What is it that governs the length of one TLC?

5. I think the scenario I have described is more attractive than efforts to explain long waves by clustering of innovations, although such a clustering, through their relation to underlying state of the art, described by Ayres in his IIASA Research Report, may be a factor.

The Authors

C. Antonelli
Politechnico di Milano
Milan
Italy

R. Ayres
IIASA
Laxenburg
Austria

H. Brooks
Harvard University
Cambridge, MA
USA

J.-L. Escudier
University of Montpellier
Montpellier
France

L. Fontvieille
CRPEE
Montpellier
France

A. Grübler
IIASA
Laxenburg
Austria

S.-D. He
Electric Power Research Institute
Palo Alto, CA
USA

A. Kleinknecht
Limburg University
Maastricht
The Netherlands

L. Klimenko
World Marxist Review
Prague
Czechoslovakia

T.H. Lee
IIASA
Laxenburg
Austria

S. Menshikov
World Marxist Review
Prague
Czechoslovakia

N. Nakicenovic
IIASA
Laxenburg
Austria

A. Prigent
CNRS
Champigny
France

G. Rosegger
Case Western Reserve University
Cleveland, OH
USA

T. Sinibaldi
Ente Nazionale Idrocarburi
(ENI)
Rome
Italy

I. Sytchova
Novosibirsk State University
Novosibirsk
USSR

I. Tchijov
IIASA
Laxenburg
Austria

A. Tylecote
University of Sheffield
Sheffield
UK

T. Vasko
IIASA
Laxenburg
Austria

Y.V. Yakovets
Academy of National Economy
Moscow
USSR

O.S. Yu
Electric Power Research Institute
Palo Alto, CA
USA

THE INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS

is a nongovernmental research institution, bringing together scientists from around the world to work on problems of common concern. Situated in Laxenburg, Austria, IIASA was founded in October 1972 by the academies of science and equivalent organizations of twelve countries. Its founders gave IIASA a unique position outside national, disciplinary, and institutional boundaries so that it might take the broadest possible view in pursuing its objectives:

To promote international cooperation in solving problems arising from social, economic, technological, and environmental change

To create a network of institutions in the national member organization countries and elsewhere for joint scientific research

To develop and formalize systems analysis and the sciences contributing to it, and promote the use of analytical techniques needed to evaluate and address complex problems

To inform policy advisors and decision makers about the potential application of the Institute's work to such problems

The Institute now has national member organizations in the following countries:

Austria
The Austrian Academy of Sciences

Bulgaria
The National Committee for Applied Systems Analysis and Management

Canada
The Canadian Committee for IIASA

Czechoslovakia
The Committee for IIASA of the Czechoslovak Socialist Republic

Finland
The Finnish Committee for IIASA

France
The French Association for the Development of Systems Analysis

German Democratic Republic
The Academy of Sciences of the German Democratic Republic

Federal Republic of Germany
Association for the Advancement of IIASA

Hungary
The Hungarian Committee for Applied Systems Analysis

Italy
The National Research Council

Japan
The Japan Committee for IIASA

Netherlands
The Netherlands Organization for Scientific Research, NWO

Poland
The Polish Academy of Sciences

Sweden
The Swedish Council for Planning and Coordination of Research

Union of Soviet Socialist Republics
The Academy of Sciences of the Union of Soviet Socialist Republics

United States of America
The American Academy of Arts and Sciences

Vol. 292: I. Tchijov, L. Tomaszewicz (Eds.), Input-Output Modeling. Proceedings, 1985. VI, 195 pages. 1987.

Vol. 293: D. Batten, J. Casti, B. Johansson (Eds.), Economic Evolution and Structural Adjustment. Proceedings, 1985. VI, 382 pages. 1987.

Vol. 294: J. Jahn, W. Krabs (Eds.), Recent Advances and Historical Development of Vector Optimization. VII, 405 pages. 1987.

Vol. 295: H. Meister, The Purification Problem for Constrained Games with Incomplete Information. X, 127 pages. 1987.

Vol. 296: A. Börsch-Supan, Econometric Analysis of Discrete Choice. VIII, 211 pages. 1987.

Vol. 297: V. Fedorov, H. Läuter (Eds.), Model-Oriented Data Analysis. Proceedings, 1987. VI, 239 pages. 1988.

Vol. 298: S.H. Chew, Q. Zheng, Integral Global Optimization. VII, 179 pages. 1988.

Vol. 299: K. Marti, Descent Directions and Efficient Solutions in Discretely Distributed Stochastic Programs. XIV, 178 pages. 1988.

Vol. 300: U. Derigs, Programming in Networks and Graphs. XI, 315 pages. 1988.

Vol. 301: J. Kacprzyk, M. Roubens (Eds.), Non-Conventional Preference Relations in Decision Making. VII, 155 pages. 1988.

Vol. 302: H.A. Eiselt, G. Pederzoli (Eds.), Advances in Optimization and Control. Proceedings, 1986. VIII, 372 pages. 1988.

Vol. 303: F.X. Diebold, Empirical Modeling of Exchange Rate Dynamics. VII, 143 pages. 1988.

Vol. 304: A. Kurzhanski, K. Neumann, D. Pallaschke (Eds.), Optimization, Parallel Processing and Applications. Proceedings, 1987. VI, 292 pages. 1988.

Vol. 305: G.-J.C.Th. van Schijndel, Dynamic Firm and Investor Behaviour under Progressive Personal Taxation. X, 215 pages. 1988.

Vol. 306: Ch. Klein, A Static Microeconomic Model of Pure Competition. VIII, 139 pages. 1988.

Vol. 307: T.K. Dijkstra (Ed.), On Model Uncertainty and its Statistical Implications. VII, 138 pages. 1988.

Vol. 308: J.R. Daduna, A. Wren (Eds.), Computer-Aided Transit Scheduling. VIII, 339 pages. 1988.

Vol. 309: G. Ricci, K. Velupillai (Eds.), Growth Cycles and Multisectoral Economics: the Goodwin Tradition. III, 126 pages. 1988.

Vol. 310: J. Kacprzyk, M. Fedrizzi (Eds.), Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making. IX, 399 pages. 1988.

Vol. 311: R. Färe, Fundamentals of Production Theory. IX, 163 pages. 1988.

Vol. 312: J. Krishnakumar, Estimation of Simultaneous Equation Models with Error Components Structure. X, 357 pages. 1988.

Vol. 313: W. Jammernegg, Sequential Binary Investment Decisions. VI, 156 pages. 1988.

Vol. 314: R. Tietz, W. Albers, R. Selten (Eds.), Bounded Rational Behavior in Experimental Games and Markets. VI, 368 pages. 1988.

Vol. 315: I. Orishimo, G.J.D. Hewings, P. Nijkamp (Eds.), Information Technology: Social and Spatial Perspectives. Proceedings, 1986. VI, 268 pages. 1988.

Vol. 316: R.L. Basmann, D.J. Slottje, K. Hayes, J.D. Johnson, D.J. Molina, The Generalized Fechner-Thurstone Direct Utility Function and Some of its Uses. VIII, 159 pages. 1988.

Vol. 317: L. Bianco, A. La Bella (Eds.), Freight Transport Planning and Logistics. Proceedings, 1987. X, 568 pages. 1988.

Vol. 318: T. Doup, Simplicial Algorithms on the Simplotope. VIII, 262 pages. 1988.

Vol. 319: D.T. Luc, Theory of Vector Optimization. VIII, 173 pages. 1989.

Vol. 320: D. van der Wijst, Financial Structure in Small Business. VII, 181 pages. 1989.

Vol. 321: M. Di Matteo, R.M. Goodwin, A. Vercelli (Eds.), Technological and Social Factors in Long Term Fluctuations. Proceedings. IX, 442 pages. 1989.

Vol. 322: T. Kollintzas (Ed.), The Rational Expectations Equilibrium Inventory Model. XI, 269 pages. 1989.

Vol. 323: M.B.M. de Koster, Capacity Oriented Analysis and Design of Production Systems. XII, 245 pages. 1989.

Vol. 324: I.M. Bomze, B.M. Pötscher, Game Theoretical Foundations of Evolutionary Stability. VI, 145 pages. 1989.

Vol. 325: P. Ferri, E. Greenberg, The Labor Market and Business Cycle Theories. X, 183 pages. 1989.

Vol. 326: Ch. Sauer, Alternative Theories of Output, Unemployment, and Inflation in Germany: 1960–1985. XIII, 206 pages. 1989.

Vol. 327: M. Tawada, Production Structure and International Trade. V, 132 pages. 1989.

Vol. 328: W. Güth, B. Kalkofen, Unique Solutions for Strategic Games. VII, 200 pages. 1989.

Vol. 329: G. Tillmann, Equity, Incentives, and Taxation. VI, 132 pages. 1989.

Vol. 330: P.M. Kort, Optimal Dynamic Investment Policies of a Value Maximizing Firm. VII, 185 pages. 1989.

Vol. 331: A. Lewandowski, A.P. Wierzbicki (Eds.), Aspiration Based Decision Support Systems. X, 400 pages. 1989.

Vol. 332: T.R. Gulledge, Jr., L.A. Litteral (Eds.), Cost Analysis Applications of Economics and Operations Research. Proceedings. VII, 422 pages. 1989.

Vol. 333: N. Dellaert, Production to Order. VII, 158 pages. 1989.

Vol. 334: H.-W. Lorenz, Nonlinear Dynamical Economics and Chaotic Motion. XI, 248 pages. 1989.

Vol. 335: A. G. Lockett, G. Islei (Eds.), Improving Decision Making in Organisations. Proceedings. IX, 606 pages. 1989.

Vol. 336: T. Puu, Nonlinear Economic Dynamics. VII, 119 pages. 1989.

Vol. 337: A. Lewandowski, I. Stanchev (Eds.), Methodology and Software for Interactive Decision Support. VIII, 309 pages. 1989.

Vol. 338: J.K. Ho, R.P. Sundarraj, DECOMP: an Implementation of Dantzig-Wolfe Decomposition for Linear Programming. VI, 206 pages. 1989.

Vol. 339: J. Terceiro Lomba, Estimation of Dynamic Econometric Models with Errors in Variables. VIII, 116 pages. 1990.

Vol. 340: T. Vasko, R. Ayres, L. Fontvieille (Eds.), Life Cycles and Long Waves. XIV, 293 pages. 1990.

This series reports new developments in mathematical economics, economic theory, econometrics, operations research, and mathematical systems, research and teaching – quickly, informally and at a high level. The type of material considered for publication includes:

1. Preliminary drafts of original papers and monographs
2. Lectures on a new field or presentations of a new angle in a classical field
3. Seminar work-outs
4. Reports of meetings, provided they are
   a) of exceptional interest and
   b) devoted to a single topic.

Texts which are out of print but still in demand may also be considered if they fall within these categories.

The timeliness of a manuscript is more important than its form, which may be unfinished or tentative. Thus, in some instances, proofs may be merely outlined and results presented which have been or will later be published elsewhere. If possible, a subject index should be included. Publication of Lecture Notes is intended as a service to the international scientific community, in that a commercial publisher, Springer-Verlag, can offer a wide distribution of documents which would otherwise have a restricted readership. Once published and copyrighted, they can be documented in the scientific literature.

Manuscripts

Manuscripts should be no less than 100 and preferably no more than 500 pages in length. On request, the publisher will supply special paper with the typing area outlined and essentials for the preparation of camera-ready manuscripts. Manuscripts should be sent directly to Springer-Verlag Heidelberg or Springer-Verlag New York.

Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33
Springer-Verlag, Tiergartenstraße 17, D-6900 Heidelberg 1
Springer-Verlag, 175 Fifth Avenue, New York, NY 10010/USA
Springer-Verlag, 37-3, Hongo 3-chome, Bunkyo-ku, Tokyo 113, Japan

ISBN 3-540-52473-8
ISBN 0-387-52473-8