A Long Run Theory of Effective Demand: Modelling Macroeconomic Systems with Hysteresis.

by

Mark Setterfield

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia,
July, 1992

ISBN 0-315-80148-4

Name MARK SETTERFIELD

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

ECONOMICS : THEORY | 0511 | U·M·I

SUBJECT TERM | SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
- Architecture 0729
- Art History 0377
- Cinema 0900
- Dance 0378
- Fine Arts 0357
- Information Science 0723
- Journalism 0391
- Library Science 0399
- Mass Communications 0708
- Music 0413
- Speech Communication 0459
- Theater 0465

EDUCATION
- General 0515
- Administration 0514
- Adult and Continuing 0516
- Agricultural 0517
- Art 0273
- Bilingual and Multicultural 0282
- Business 0688
- Community College 0275
- Curriculum and Instruction 0727
- Early Childhood 0518
- Elementary 0524
- Finance 0277
- Guidance and Counseling 0519
- Health 0680
- Higher 0745
- History of 0520
- Home Economics 0278
- Industrial 0521
- Language and Literature 0279
- Mathematics 0280
- Music 0522
- Philosophy of 0998
- Physical 0523
- Psychology 0525
- Reading 0535
- Religious 0527
- Sciences 0714
- Secondary 0533
- Social Sciences 0534
- Sociology of 0340
- Special 0529
- Teacher Training 0530
- Technology 0710
- Tests and Measurements 0288
- Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
- Language
  - General 0679
  - Ancient 0289
  - Linguistics 0290
  - Modern 0291
- Literature
  - General 0401
  - Classical 0294
  - Comparative 0295
  - Medieval 0297
  - Modern 0298
  - African 0316
  - American 0591
  - Asian 0305
  - Canadian (English) 0352
  - Canadian (French) 0355
  - English 0593
  - Germanic 0311
  - Latin American 0312
  - Middle Eastern 0315
  - Romance 0313
  - Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
- Philosophy 0422
- Religion
  - General 0318
  - Biblical Studies 0321
  - Clergy 0319
  - History of 0320
  - Philosophy of 0322
- Theology 0469

SOCIAL SCIENCES
- American Studies 0323
- Anthropology
  - Archaeology 0324
  - Cultural 0326
  - Physical 0327
- Business Administration
  - General 0310
  - Accounting 0272
  - Banking 0770
  - Management 0454
  - Marketing 0338
- Canadian Studies 0385
- Economics
  - General 0501
  - Agricultural 0503
  - Commerce-Business 0505
  - Finance 0508
  - History 0509
  - Labor 0510
  - Theory 0511
- Folklore 0358
- Geography 0366
- Gerontology 0351
- History
  - General 0578
  - Ancient 0579
  - Medieval 0581
  - Modern 0582
  - Black 0328
  - African 0331
  - Asia, Australia and Oceania 0332
  - Canadian 0334
  - European 0335
  - Latin American 0336
  - Middle Eastern 0333
  - United States 0337
- History of Science 0585
- Law 0398
- Political Science
  - General 0615
  - International Law and Relations 0616
  - Public Administration 0617
- Recreation 0814
- Social Work 0452
- Sociology
  - General 0626
  - Criminology and Penology 0627
  - Demography 0938
  - Ethnic and Racial Studies 0631
  - Individual and Family Studies 0628
  - Industrial and Labor Relations 0629
  - Public and Social Welfare 0630
  - Social Structure and Development 0700
  - Theory and Methods 0344
- Transportation 0709
- Urban and Regional Planning 0999
- Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
- Agriculture
  - General 0473
  - Agronomy 0285
  - Animal Culture and Nutrition 0475
  - Animal Pathology 0476
  - Food Science and Technology 0359
  - Forestry and Wildlife 0478
  - Plant Culture 0479
  - Plant Pathology 0480
  - Plant Physiology 0817
  - Range Management 0777
  - Wood Technology 0746
- Biology
  - General 0306
  - Anatomy 0287
  - Biostatistics 0308
  - Botany 0309
  - Cell 0379
  - Ecology 0329
  - Entomology 0353
  - Genetics 0369
  - Limnology 0793
  - Microbiology 0410
  - Molecular 0307
  - Neuroscience 0317
  - Oceanography 0416
  - Physiology 0433
  - Radiation 0821
  - Veterinary Science 0778
  - Zoology 0472
- Biophysics
  - General 0786
  - Medical 0760

EARTH SCIENCES
- Biogeochemistry 0425
- Geochemistry 0996
- Geodesy 0370
- Geology 0372
- Geophysics 0373
- Hydrology 0388
- Mineralogy 0411
- Paleobotany 0345
- Paleoecology 0426
- Paleontology 0418
- Paleozoology 0985
- Palynology 0427
- Physical Geography 0368
- Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
- Environmental Sciences 0768
- Health Sciences
  - General 0566
  - Audiology 0300
  - Chemotherapy 0992
  - Dentistry 0567
  - Education 0350
  - Hospital Management 0769
  - Human Development 0758
  - Immunology 0982
  - Medicine and Surgery 0564
  - Mental Health 0347
  - Nursing 0569
  - Nutrition 0570
  - Obstetrics and Gynecology 0380
  - Occupational Health and Therapy 0354
  - Ophthalmology 0381
  - Pathology 0571
  - Pharmacology 0419
  - Pharmacy 0572
  - Physical Therapy 0382
  - Public Health 0573
  - Radiology 0574
  - Recreation 0575
  - Speech Pathology 0460
  - Toxicology 0383
- Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
- Chemistry
  - General 0485
  - Agricultural 0749
  - Analytical 0486
  - Biochemistry 0487
  - Inorganic 0488
  - Nuclear 0738
  - Organic 0490
  - Pharmaceutical 0491
  - Physical 0494
  - Polymer 0495
  - Radiation 0754
- Mathematics 0405
- Physics
  - General 0605
  - Acoustics 0986
  - Astronomy and Astrophysics 0606
  - Atmospheric Science 0608
  - Atomic 0748
  - Electronics and Electricity 0607
  - Elementary Particles and High Energy 0798
  - Fluid and Plasma 0759
  - Molecular 0609
  - Nuclear 0610
  - Optics 0752
  - Radiation 0756
  - Solid State 0611
- Statistics 0463

Applied Sciences
- Applied Mechanics 0346
- Computer Science 0984
- Engineering
  - General 0537
  - Aerospace 0538
  - Agricultural 0539
  - Automotive 0540
  - Biomedical 0541
  - Chemical 0542
  - Civil 0543
  - Electronics and Electrical 0544
  - Heat and Thermodynamics 0348
  - Hydraulic 0545
  - Industrial 0546
  - Marine 0547
  - Materials Science 0794
  - Mechanical 0548
  - Metallurgy 0743
  - Mining 0551
  - Nuclear 0552
  - Packaging 0549
  - Petroleum 0765
  - Sanitary and Municipal 0554
  - System Science 0790
- Geotechnology 0428
- Operations Research 0796
- Plastics Technology 0795
- Textile Technology 0994

PSYCHOLOGY
- General 0621
- Behavioral 0384
- Clinical 0622
- Developmental 0620
- Experimental 0623
- Industrial 0624
- Personality 0625
- Physiological 0989
- Psychobiology 0349
- Psychometrics 0632
- Social 0451

Dedicated to my Mother, Father and Sister for their love and support.

Table of Contents

Table of Contents v

Illustrations and Tables vii

Abstract viii

Introduction 1

Part I: Economic Modelling: Some Methodological Considerations

Introduction 10

Chapter 1: Economic modelling and the concept of equilibrium 13

Chapter 2: An alternative approach - modelling with hysteresis 38

Part II: Modelling Macroeconomic Systems With Hysteresis: Growth and Structural Change in Capitalist Economies

Introduction 84

Chapter 3: A model of cumulative causation 87

Chapter 4: The supply side and macroeconomic performance I: technological evolution 117

Chapter 5: The supply side and macroeconomic performance II: institutional evolution 146

Part III: The Evolution of a Macroeconomy: Growth and Structural Change in the British Economy Since 1780

Introduction 185

Chapter 6: A brief history of the British economy since 1780 189

Chapter 7: Technological interrelatedness and lock in in the British economy 1870-1930 211

Chapter 8: Institutional hysteresis and the performance of the British economy after 1914 243

Conclusion. 275

Appendix A: Growth as transformation, and the existence of large scale increasing returns 284

Appendix B: Labour transfer and productivity growth - an illustrative example 288

References 293

Illustrations and Tables

Rates of diffusion of a new technique of production in two regions	133
"Selecting out and grooving in:" hysteretic "natural selection" and lock in	170
Hysteresis and the process of endogenous institutional procreation	174
Shares of world exports of manufactures for the UK, US and Germany 1899-1950	193
Rates of growth of total output for selected OECD economies 1870-1950 (percent per annum)	194
Rates of growth of output per capita for selected OECD economies 1870-1950 (percent per annum)	194
Rates of growth of output per person hour for selected OECD economies 1870-1950 (percent per annum)	195
Rates of growth of real GDP for selected OECD economies 1950-85 (percent per annum)	198
Rates of growth of real GDP per capita for selected OECD economies 1950-90 (percent per annum)	199
Rates of growth of output per person hour in manufacturing for selected OECD economies 1963-85 (percent per annum)	200
Shares of world exports of manufactures for selected OECD economies 1950-85 (percent)	201
GDP per capita in selected OECD economies 1953-84 ($US 1975)	203
Index of relative strike volume in the UK 1927-72 (1968-72 = 100)	247
Incidence, duration and size of strikes in the UK 1959-73	250

Abstract

Do high rates of economic growth create conditions favourable to their own maintenance, or can a period of high growth "sow the seeds of its own destruction?" This question is addressed by applying self reinforcing mechanisms (hysteresis, cumulative causation and lock in[1]) to the theory of growth and structural transformation.

We begin by contrasting an economic methodology based on hysteresis with the methodology underlying conventional equilibrium analysis. A model of cumulative causation is then developed, the dynamics of which are such that high growth in the present period creates conditions conducive to high rates of growth in subsequent periods. When this model is extended to allow for endogenously induced technological and institutional changes, however, it becomes clear that inefficient technological and/or institutional evolution can cause the dynamics of high growth to break down. The extended model is shown to be of a more generally hysteretic nature than its predecessor.

Finally, the theoretical insights of the thesis are applied to the historical experience of the British economy. Evidence on this experience suggests that Britain's relative economic rise and subsequent relative decline exhibit elements of the dynamics of self reinforcing mechanisms.

---

[1] See, for example, R.Cross and A. Allen (1988) "on the history of hysteresis," in R. Cross (ed.) Unemployment, Hysteresis and the Natural Rate, Oxford: Basil Blackwell on hysteresis, N. Kaldor (1972) "The irrelevance of equilibrium economics," Economic Journal, 82, 1237-55 on cumulative causation, and W.B. Arthur (1988) "Self reinforcing mechanisms in economics," in P.W. Anderson, K. Arrow and D. Pines (eds.) The Economy as an Evolving Complex System, Reading, Mass.: Addison-Wesley on lock in.

Introduction.

I Understanding the growth and development of capitalism: facts, theories and methodological approaches.

Of the many "stylized facts" about growth, two are outstanding. First, growth rates differ between capitalist economies at any given point in time. Second, growth rates within individual capitalist economies differ over time.

One consequence of these facts in terms of the historical development of capitalism has been that relative growth rates in capitalist economies have varied over time. This point has recently found expression in the form of a "new" stylized fact about growth, closely associated with the Regulation and Social Structure of Accumulation theories (see, for example, Boyer, 1990 and Bowles, Gordon and Weisskopf, 1990 respectively). This states that the growth of capitalism can be periodised into distinct "epochs," during which rapid growth and development is associated with markedly different technological and institutional structures. Furthermore, these epochs are characterised by the dominance of the world economy by different nation states, the dominant economy in each epoch being the one with the currently most successful technological and institutional structures.

The view of capitalistic growth that this "new" stylized fact inspires is that of a process intrinsically related to structural change. From the point of view of "dominant" individual nation states, it also suggests that a prolonged phase of relatively high growth may be superseded by a subsequent period of relative economic failure, as a new and more dynamic competitor gains international pre-eminence. The significance of this idea is already well known to scholars of the British economy, who have long lamented the relative economic decline of Britain since its nineteenth century industrial heyday.[1] It is also an idea that has tremendous contemporary significance, as evidenced by the increasing concern of economists with the apparent relative decline of the US economy and the emerging Japanese/East Asian challenge since the late 1960s (see, for example, Baumol, Blackman and Wolff, 1989; Bowles, Gordon and Weisskopf, 1990; Lazonick, 1990).

Concern with the long run competitive fate of nations - and in particular, with the idea that a period of sustained and relatively high growth may be followed by one of relative economic decline - is not new. The prospects of an established commercial society in an international trading environment occupied an important place in the debate within eighteenth century Scottish classical political economy.[2] One feature of the Scottish discourse was a concern as to whether accumulation could indefinitely extend the wealth of a rich

---

[1] There exists a voluminous literature relating to this concern. See, for example, Pollard (1962, 1989) and Eatwell (1984) for an introduction to the debates.

[2] See Hont (1983) for an overview of this debate.

nation (subject to its securing sufficient markets for final output), or whether the acquisition of wealth was an inherently self defeating process, by virtue of its adverse effects on the industry and ingenuity of an economy's populace.

The issue of the long run competitive fate of nations is therefore both timely and timeless. This thesis addresses the question of whether the accumulation of wealth tends to favour or hinder its subsequent accumulation, with the aim of relating the answer to the growth "epochs" described above. More specifically, it investigates the growth dynamics of capitalist economies, illustrating under what circumstances high growth may be self perpetuating, and how high growth may ultimately bequeath the conditions for its own future deceleration, and even give rise to a subsequent period of relative economic decline.

The most important methodological feature of this exercise is its conception of capitalism as an endogenously evolving system. Following Cornwall (1977), growth is interpreted as a process of structural transformation - of changes in the sectoral composition of final output, and of changes in the technological and institutional structures of an economy. Furthermore, structural transformation is not seen as being a smooth and continuous process. The transition between different technological, institutional and output structures may not always be easy for an individual economy to undergo.

This conception of the dynamics of capitalism involves a rejection of steady state approaches to modelling growth. In keeping with the "stylized facts" outlined above, the historical growth and development of capitalism is interpreted as a history of *change*, involving variations in national growth rates over time, and more fundamental changes in international economic "leadership." Adopting a steady state framework of analysis necessarily precludes all discussion of change, unless we entertain the notion of exogenous shocks. Since shocks constitute an unexplained and non-systematic element in a steady state framework, however, they cannot, by definition, be construed as a systematic explanation of change. Yet even the most cursory reading of the historical growth record of capitalism suggests that the systematic explanation of change is precisely what is required.

To this end, the steady state approach is eschewed in favour of a framework of analysis which relies upon "self reinforcing mechanisms" in order to explain macroeconomic evolution. The common methodological feature of these concepts is that they rest on the analysis of uni-directional, sequential patterns of economic activity; hence final economic outcomes are seen as being inherently dependent on the historical conditions that preceded them. This methodology contrasts markedly with that of standard equilibrium analysis (of which steady state analysis forms a part), which conceives long run economic outcomes as equilibria, defined and reached without reference to the path taken towards them. These equilibrium outcomes are ultimately "explained" by data, imposed upon the economic system from without. As suggested above, this framework of analysis cannot truly explain change. In contrast, self reinforcing mechanisms may assist in this task. These mechanisms posit that economies are influenced not just by exogenous data, but also by the nature of endogenous adjustment processes which may, in and of themselves, affect final economic outcomes. By explicitly analysing the economy as an intertemporal sequence of events, self reinforcing mechanisms provide an analytical framework for understanding the nature and significance of change.

II Outline of the thesis.

This thesis is divided into three parts, which broadly pursue methodological, theoretical and empirical themes respectively.

In Part I, the traditional equilibrium approach to economic analysis, which conceives final economic outcomes as being defined and reached without reference to the path taken towards them, is contrasted with a methodology inspired by the concept of hysteresis, in which long run outcomes are held to be path dependent. Chapter 1 suggests that although there is no unique definition of the concept of "economic equilibrium," equilibrium theories display sufficient methodological similarities for us to be able to identify a "traditional equilibrium approach" to economic analysis. This approach is critiqued in terms of its lack of realism, frequent neglect of multiple equilibria, inattention to the problems of disequilibrium adjustment, and finally, in terms of the concept of hysteresis.

In chapter 2, hysteresis is formally defined as existing when the long run outcomes of an economic system depend on the (historical) path taken towards them. It is shown that recent attempts to characterise this process mathematically have relegated hysteresis to the status of a special case. However, these characterisations are themselves rejected as a special case, and hysteresis is "reconstructed" as a potentially pervasive feature of economic systems. Hysteresis is shown to be compatible with other self reinforcing mechanisms such as cumulative causation and lock in, and complementary to the pedagogical use of the Keynesian concept of short run under employment equilibrium. This suggests the possibility of analysing the evolution of capitalist economies in terms of a long run theory of effective demand.

The purpose of Part II is to attempt to construct this theory, and illustrate its potential implications for the long run competitive fate of an open economy in an interregional trading system. Chapter 3 takes the Kaldorian heritage of non-equilibrium growth modelling as a starting point. The hysteretic potential of dynamic increasing returns to scale is discussed, and a model of cumulative causation is then constructed to illustrate the potential implications of dynamic increasing returns for growth.

Chapter 4 begins by identifying an important potential caveat in the process of cumulative causation - its apparent implication that given only initial conditions (and abstracting from the influence of exogenous shocks) the long run competitive fate of a region is known with certainty. Success (i.e., high growth) breeds success and failure (i.e., low growth) begets failure indefinitely. It is then shown that this caveat can be overcome in a manner consistent with the concept of hysteresis by considering the technological and institutional context in which growth occurs. Hence the technological (and in chapter 5, institutional) evolution of a growing economy are themselves a function of the economy's growth and development - and these processes of technological and institutional evolution need not be efficient, in the sense of being functional to the maintenance of high growth dynamics. Particularly useful for demonstrating this is the concept of lock in (Arthur, 1988), which describes the propensity of dynamic economic systems to become stuck in "grooves" from which they subsequently find it difficult to deviate. Taking the example of a high growth "virtuous circle" economy, the potential for an endogenously induced breakdown in the growth dynamics of cumulative causation is illustrated, with the possible future consequence that an initially fast growing economy may enter a period of relative economic decline. In terms of the debate over the accumulation of wealth and its perpetuation, then, it is shown that there exist mechanisms in a capitalist economy which may act so as to sustain initially high growth (such as dynamic increasing returns), and mechanisms which may cause high growth dynamics to break down (such as inefficient technological and/or institutional evolution). Ultimately, it is the historical balance of these forces which determines the long run competitive fate of a dominant (i.e., relatively fast growing and internationally pre-eminent) economy.

The purpose of Part III is to illustrate this by means of reference to the evolution of the British economy. Britain is identified as an initially fast growing economy which failed to undergo the type of structural transformation (primarily of a technological and institutional nature) necessary to maintain, in the face of emerging international competition, its position of relative economic dominance. Chapter 6 begins with a brief outline of the history of export, output and productivity growth in Britain since 1780, designed to illustrate that these variables have displayed the cumulative patterns of behaviour hypothesised in Part II. Chapters 7 and 8 then concentrate on explaining Britain's relative economic retardation after 1870, which followed its unprecedented growth as the world's "first industrial nation." In chapter 7, the influence of lock in on Britain's technological evolution is examined. The latter is found to have been inefficient, with negative consequences for the dynamics of key industries in the British economy. Chapter 8 then repeats this exercise, this time focusing on Britain's institutional evolution. This is again found to have been inefficient, and the adverse consequences for industries central to Britain's relative economic performance are again highlighted.

Finally, the conclusion offers a summary of the main findings of the thesis. It is suggested that capitalism can be interpreted as an endogenously evolving system, comprised of mechanisms whose very operation can influence the long run configuration of the system as a whole.

Part I.

Economic Modelling: Some Methodological Considerations.

Introduction.

This section assesses two different approaches to modelling the macroeconomic performance of an economy.[1] The first, identified as the traditional equilibrium approach, is based on a conception of long run economic outcomes as being determinate (Kaldor, 1933). This approach is common to most macroeconomic theorising, regardless of whether it is static or dynamic, or whether it is of a predominantly neoclassical or Keynesian nature. However, the use of the orthodox notion of equilibrium as an organizing concept does vary between neoclassical and Keynesian schools of thought. As will be seen, this has potentially important implications for macroeconomic modelling.

The second approach, which is identified with the increasingly popular concept of hysteresis, may be seen to have developed initially as a critique of the traditional equilibrium approach. Hence most authors in this relatively

[1]. The term "macroeconomic performance" is used here in general reference to indicators such as the level, growth and distribution of national income, the level of unemployment, the rate of inflation and so on. It may be noted, however, that subsequently, we shall be particularly interested in the rates of growth of exports, output and productivity when discussing macroeconomic performance.

new tradition have concentrated on demonstrating the implications of hysteresis for traditional equilibrium models such as the natural rate hypothesis. Furthermore, their methodology has not, as yet, progressed far beyond the confines of the traditional equilibrium approach.[2]

What arises from this seemingly narrow critique, however, is nothing less than a distinct approach to macroeconomic modelling - one which has logical antecedents in concepts such as cumulative causation, and which provides a dynamic alternative to the traditional equilibrium approach that is broadly consistent with the notion of real historical time. Allowing for historical time takes account of the fact that economic events occur in a uni-directional sequence. In historical time, any event occurring in the present exists in the context of a series of prior events corresponding to the periods which make up the past. In other words, what is current is so in the context of what has gone before.

As both Robinson (1980, p.255) and Shackle (1958, p.13) note, traditional equilibrium economics (both static and dynamic) treats time, if at all, as a mathematical space rather than as the historical phenomenon described above. It is because of this that traditional equilibrium models may be criticised for being ahistorical. As a consequence, traditional equilibrium models tend to obscure precise

---

[2]. This is made clear in the discussion of hysteresis and persistence in chapter 2.

(historical) sequences of events, and in so doing, deflect attention from the important possibility that such sequences might actually matter, in the sense that they may influence final economic outcomes. In contrast, hysteresis is complementary to the notion of real historical time. Indeed, it is precisely the principle that it is sequences of events that determine the nature of final outcomes on which the concept of hysteresis is based.

Part I is organised as follows. In chapter 1, the traditional equilibrium approach is discussed, and a critique of this approach is seen to give rise to the concept of hysteresis. In chapter 2, this concept is investigated and developed, and is shown to lead to an alternative approach to macroeconomic modelling which is related to the earlier methodologies of authors such as Kaldor (1972, 1985.) Furthermore, an assessment of what are identified as different equilibrium organizing concepts within the traditional equilibrium approach is shown to offer some suggestions as to what tools may be appropriate for the construction of a fully hysteretic model of a macroeconomic system.

Chapter 1.

Economic modelling and the concept of equilibrium.

I The traditional equilibrium approach to economic modelling.

There exist a number of definitions of the concept of economic equilibrium. One elementary idea of equilibrium is that of a situation characterised by "offsetting forces" - as, for example, when supply equals demand. A somewhat broader conception defines equilibrium as any state of rest which displays no endogenous tendencies to change over time.[1] There also exist a variety of model specific definitions of equilibrium. For example, in equilibrium business cycle models, equilibrium is characterised in terms of the way that individuals form conditional expectations.[2]

In spite of these differences in definition, however, it is possible to identify common methodological traits underlying what may be broadly construed as equilibrium analysis in economics. This "traditional equilibrium approach" to economic modelling has two outstanding features. First, it involves identifying an array of structural equations, which define relationships determining the values of a set of

---

[1] This encompasses the preceding "market clearing" definition, and also allows for stationary but non-market clearing configurations.

[2] Specifically, individuals are in equilibrium when they are forming the "correct" (i.e., strong form rational) expectations on the basis of available information.

endogenous variables, and (most importantly) which derive their structure from a set of *exogenous* variables and coefficients imposed upon the system from without. For example, in neoclassical partial equilibrium analysis of consumer and producer behaviour, individual consumption and production decisions (the endogenous variables) are related to exogenous variables such as relative prices, and coefficients representing preferences and technology.[3] These exogenous variables and coefficients are taken as given, forming what may be referred to as the "data" of the system, determined independently of individual behaviour by the market, consumer psychology and the state of science respectively.

Second, traditional equilibrium models are typically constructed so as to yield stable equilibria - that is, points of rest to which the system will return following any arbitrary displacement.[4] This concern with stability follows naturally from concern with economic equilibria, since equilibrium configurations are not of general interest unless stable.

Stability is also important because it implies that an

---

[3] Prices are, of course, endogenous to the market process, but assuming perfectly competitive conditions, they are exogenous to the individual decision maker.

[4] The notion of stability used here corresponds to the formal mathematical concept of asymptotic stability. This implies that following some initial disturbance, all subsequent motion will lead a system towards convergence, in the limit, with the equilibrium from which it was initially disturbed.

equilibrium configuration *defined* in terms of exogenous data may also be *reached* by the system to which it pertains from *any* arbitrary starting point. What the preceding considerations suggest, then, is a definition of the traditional equilibrium approach to economic analysis as one in which the long run or final outcomes of economic systems are seen as being determinate (Kaldor, 1933); they are both defined and reached without reference to the (historical) adjustment path taken towards them.

a. The use of traditional equilibrium methodology in Keynesian and neoclassical macroeconomics.

The widespread use of the traditional equilibrium approach in macroeconomics can be illustrated by considering two diverse examples of this methodology, both of which are characterised by stable equilibrium configurations defined in terms of exogenous data. Equations [1.1] and [1.2] below comprise a highly simplified Keynesian income - expenditure model:

$$Y_t = C_t + I \qquad [1.1]$$

$$C_t = \alpha Y_{t-1} \quad , \quad 0 < \alpha < 1 \qquad [1.2]$$

where Y_t represents national income, C_t represents realised aggregate consumption, α is the propensity to consume, I is the (exogenously given) desired level of investment, and the subscripts denote time periods.

In this system, there are two structural equations in two unknowns (Y_t and C_t), with Y_{t-1}, I and α taken as given. Substituting equation [1.2] into equation [1.1], we arrive at:

$$Y_t = \alpha Y_{t-1} + I \qquad [1.3]$$

Setting $Y_t = Y_{t-1} = Y^E$ now yields:

$$Y^E = \frac{1}{1 - \alpha} . I \qquad [1.4]$$

- an equilibrium income configuration, Y^E, defined in terms of the exogenous data I and α.

Suppose now that in some period t, a transitory shock gives rise to a once over increase in the value of income, so that $Y_t > Y_{t-1} = Y^E$. The effects of this are captured by the general solution of the difference equation [1.3], which can be written as:

$$Y_t = \alpha Ab^{t-1} + I/(1 - \alpha)$$

where $Y_t = Ab^t$, or[5]

[5] Equation [1.5] follows from the homogeneous function derived from equation [1.3], which can be written:

$$Y_t = \alpha Y_{t-1}$$

$$Y_t = A\alpha^t + I/(1 - \alpha) \qquad [1.5]$$

From equation [1.5] we can see that as $t \to \infty$, $A\alpha^t \to 0$ if $-1 < \alpha < 1$ - i.e., that Y_t will converge to the steady state defined in equation [1.4] if $-1 < \alpha < 1$, which is clearly satisfied when $0 < \alpha < 1$. Hence the specification of the model in equations [1.1] and [1.2] ensures that this model is stable around the equilibrium Y^E following any arbitrary disturbance.

Now consider the following natural rate model:

$$(U_t - U_n) = \alpha(\dot{p}_t - \dot{p}^e_t) + e_t \quad , \quad e_t \sim (0, \sigma^2_e) \qquad [1.6]$$

$$\dot{p}^e_t = \dot{p}^e_{t-1} + \beta(\dot{p}_{t-1} - \dot{p}^e_{t-1}) \ , \ 0 < \beta < 1 \qquad [1.7]$$

$$\dot{p}_t = \dot{m}_t \qquad [1.8]$$

where $\dot{p}_t$ is the rate of price inflation, $\dot{p}^e_t$ is the expected rate of price inflation, U_n is the (exogenously given) natural rate of unemployment, U_t is the actual rate of unemployment, $\dot{m}_t$ is the (exogenously given) rate of money supply growth, e_t is a random error term with zero mean and constant variance, and subscripts again denote time periods.

This time, there are three equations in three unknowns (U_t, $\dot{p}_t$, and $\dot{p}^e_t$) with α, β, $\dot{m}_t$, $\dot{p}^e_{t-1}$ and U_n taken as given. Setting $\dot{p}_t = \dot{p}^e_t = \dot{p}^e_{t-1} = \dot{p}^*$, equation [1.6] can be written:

Recalling that $Y_t = Ab^t$, this implies that:

$$Ab^t - \alpha Ab^{t-1} = 0$$

$$\Rightarrow \quad b = \alpha$$

$$U_t = U_n + e_t$$

Recalling that $E(e_t) = 0$, we arrive at:

$$E(U_t) = U_n \qquad [1.9]$$

- an equilibrium unemployment solution defined in terms of the exogenous data U_n.

Suppose now that we introduce a once over increase in $\dot{m}$ in period t-1 so that $\dot{p}^* = \dot{p}_t > \dot{p}^e_t$. This time, the effects of a disturbance to the model are captured by the general solution to the difference equation [1.7], which can be written as:

$$\dot{p}^e_t = (1 - \beta)Bd^{t-1} + \dot{p}^*$$

where $\dot{p}^e_t = Bd^t$, or[6]

$$\dot{p}^e_t = B(1 - \beta)^t + \dot{p}^* \qquad [1.10]$$

---

[6] Again, equation [1.10] follows from the homogeneous function derived from equation [1.7], which can be written:

$$\dot{p}^e_t = (1 - \beta)\dot{p}^e_{t-1}$$

Recalling that $\dot{p}^e_t = Bd^t$, this implies that

$$Bd^t - (1 - \beta)Bd^{t-1} = 0$$

$$\Rightarrow \quad b = (1 - \beta)$$

from equation [1.10], we can see that as $t \to \infty$, $B(1 - \beta)^t \to 0$ if $-1 < (1 - \beta) < 1$ - i.e., that $\dot{p}^e_t$ will converge to $\dot{p}^*$ as long as $-1 < 1 - \beta < 1$, which is satisfied when $0 < \beta < 1$. The structure of the model in equations [1.6] - [1.8] again ensures that this model is stable in response to some initial disturbance, this time around the equilibrium configuration defined in equation [1.9].

Each of the examples chosen above is dynamic, which allows us to explicitly illustrate that the long run outcomes deriving from them are both defined and reached without reference to the path taken towards them - that is, the long run outcomes are determinate. Whilst many macroeconomic models are formally static, so that only the existence of equilibrium can be formally specified, these models nevertheless assume or allege stability in an intuitive manner. For example, the assumption of stability is implicit in comparative static exercises, such as those commonly performed with the IS-LM system, and with the "Marshallian cross" in the analysis of supply and demand.

b. Towards a critique of the traditional equilibrium approach.

The traditional equilibrium approach to macroeconomic modelling suggests both the existence and stability of a configuration which is independent of the past history of the system to which it pertains. In other words, it conceives long run economic outcomes as being of a determinate nature, both defined and reached without reference to the path taken towards them.

The diversity of the examples provided above illustrates the popularity of this approach in macroeconomics. However, the traditional equilibrium approach to economic modelling can be subject to a number of important criticisms.

II The "realism" of traditional equilibrium models.

Some authors (for example, Kaldor, 1972) have criticised particular variants of the traditional equilibrium approach on the basis that their assumptions lack realism.[7] These assumptions or axioms influence not only the choice of data in a model (i.e., what is taken as "given") but also the form of its structural equations and its general methodology. Indeed, they may even influence the type of questions a model asks (i.e., what precisely it seeks to explain). Kaldor criticises general equilibrium theory (and by extension, neoclassical theory as a whole) for making assumptions that are either impossible to falsify (for example, consumers "maximise" utility) or which are at variance with what we

[7]. Kaldor is quite explicit in identifying the variant he wishes to criticise as "the general economic equilibrium originally formulated by Walras, and developed ..... by the mathematical economists of our own generation." (Kaldor, 1972, p. 1237.) However, many of the criticisms he makes could be applied to neoclassical theory in general.

observe in real world economies[8] (for example, perfect competition and wholly impersonal market relationships.)

The idea that it is legitimate to criticise a theory or group of theories for being "unrealistic" has not met with unanimous approval. Friedman (1953) argues from a positivist perspective that a theory should be judged not by its realism, but on the basis of its predictive abilities. Hahn (1974), meanwhile, argues that by clearly specifying the restrictions necessary to obtain its results, even a patently unrealistic construct (such as the Arrow-Debreu model of general equilibrium) can be of great value in indicating what we likely *cannot* claim about the functioning of real world economies.

Coddington (1972, 1975) criticises and rejects the methodological positions of both Friedman and Hahn. He argues that positivism is ultimately unworkable, due to difficulties associated with unambiguously falsifying the predictions of economic hypotheses. These difficulties include data deficiencies, and the problem that economic hypotheses come in "clusters," so that the apparent falsity of any particular hypothesis can always be blamed on the inadequacy of the supporting hypotheses necessary to generate the original testable proposition (see also McCloskey, 1983). Meanwhile, Coddington argues that Hahn's position is itself positivist;

---

[8]. The term "real world economy" refers to an actual, data generating economy, as opposed to a theoretical construct.

it rests on the idea that the mathematical precision of the Arrow Debreu model is useful because precise statements permit the claims of a model to be subject to the test of falsification.

However, in criticising Friedman and Hahn, Coddington (1975) establishes a further basis for rejecting the criteria of realism. For Coddington, all theories are necessarily abstract, and hence unrealistic. Criticising a theory for lacking realism involves the mistaken belief that theories should replicate their subject matter - something which they are not intended to do and, indeed, are incapable of doing. According to Coddington, the value of theory is that it "can 'impose coherence' and provide 'insight' without being aimed primarily at description" (Coddington 1975, p.542).

Coddington is correct to point out that any abstraction and hence any theory is, by definition, unrealistic. However, since different models make different abstractions, a theory may be criticised as unrealistic not because it makes abstractions per se, but because it makes the *wrong* abstractions. For example, traditional equilibrium models will subsequently be criticised for treating as exogenous variables which should be thought of as being endogenous to macroeconomic activity. In this particular case, the "wrong abstraction" compromises traditional equilibrium models by leading them to overlook the effects of intertemporal adjustment processes in an economy on final economic outcomes.

Essentially, what is being suggested is that if, following Coddington (1975), Wilber and Harrison (1978) and McCloskey (1983), we reject the positivist position of evaluating theory solely on the basis of its predictions, and if instead we interpret economic theory as a collection of "metaphors" and "parables" designed to "provide insight without being aimed primarily at description," then it is legitimate to question how suitable any particular theory is as a metaphor for reality - i.e., how realistic it is.

III The possibility of multiple equilibria.

Another objection to the traditional equilibrium approach involves the frequently overlooked possibility of multiple equilibria. This implies that even when we can adequately characterise an economic process in terms of a system of stable structural equations defined in terms of a given set of exogenous data, the solution of this system need not yield a *unique* equilibrium configuration. The importance of this possibility is twofold. First, multiple equilibria may be Pareto rankable - that is, different equilibrium configurations of the same system may have different welfare implications. In this case, the notion of equilibrium loses much of the normative significance that is usually attached to it;[9] a system can be "in equilibrium" without this outcome being globally optimal.

Second, when there exist multiple equilibria, then by definition no one equilibrium configuration can be *globally* stable. This immediately raises the possibility that final outcomes in an economic system may depend on initial conditions, and hence which of the locally stable multiple equilibria the system converges towards.[10]

The possibility of multiple equilibria is given more explicit recognition in some variants of the traditional equilibrium approach than in others. For example, the Second Fundamental Theorem of Welfare Economics recognises that any Pareto optimal point is a market equilibrium for some initial distribution of resources.[11] Market equilibrium is not unique, therefore, in any circumstances where initial endowments are not fixed (for example, when there exists a government with redistributive functions.) Furthermore, even if initial endowments are fixed, it is well known that a

---

[9] A strong association exists in economics between equilibrium and optimal outcomes, although it is well known from situations such as the static prisoner's dilemma problem that this association does not necessarily hold.

[10] This point pre-empts the notion that final outcomes may not be independent of the path taken towards them, which is dealt with at length in section V and chapter 2 below.

[11] The Second Fundamental Theorem of Welfare Economics is a property of general competitive systems. The First Fundamental Theorem of Welfare Economics states that a Walrasian (market) equilibrium is Pareto efficient. See Varian (1984, pp. 220-1).

general equilibrium may not be unique - the number of equilibria depends on the shape of preferences (see, for example, Varian 1984, p.241).

Similarly, explicit recognition of the possibility of multiple equilibria is made in game theory, especially when non-cooperative solutions are considered. It is widely acknowledged that these solutions may proliferate in any n-person non-cooperative game (Shubik 1981, p.307).[12] This can be illustrated in the context of the following two player strategic form game, in which all of the payoff vectors constitute non-cooperative equilibria:

		Player 2: A	Player 2: B
Player 1	A	(5,1)	(3,1)
	B	(5,10)	(3,10)

In this game, regardless of the behaviour of their opponent, neither player will express a preference between his/her available strategies. For example, if player 2 is known to have chosen strategy A, player 1 will be indifferent between strategies A and B (assuming that he/she is an individual welfare maximiser with no "desire to punish")

[12] Note that the possibility of multiple equilibria in non-cooperative games of strategy is not contingent on assumptions made about the number of players.

because they will each yield an identical payoff. Hence any of the outcomes of the game may be sustained as non-cooperative equilibria since, ceteris paribus, for any arbitrary strategy pair, no individual has any incentive to change his/her current behaviour.

However, in many variants of the traditional equilibrium approach, and particularly in macroeconomic analysis, the possibility of multiple equilibria is ignored. For example, the original formulation of the natural rate hypothesis postulates a natural rate of unemployment representing

> the level that would be ground out by the Walrasian system of general equilibrium equations.
>
> (Friedman, 1968, p.8.)

Friedman simply assumes that this Walrasian equilibrium is unique, and indeed this assumption has continued to advise virtually all subsequent theorising on the natural rate hypothesis.

A second example is provided by the theory of equilibrium business cycles in New Classical macroeconomics. Basic representations of this theory assume that the economy permanently gravitates around a unique Walrasian equilibrium, subject only to the perturbations caused by serially correlated random errors (see, for example, Lucas 1975). Again, the possible implications of multiple equilibria are overlooked.

IV "Getting into" equilibrium.

A third criticism of the traditional equilibrium approach is related to the phenomenon of disequilibrium, or more precisely, the process commonly referred to as disequilibrium adjustment, by which a system is supposed to regain equilibrium following some initial disturbance.[13]

The informal characterisation of an economic system "getting into" equilibrium is one of the most frequently used and powerful allegories in the "rhetoric of economics" (McCloskey, 1983). Robinson (1974a) argues that the most curious feature of equilibrium in this allegory is that it is conceived as representing the end of an economic process - a state that attains after a discrete interval of time during which the economy has been in disequilibrium. For Robinson, it appears counter-intuitive that a situation where economic agents are forming the "wrong" present expectations on the basis of the "wrong" current prices, levels of output, capital stock etc. (as is the case under conditions of disequilibrium) should lead to the state of order that is equilibrium. Concerns of this nature can be summarised by the following question: once an economic system is *out* of equilibrium, can it be supposed that forces will operate so as to get it *back*

[13]. It should be noted that some rational expectations models circumvent the issues discussed here by conceiving the economy as a moving equilibrium i.e., as a system which never experiences disequilibrium (although this does not explain how the economy is supposed to achieve equilibrium in some initial period.)

into equilibrium?

Considerable scepticism has been expressed on this issue, not least by some eminent theorists from within the equilibrium tradition. Hence Hahn (1970) notes that whilst

> (e)quilibrium economics,[14] because of its well known welfare economics implications, is easily convertible into a apologia for existing economic arrangements.... (there is)..... no support for the view that any of the traditional methods of response of various agents to changes in their economic environment makes the '(invisible) hand' perform as it is often taken to perform."
>
> (Hahn, 1970, p.1)

One conventional approach to the problem of disequilibrium adjustment is to postulate the existence of an auctioneer, whose task is to reconcile all economic activity by means of adjusting prices before any activity actually takes place. In effect, this "solves" the problem of disequilibrium adjustment by removing the phenomenon from the sphere of economic activity. The auctioneer is implicit in the notion of instantaneous disequilibrium adjustment,[15] and is therefore more widely appealed to than is commonly recognised. However, the assumption of what amounts to a central planner

---

[14]. The term "equilibrium economics" used here may be thought of as referring broadly to the Walrasian and Marshallian traditions in neoclassical economics.

[15]. Although his/her role may be taken by strong form rational expectations, in which agents know the "true" model of the economy and hence instantaneously adjust their behaviour to that which is compatible with equilibrium.

is clearly unrealistic in the context of a decentralised economy. Acknowledging this forces us to conceive disequilibrium adjustment as a process of sequential recontracting endemic to the economy, and occurring in real historical time. But this in turn means that we must contend with the reservations of authors such as Robinson and Hahn, regarding the capacity of such a process to guide an economic system into equilibrium.

Unfortunately, economic theory has not gone out of its way to address these reservations. Indeed, the majority of investigations into the disequilibrium behaviour of economic systems do not extend far beyond pure allegory, usually based on the statement of heuristic adjustment rules such as "excess demand leads to a change in price." Such adjustment rules posses no innate features that guarantee the type of adjustments necessary to lead a system back into equilibrium. The mere fact that economic agents in a state of disequilibrium will seek to change their behaviour does not in and of itself imply that these changes will subsequently result in the attainment of equilibrium.

As illustrated earlier, it is possible to specify the mathematical conditions necessary for stability in formal dynamic models. However, even then, attention is seldom paid to the necessary *economic* conditions that these mathematical conditions imply.[16] In spite of the contribution made by mathematical formulations of stability conditions, the following problems remain regarding the process of disequilibrium adjustment and the prospects for "getting into" equilibrium.

a. The problem of control over economic variables.

Models which make the assumption of perfect competition suffer the conceptual weakness of failing to specify precisely who is supposed to signal disequilibrium by changing prices (Arrow, 1959). Recall that in a perfectly competitive economy, no one individual has any direct control over prices. If perfect competition is interpreted literally, then, models which make this assumption require an auctioneer to make disequilibrium price adjustments. But the assumption of what amounts to central planning in a decentralised economy is, as we have already seen, untenable.

This problem is, of course, easily overcome by relaxing the assumption of perfect competition. However, this cannot be done without forfeiting the accompanying welfare propositions which, as Hahn notes, motivate a good deal of the attention

---

[16] In view of this, and in light of the discussion in section II above, formal dynamic models are open to criticism as being "unrealistic" in their representation of stability conditions. That such a criticism is possible is itself testimony to McCloskey's (1983) argument that mathematical economic reasoning is inherently metaphorical, and can therefore be questioned with regard to its aptness as a characterisation of any given economic process.

which is paid to models within the traditional equilibrium approach.

b. Stocks and flows, and their impact on disequilibrium adjustment.

Further problems arise if we allow for the fact that some goods may accumulate in the form of stocks, so that in the process of disequilibrium trading, "mistakes" will carry over from period to period[17]. Unless all goods are flows, we cannot postulate that all agents begin each period of activity with the same initial endowments; hence in the absence of an auctioneer to prevent trade from occurring until equilibrium is reached, the process of tatonnement breaks down (Varian, 1984 p.247).[18] Instead, we are forced to postulate that disequilibrium adjustment occurs via a non-tatonnement process, where stocks of goods are allowed to carry over from period to period. However, such a process does not permit the prior specification of the long run outcome that will arise from any given starting point (i.e., vector of initial

---

[17]. For example, this problem will certainly arise in any model in which the process of production takes place, unless we assume that production decisions have no concrete manifestation in physical assets i.e., that we are working with "putty capital." See, for example, Robinson (1974b.)

[18] Tatonnement is a disequilibrium adjustment process in which agents "grope" for an equilibrium outcome over time basing their demands and supplies in each period on the *same* initial endowment. Non-tatonnement processes, on the other hand, allow endowments to change in the course of disequilibrium trades.

endowments.) Initial disequilibrium trades at disequilibrium prices will alter individuals endowments so that in the next "period," new disequilibrium trades at new disequilibrium prices will take place which will further alter individual endowments and so on. In sum, because trading activity in non-tatonnement processes is a function of current rather than initial endowments, and because current endowments are influenced by past "mistakes" made in the course of disequilibrium trading which cannot be predicted a priori, the equilibrium to which non-tatonnement processes lead is determined ex post rather than ex ante.[19]

c. The speed of disequilibrium adjustment.

A third problem concerns the speed with which disequilibrium adjustment is supposed to take place. Although assumed "given," the data of a traditional equilibrium model may change over time. For example, we would not expect consumer preferences to remain fixed indefinitely. It is quite common in the context of the traditional equilibrium approach to postulate exogenous changes in the data of a model in order to motivate comparative static/dynamic analysis. An important and neglected point, therefore, is how fast the process of disequilibrium adjustment takes place relative to the speed

---

[19] Hence non-tatonnement processes do not give rise to determinate outcomes. In fact, the outcomes of non-tatonnement processes are subject to what will subsequently be identified as a process of hysteresis.

with which the "data" of a model is changing over time. As Cornwall (1991) notes:

> ... if ... real world change(s) in tastes, technologies and other institutional features are very rapid relative to the rate at which the economy can adjust, the convergence properties of the model take on much less interest and importance than the institutional changes themselves.
>
> (Cornwall, 1991, p.107.)

d. Connections between markets.

Each problem with the process of disequilibrium adjustment that has been discussed so far applies equally to individual market (i.e., "partial equilibrium") and economy wide (i.e., "general equilibrium") analysis. There exists a further problem, however, which relates specifically to general equilibrium analysis, and concerns the existence of connections between markets. This problem is central in Keynesian macroeconomic analysis. Hence for Keynes (1936), wages play an important dual role, as both a cost of production, and as a source of aggregate demand in the circular flow of income. Part of Keynes' contribution was to argue that this neglected interrelatedness between the labour market and the demand side of the product market denies the possibility of automatic adjustment towards a full employment equilibrium, since strategies designed to reduce real labour costs (i.e., real wages) will simultaneously reduce aggregate demand, and hence the derived demand for labour.[20]

e. Signalling problems in disequilibrium.

A final problem connected with the process of disequilibrium adjustment concerns the issue of signalling. Some proponents of the traditional equilibrium approach insist that equilibrium must be reached, since otherwise unexploited gains from trade will persist, and this is inconsistent with the postulate of individual rationality.

However, there is an important caveat in this argument. Unexploited gains from trade cannot be regarded as a sufficient condition to entice rational individuals to change their current behaviour in an equilibrating manner. As Hahn (1982) points out, if mutually beneficent trades are to take place, the opportunities for these trades must be both recognised and successfully signalled. But the extent to which such opportunities will be successfully signalled depends vitally on the ways in which potential parties to a trade are able to communicate, and hence what signals can be sent and

[20]. Note that this Keynesian critique relates primarily to the notion of adjustment towards the type of general equilibrium described by Walras, and in turn makes use of a new equilibrium concept - the short run under employment equilibrium. This concept has already been identified as part of the traditional equilibrium approach; hence adjustment towards *it* may be subject to some of the criticisms previously outlined (for example, the problem regarding the speed of adjustment relative to the rate of change of the data of the model may be apposite.) The relationship between Keynesian and neoclassical equilibrium concepts and their relative usefulness is taken up in more detail in chapter 2.

how they are interpreted.

In a decentralised economy with incomplete information in which individuals communicate impersonally, the problems associated with signalling are enormous. Indeed, the interpretation of even the most elementary signal, such as a fall in price, is open to doubt. For example, if price is taken as an indication of quality, lowering the selling price of a good may be interpreted as a signal of its inferiority. Rather than increasing the number of prospective buyers as in conventional exchange theory, a fall in price may therefore reduce the demand for a product (Stiglitz, 1987). In sum, if the interpretation of market signals under conditions of imperfect information is uncertain, it is not clear whether the communication that takes place between agents in a state of disequilibrium will be of a nature that will lead these agents into a state of equilibrium.

V Hysteresis.

The arguments of the preceding section question the stability of equilibrium configurations in the traditional equilibrium approach by asking whether it is conceivable for an economic system to "get into" equilibrium as the result of a prior period of disequilibrium adjustment. However, the stability issue can be pursued further. In particular, it is important to ask whether it is at all legitimate to model economic processes in terms of determinate equilibria - i.e., as having long run outcomes which can be defined and reached without reference to the adjustment path taken towards them.

Recall that the orthodox notion of equilibrium constitutes a stable configuration[21] derived from a series of structural equations defined in terms of exogenously given data. What is especially interesting about this notion is that the equilibrium configuration is conceived as being completely independent of the process of disequilibrium adjustment - i.e., it is independent of the path the system takes towards equilibrium. Traditional equilibrium models postulate equilibrium configurations that depend only on the form of the (alleged stable) structural equations of the model and, most importantly (since this component is left unexplained by the models themselves,) the nature of the alleged exogenous data. The equilibria of such models are thus entirely independent of the past history of the systems from which they derive.

As was noted earlier, the intertemporal adjustment behaviour of economic systems is seldom explicitly considered by traditional equilibrium models. Yet as Kaldor (1933) notes:

> It is not possible .... to determine the position of equilibrium from a given system of data, since every successive step taken in order to reach equilibrium will alter the conditions of equilibrium .... and thus change the final position - unless the conditions are such that either (1) an equilibrium .... (is) established immediately, or (2) the (situation) ....

[21]. Or configurations, if we allow for the possibility of multiple equilibria. In this case, of course, any individual equilibrium configuration can only be locally stable.

> actually established leaves the conditions of equilibrium unaffected (in which case the final position will be independent of the route followed.)
>
> (Kaldor, 1933, p.124)

It is precisely these limiting assumptions which are implicit in most variants of the traditional equilibrium approach. Clearly, the first assumption must be rejected if we are to take serious account of the genuinely intertemporal nature of economic processes. What remains to be seen is whether the second assumption should also be rejected. Is it appropriate to conceive long run economic outcomes as being independent of the prior adjustment paths taken by the systems from which they derive? This issue brings us to the concept of path dependency, or hysteresis.

Chapter 2.

An alternative approach - modelling with hysteresis.

I Introduction.

If we assume that an economy is permanently in equilibrium (i.e., that it will adjust instantaneously to this state from any arbitrary starting point), then questions posed by the nature of the intertemporal behaviour of the economy are of little importance.[1] As long as economic activity is governed by an equilibrium, which is itself determined at any point in time by stable structural relations and a set of exogenously given data, the economy will simply replicate itself through time in the form of a steady state growth path. There will be no endogenous tendencies for the existing configuration of the economy to change, whilst the disturbing effects of random shocks will be instantaneously corrected.

However, in chapter 1, it was suggested that equilibrium

---

[1] One exception to this occurs in equilibrium business cycle models. It was noted earlier that these models seem to circumvent the problems posed by the process of disequilibrium adjustment, by postulating that individuals are always in equilibrium by virtue of the fact that their conditional expectations are (strong form) rational. However, equilibrium business cycle models do not avoid the problem of path dependency or hysteresis. This is because the sub-optimal short run equilibria that they postulate as arising from conditional expectations formed in the presence of "price surprises" will necessarily create a sub-optimal history of the economy in the short run. The "mistakes" made during this period may influence the future behaviour of the economy - in particular, the nature of the long run equilibrium about which it is presumed to oscillate.

is traditionally conceptualised as being the end of an economic process. Even within the traditional equilibrium approach to macroeconomic modelling, then, there is tacit recognition of the fact that economic outcomes arise from sequential patterns of economic activity (in this case, a sequence of disequilibrium trades) - that is, that economic processes occur in real historical time. Furthermore, it was seen that consideration of the genuinely intertemporal nature of economic activity poses important problems for traditional equilibrium analysis. The prospects of an economic system "getting into" equilibrium become less favourable. There is also the possibility, indicated towards the end of chapter 1, that by neglecting to consider the precise adjustment path taken by an economic system in a state of disequilibrium, traditional equilibrium analysis overlooks an important potential influence on the determination of final economic outcomes. It is to the discussion of the far reaching implications of this latter possibility that the present chapter is devoted.

II The importance of adjustment paths in determining economic outcomes: the process of hysteresis.

The term hysteresis originates in the natural sciences where, for example, it is used to describe the electromagnetic properties of ferric metals in the process of magnetization and subsequent demagnetization[2]. Hysteresis exists when the long run or final value of a variable depends on the value of the variable in the past, by virtue of the influence of this past value on the current alleged exogenous variables, coefficients and structural equations which characterise the system that determines the variable.[3] The central characteristic of hysteresis is, therefore, that it causes the *long run or final* outcome of a system to depend on its previous outcomes - that is, the long run or final outcome depends on the path taken towards it. Notice that this distinguishes hysteretic systems from other dynamic systems whose short run outcomes are path specific, but which ultimately converge to configurations defined independently of the path taken towards them.[4]

Recent attempts to mathematically characterize the process of hysteresis (for example, Wyplosz, 1987; Franz, 1990) have tended to proceed in terms of systems of equations similar to the following:

---

[2]. See Cross and Allen (1988) for a fuller discussion of this and other points relating to the history of hysteresis.

[3] Note that this definition of hysteresis is consistent with Elster's (1976) claim that the past can only be seen to influence the present by virtue of the traces left by the past in the present. In the preceding definition, these "traces of the past" are reflected by the sensitivity of the current structural model to past values of the variable of interest.

[4] This point is illustrated in detail in sections IIa and IIb below.

$$X_t = \Omega W_t \qquad [2.1]$$

$$W_t = f(X_{t-1}, Z_t) \qquad [2.2]$$

In this "mainstream" characterization, hysteresis may occur as a result of the influence of past values of the endogenous variable X on the current values of the alleged exogenous variable W; both the coefficient Ω and the variable Z_t are treated as exogenous data. Consider the case where equation [2.2] can be given the simple linear functional form:

$$W_t = \beta X_{t-1} + \gamma Z_t \qquad [2.3]$$

Substituting [2.3] into [2.1] yields

$$X_t = aX_{t-1} + \phi Z_t \qquad [2.4]$$

where $a = \Omega\beta$ and $\phi = \Omega\gamma$. By setting $X_t = X_{t-1} = X^*$, equation [2.4] may be solved to obtain:

$$X^* = \phi Z^*/(1 - a) \qquad [2.5]$$

where an asterisk (*) represents the steady state value of a variable. However, if a = 1, that is if the difference equation [2.4] possesses a unit root, equation [2.5] cannot be solved for X^*. Instead, given the existence condition $Z^* = 0$,[5] we can only define from [2.4] the system of equations:

$$X_t = X_{t-1} + \phi Z_t$$

$$X_{t-1} = X_{t-2} + \phi Z_{t-1}$$

$$\vdots$$

$$X_1 = X_0 + \phi Z_1$$

which can be solved to yield:

$$X_t = X_0 + \phi . \sum_{i=1}^{t} Z_i \qquad [2.6]$$

where Z_i represents the time dependent values of Z.[6] This

---

[5] Note that from [2.4], if a = 1, in the steady state,

$$X_t - X_{t-1} = 0 = \phi Z^*$$

which implies $Z^* = 0$ for all non-trivial values of ϕ.

[6] Note that only time dependent values of Z are of interest in this formulation. If we consider the steady state value of Z, equation [2.6] reduces to

$$X_t = X_0 + t\phi Z^*$$

Recalling the existence condition $Z^* = 0$, this implies that

$$X_t = X_0$$

which simply states that X has reached a steady state value - precisely the assumption originally made to generate the existence condition.

suggests that at any point t in time, all that can be derived is some "contemporaneous" value of the variable X which depends on its own past history. Under these conditions, then, the simple linear model in [2.4] is hysteretic.

The model summarised in [2.6] is useful for illustrating an important feature of hysteretic systems. Note that in [2.6], X_t depends on the entire timepath of Z.[7] This implies that the system has a "long memory," the importance of which is that it will never "forget" any feature of its timepath - even what appear to be transitory shocks. Hence in complete contrast to, for example, New Classical macroeconomics, hysteresis suggests that even random disturbances may have long lasting effects which permanently influence the future behaviour of an economic system, and hence its long run configuration.

a. Hysteresis and persistence.

Hysteresis occurs in equation [2.6] only as a result of the assumption that a = 1, which was neither explained nor justified. This leads authors such as Wyplosz (1987) and Franz (1990) to conclude that hysteresis is a special case, which occurs only when a dynamic system comprising a set of

---

[7] Not all hysteretic systems are influenced by their entire past history, however. See section II.2 below and Cross (1991).

difference equations possesses a unit root,[8] and which is encompassed by the more general traditional equilibrium approach. Nowhere is acceptance of this more evident than in the literature that claims to investigate the incidence of hysteresis in the natural rate of unemployment. Following Gordon (1989), this may be illustrated in the context of the following version of the natural rate hypothesis:

$$\dot{p}_t = \alpha\dot{p}_{t-1} + \beta(U_t - U^*) \qquad [2.7]$$

where $\dot{p}_t$ and $\dot{p}_{t-1}$ are the rates of inflation in periods t and t-1 respectively, and $U_t - U^*$ represents the difference between the actual and the natural rates of unemployment in period t. Given [2.7], it is usually argued that hysteresis exists if U^* depends on the past history of unemployment, as well as other microeconomic variables captured by the vector Z_t.[9] Hence we can write:

$$U^* = \eta U_{t-1} + \gamma Z_t \qquad [2.8]$$

Substituting [2.8] into [2.7], we arrive at:

---

[8] The analogous condition in continuous time is that a system of differential equations possess a zero root. See Giavazzi and Wyplosz (1985).

[9]. Note that these microeconomic variables are assumed *exogenous*, a vitally important point to which we shall subsequently return.

$$\dot{p}_t = \alpha\dot{p}_{t-1} + \beta(1 - \eta)U_t + \beta\eta.\Delta U_t - \beta\gamma Z_t \qquad [2.9]$$

Under the assumptions of the natural rate hypothesis, we may set $\alpha = 1$. In the steady state, $\dot{p}_t = \dot{p}_{t-1} = \dot{p}^*$, and $U_t = U_{t-1} = U^*$, which would conventionally allow us to solve for the natural rate as follows:

$$U^* = \gamma Z^*/(1 - \eta) \qquad [2.10]$$

However, if $\eta = 1$, equation [2.10] cannot be solved for U^*. Indeed, there exists no determinate natural rate in this case, which is referred to as "full hysteresis." But if $0 < \eta < 1$,[10] equation [2.10] does admit a determinate solution for U^*, although there may exist situations in the short run where, with $\dot{p}_t = \dot{p}_{t-1} = \dot{p}^*$, we have:

$$U_t^* = \gamma Z_t + \eta U_{t-1} \qquad [2.11]$$

Here, U_t^*, the "contemporaneous" natural rate, represents "... that level of unemployment which is consistent with stable inflation *during the current period* ..." (Layard, Nickel and Jackman, 1991, p.382). This case is referred to as "persistence". U_t^* is intended to represent a temporary

[10] Note that the case where $-1 < \eta < 0$ is not considered since, although mathematically admissible, it has no economic interpretation in equation [2.8] in terms of the theory of hysteresis in the labour market.

equilibrium sensitive to past rates of unemployment (Cross, 1991, p.10), as distinct from the determinate long run or steady state equilibrium defined in [2.10]. Thus whilst full hysteresis is regarded as a special case, "hysteresis" may nevertheless occur in the short run in the form of unemployment persistence, as indicated by equation [2.11]. However, in the case of persistence, there continues to exist a determinate long run or steady state value of the natural rate defined solely in terms of exogenous data, and uninfluenced by the timepath of previous unemployment. Hence the system ultimately gravitates towards a configuration which is independent of the path taken towards it; it is ahysteretic, and belongs in the domain of traditional equilibrium analysis.

b. A familiar example.

The idea of persistence is not, in fact, new. In order to understand it more fully, it is constructive to consider a special case of a familiar macroeconomic model which displays persistence. The system of equations below will immediately be recognised as Samuelson's multiplier-accelerator model, where Y is the level of national income, C represents aggregate consumption, I and G are the levels of investment and government expenditure respectively, and the subscripts denote time periods:

$$Y_t = C_t + I_t + G \qquad [2.12]$$

$$C_t = \alpha Y_{t-1} \qquad [2.13]$$

$$I_t = \beta(C_t - C_{t-1})$$

$$= \alpha\beta(Y_{t-1} - Y_{t-2}) \qquad [2.14]$$

Substituting [2.13] and [2.14] into [2.12] yields:

$$Y_t = \alpha(1 + \beta)Y_{t-1} - \alpha\beta Y_{t-2} + G_t \qquad [2.15]$$

and if we assume that $\alpha\beta < 1$ and $\alpha < 4\beta/(1 + \beta^2)$, it can be shown that the general solution of this model will yield an income time path which displays damped cycles[11].

It appears from the nature of these equations that the model is hysteretic, since it is apparent from equation [2.15] that the variable Y depends, in any period t, on its own past values. However, the specific values of the parameters chosen above ensure that following any arbitrary disturbance, the system will converge to the determinate equilibrium:

---

[11]. See, for example, Henin (1986). Attention is confined to the case of damped cycles to avoid the possibility of divergence in the multiplier-accelerator model, which cannot be interpreted as persistence. Indeed, such behaviour may be closer to the notion of hysteresis (see also Cross, 1987, p.79). However, it is important to note that hysteresis is not usually characterized in terms of model instability, and indeed the two are not generally equivalent. Hence models displaying hysteresis may approach states of rest, although any such equilibria they reach will, of course depend on the prior disequilibrium adjustment path taken towards them.

$$Y^* = G/(1 - \alpha)$$

What the model in fact describes is an elaborate process of oscillatory disequilibrium adjustment, in which the current or "contemporaneous" value of income in equation [2.15] depends on its values in the past, but in which the level of income ultimately converges to an equilibrium which is independent of its prior adjustment path. The simple multiplier-accelerator model is therefore ahysteretic, although it does display what has been described above as persistence.

III Criticisms of the hysteresis/persistence dichotomy: reconstructing hysteresis as a general case.

Whilst the characterizations outlined above suggest that hysteresis is a special case, closer inspection suggests the contrary - that the importance of unit roots and the distinction between hysteresis and persistence are more apparent than real, and that hysteresis is a potentially general phenomenon.

a. Hysteresis and the process of disequilibrium adjustment.

If $a \neq 1$ the system of equations [2.1] and [2.3] yields the steady state solution:

$$X^* = \phi Z^*/(1 - a) \qquad [2.5]$$

However, for this to be a relevant description of the system at any point in time, we must be able to show that the configuration in [2.5] can be reached. As was pointed out in chapter 1, regardless of the existence of a steady state, we cannot conclude that a system will "get into" equilibrium without specifying the appropriate dynamic convergence process. Fortunately, the system that is currently being contemplated is of a simple nature. The only form of hysteresis that is being admitted is the influence of past values of the endogenous variable X on the alleged exogenous variable W. Other channels of hysteresis (such as the possible influence of past values of X on the parameters and the form of the structural equations of the system) are being ignored.

This enables us to characterize equations [2.1] and [2.3] as a simple difference equation of the form:

$$X_t = aX_{t-1} + \phi Z_t \qquad [2.4]$$

Studying the general solution of equation [2.4] enables us to be quite specific about the behaviour of the timepath of the variable X. This general solution can be written:

$$X_t = aAb^{t-1} + \phi Z^*/(1 - a)$$

where A and b are constants, and $X_t = Ab^t$. This equation can in turn be rewritten as:[12]

$$X_t = Aa^t + \phi Z^*/(1 - a)$$

which implies that X_t will converge to the equilibrium value indicated by equation [2.5] for all values of a such that $-1 < a < 1$.

From a mathematical point of view, this suggests that under conditions more general than those previously invoked to generate hysteresis, the equilibrium in equation [2.5] is determinate. However, it is important to consider what assumptions are necessary for this result to be relevant for an *economic* system. Perhaps the most important, as noted in section IVc of chapter 1, is that the speed of adjustment of X towards the equilibrium in [2.5] must be fast relative to the speed at which the data underlying this equilibrium are themselves changing. But suppose that the convergence of X is a long run phenomenon, or at least of sufficient longevity to allow the equilibrium in [2.5] to alter its configuration

---

[12]. Note that the homogeneous function derived from the general solution to [2.4] can be written as:

$$X_t - aX_{t-1} = 0$$

Recalling that $X_t = Ab^t$, this yields:

$$Ab^t - aAb^{t-1} = 0$$

$$\Rightarrow \quad b - a = 0$$

$$\Rightarrow \quad a = b$$

before it is ever reached. Essentially, we are assuming that persistence occurs over a sufficiently long period of time to allow the data which determine long run equilibrium to change. Then this long run equilibrium will not provide an appropriate description of the configuration of the system at any point in time. In this case, the "contemporaneous" value of X in equation [2.4] is all that is of interest. The behaviour of X will be described by the equations:

$$X_t = aX_{t-1} + \phi Z_t$$
$$X_{t-1} = aX_{t-2} + \phi Z_{t-1}$$
$$\vdots$$
$$X_1 = aX_0 + \phi Z_1$$

which upon substitution yields:

$$X_t = a^t X_0 + \phi . \sum_{i=1}^{t} a^{t-i} Z_i$$

which in turn implies that the value of X in any period t is specific to its prior adjustment path - i.e., that it is hysteretic.

What this illustrates is that an important implicit assumption in the literature that distinguishes between hysteresis and persistence is that the speed of adjustment towards long run equilibrium is relatively fast. If, however, persistence occurs over a sufficient period of time to allow the data underlying this long run equilibrium to change - that is, if the speed of adjustment is relatively slow - long run equilibrium may never be reached. Instead, as illustrated above, the distinction between hysteresis and persistence breaks down; the current value of X *always* depends on its prior adjustment path, even in the long run. This suggests that hysteresis can exist even in systems with unique equilibria defined independently of the path taken towards them, if these equilibria fail to govern the long run behaviour of such systems, which instead remain in a constant state of disequilibrium.

b. The importance of functional forms.

The second problem with mainstream characterizations of hysteresis is that the results to which they give rise are contingent on linear functional forms. Suppose that instead of equations [2.1] and [2.3] we have:

$$X_t = \Omega W_t \qquad [2.1]$$

$$W_t = \delta X_{t-1} . Z_t \qquad [2.3a]$$

where the coefficient δ is again taken as given. Substituting [2.3a] into [2.1] now yields

$$X_t = bX_{t-1}.Z_t \qquad [2.16]$$

where b = $\Omega\delta$. Imposing steady state conditions on [2.16] admits no solution for X^*. Instead given the existence condition $bZ^* = 1$,[13] we can only define the system of equations:

$$X_t = bX_{t-1}.Z_t$$

$$X_{t-1} = bX_{t-2}.Z_{t-1}$$

$$\vdots$$

$$X_1 = bX_0.Z_1$$

which can be summarised as:

$$X_t = b^t X_0 . \prod_{i=1}^{t} Z_i \qquad [2.17]$$

---

[13] Note that from equation [2.16]:

$$\left(\frac{X_t}{X_{t-1}}\right) = bZ_t$$

$$\Rightarrow bZ^* = 1$$

in the steady state.

This indicates that the value of X depends on its prior timepath for any value of $b \neq 0$ - i.e., that the system is hysteretic for any non-trivial value of b.[14]

The generality of hysteresis in the presence of non-linearities is also recognised by Cross (1990, 1991), who provides a more general discussion of non-linear hysteretic models based on the work of Krasnosell'skii and Pokrovskii (1989).[15] One important feature of the systems considered by Cross is that their current outcomes are influenced only by the *dominant extrema* of past "inputs" to the system, rather than by all past inputs.[16] Hence models of this nature have

---

[14] Once again, time dependent values of Z appear in equation [17] since the case where $Z = Z^*$ is uninteresting. This can be seen by rewriting [17] as

$$X_t = X_0 . \prod_{i=1}^{t} bZ^*$$

$$= X_0 . (bZ^*)^t$$

$$= X_0$$

(recalling that $bZ^* = 1$), which simply states that X has reached a steady state value, as initially assumed in generating the existence condition.

Note that it is possible to conceive of instances when X_t in equation [17] will converge to zero regardless of its prior timepath (for example, if $-1 < b < 1$ and $Z_i < 1$ for all values of $i = 1, \ldots ,t$). These cases are overlooked as being of little general economic interest.

[15] For a treatment of hysteresis in non-linear systems, see also Mayergoyz (1985, 1991).

[16] To clarify the terminology used here, note that Z_i (for all $i = 1, \ldots, t$) represents the "input" of the system summarised in equation [2.17].

"selective memories," in the sense that only *some* of their history is relevant in the determination of current outcomes arising from them (see Cross, 1991, pp. 32-34).

c. Hysteresis in linear models: the concept of "deep endogeneity."

The single most important criticism of the characterizations of hysteresis outlined in section I is that they embody a misapplication of the concept of hysteresis itself. To fix ideas, consider once again the steady state solution to the system described by equations [2.1] and [2.3]:

$$X^* = \phi Z^*/(1 - a) \qquad [2.5]$$

Now suppose that we differentiate this expression totally with respect to X_{t-1}, where X_{t-1} is a value from the prior disequilibrium adjustment path taken by X towards the steady state equilibrium X^*. We thus obtain:

$$\frac{dX^*}{dX_{t-1}} = \frac{\partial X^*}{\partial a} \cdot \frac{da}{dX_{t-1}} + \frac{\partial X^*}{\partial Z^*} \cdot \frac{dZ^*}{dX_{t-1}} \qquad [2.18]$$

where ϕ is taken as given, and dX^*/dX_{t-1} captures the sensitivity of X* to the past (non-steady state) timepath of X.

Normally, we would expect the expression in [2.18] to equal zero, since equation [2.5] is intended to represent a determinate steady state solution. But in order for this result to hold, it must be that $da/dX_{t-1} = 0$ and $dZ^*/dX_{t-1} = 0$ – i.e., the coefficient a and the steady state value of the variable Z must be independent of the time path of X.[17] However, a central point of the hysteresis critique of traditional equilibrium analysis is to question whether conventional "data" can in fact be regarded as exogenous in this manner. Consider, for example, the literature that interprets traditional "data" such as tastes and technology as being endogenous to the economic outcomes (consumption and production) they are held to explain.[18]

Suppose, then, that in addition to equations [2.1] and [2.3] we have:

$$Z_t = f(X_{t-1}) \quad , \quad f' \neq 0 \qquad [2.19]$$

that is, Z is dependent on the time path of X, or is what may be described as "deeply endogenous."[19] Assuming for

---

[17] The possibility that $(da/dX_{t-1}) = -(dZ^*/dX_{t-1}) \neq 0$ is overlooked as an unlikely special case. If traditional equilibrium analysis were, in fact, to depend on this possibility, it is unlikely that this methodology would command the widespread attention that it does.

[18]. See, for example. Georgescu-Roegen (1971), Arthur (1988), and Witt (1990).

[19] Following the Lucas (1976) critique of econometric policy evaluation, as a result of which conventional "data" such as tastes and technology have been referred to as "deep parameters," the term deep endogeneity is deliberately

simplicity that da = 0, equation [18] now becomes

$$\frac{dX^*}{dX_{t-1}} = \frac{\partial X^*}{\partial Z^*}.\frac{dZ^*}{dX_{t-1}} = \frac{\Phi}{1 - a}.f'$$

and since $f' \neq 0$ by [2.19], this implies that the long run value of X depends on its own past values, by virtue of the influence of these past values on the variable Z. The system is consequently hysteretic, regardless of the value of the coefficient a.

Alternatively, consider the case where equations [2.1] and [2.3] are augmented by the expression:

$$a = g(X_{t-1}) \qquad , \qquad g' \neq 0 \qquad [2.20]$$

that is, where a is a deeply endogenous coefficient.[20] If we now assume for simplicity that $dZ^* = 0$, equation [2.18] becomes:

---

contrived to suggest the potential sensitivity of at least some conventional "data" to the economic outcome that is being explained.

[20] It is common to overlook the potential endogeneity of coefficients in formal economic models. The importance of allowing for this possibility is illustrated in detail in Part II, where coefficients representing technological and institutional structures are considered as deeply endogenous in a long run model of growth. A simple example which, for now, suffices to indicate the significance of equation [2.20] is the endogeneity of individuals' fertility behaviour to the level of economic development.

$$\frac{dX^*}{dX_{t-1}} = \frac{\partial X^*}{\partial a} \cdot \frac{da}{dX_{t-1}} = \frac{\phi Z^*}{(1 - a)^2} \cdot g'$$

Since $g' \neq 0$ by [2.20], this means that the long run value of X depends on its own past values, by virtue of the influence of these past values on the coefficient a. The system is again hysteretic, regardless of the value of a.

d. The nature of deep endogeneity in macroeconomic systems: the importance of adjustment asymmetries.

The idea that even traditional "data" such as tastes and technologies are deeply endogenous appears to suggest that we must be willing to treat the entire structure of macroeconomic models as simultaneously endogenous in order to generate hysteresis. In fact, this is not the case, as is illustrated by Setterfield (1992). However, the question which then arises is whether deep endogeneity actually matters. Is it not possible that changes in the endogenous components of a model induced by its prior adjustment path will "cancel out" over time, leaving long run variables a determinate function of genuinely exogenous data?

To further understand the nature of this question, consider the following expression:[21]

$$Z_t = f(X_0, \ldots, X_{t-1}) \qquad [2.21]$$

Equation [2.21] is an analog of equation [2.19], expressing Z as a function of the disequilibrium adjustment path of X, (X_0, ..., X_{t-1}). Under what circumstances, then, can Z^*, the long run value of Z, be treated as independent of the timepath of X, making the original steady state in [2.5] determinate?

Suppose that $DX = \sum_{i=0}^{t-1} dX_i = 0$ – i.e., that following some initial disturbance, X follows a disequilibrium adjustment path along which it returns to its original value over time. Suppose further we assume that:

i) $f_i' \neq 0$ for some $i = 0, \ldots, t-1$ and

ii) $\sum_{i=0}^{t-1} f_i'.dX_i = 0$

What this means is that changes in X along its disequilibrium adjustment path (dX_i) have a short run effect on Z ($f_i' \neq 0$ for some i in assumption (i), so that $f_i'.dX_i \neq 0$ for some i) but

[21] In what follows, we formally examine the deep endogeneity of the variable Z. Similar conclusions can be reached regarding the treatment of deeply endogenous coefficients such as a. This is illustrated in the context of the model developed in Part II.

that these variations in Z "cancel out" over the course of the adjustment path ($\sum_{i=0}^{t-1} f_i'.dX_i = 0$ in assumption (ii)). Ultimately, then, the long run value of Z, Z^*, is unaffected by the disequilibrium timepath of X, and so the long run value of X, X^*, will also be unaffected by its own prior timepath. Instead, X^* will be given by the determinate equilibrium in equation [2.5].

This is precisely the type of reasoning implicit in the hysteresis/persistence dichotomy. This dichotomy admits the possibility of some heretofore overlooked endogeneity in economic relationships, but assumes that endogenously induced changes in economic "data" have no lasting significance - they effectively "cancel out" over time. The result of this assumption is that whilst short run outcomes display persistence, long run outcomes are nevertheless determinate.

This is a highly particular treatment of the nature of endogeneity in adjustment processes, which depends critically on assumption (ii). Suppose, then, that we retain assumption (i), but replace assumption (ii) with:

iii) $$\sum_{i=0}^{t-1} f_i'.dX_i \neq 0$$

This time, the variations in Z_t with respect to X_t do not "cancel out" over time, so that the long run value of Z will depend on the prior timepath of X. This, in turn, implies that the long run value of X will depend on its own past values - i.e., we will have hysteresis, and under more general formal conditions than those necessary to generate a determinate equilibrium outcome.

It only remains to show what *economic* conditions are implied by assumption (iii). In general, the outcome in assumption (iii) will be observed whenever there exist "adjustment asymmetries" in economic relationships displaying deep endogeneity. In our example, the deeply endogenous variable Z can be said to display adjustment asymmetries if a change in this variable induced by a change in the variable X cannot be reversed simply by restoring the original value of X. Suppose, for example, that an increase in X along its disequilibrium adjustment path induces an increase in Z. Then if Z is subject to adjustment asymmetries, a series of offsetting reductions in X in subsequent periods will not induce Z to return to its original value. As a result, the adjustment path followed by X will induce changes in the long run value of Z, which will in turn induce changes in the long run value of X itself - i.e., the long run value of X will depend on its own past values, and will therefore display hysteresis.

There are already well known examples of how such adjustment asymmetries can occur, and how they can give rise to hysteresis. Hence there may exist adjustment asymmetries in behavioural norms, due, for example, to "revenge motives." Cornwall (1991) argues that labour may feel that it has been "made to pay" the costs of an anti inflationary policy which relies on high unemployment to reduce wage demands. Hence even if unemployment subsequently falls, there may exist a strong desire to "get even" for the perceived inequities of previous periods, which will change workers attitudes towards the wage bargaining process. This can give rise to hysteresis in the long run Phillips curve (Cornwall, 1991). Adjustment asymmetries may also arise in firms costs of production, due to sunk costs which are incurred as a result of entry into a market. Sunk costs are, by definition, expenditures which cannot be recovered if the action that led to them being incurred is subsequently reversed. This can give rise to hysteresis in international trade flows (Dixit, 1989).

e. An illustration.

The role of deep endogeneity and adjustment asymmetries in generating hysteresis can be illustrated in terms of the natural rate model developed earlier. According to this model, the natural rate of unemployment can be expressed as a function of past values of the unemployment rate, and a set of alleged exogenous microeconomic variables affecting individuals willingness and ability to find market work.[22]

---

[22]. This relationship is captured by equation [2.8] in the model developed earlier.

However, numerous contributions to the debate over hysteresis in the natural rate have emphasised that these "exogenous" variables are themselves systematically influenced by the past history of unemployment (see, for example, Phelps, 1972; Hargreaves Heap, 1980; Jenkinson, 1987.) Hence the scrapping of physical capital (i.e., a rundown in the size of the capital stock) which may accompany periods of high unemployment will affect the quantity constraints faced by individuals in the labour market. Human capital may also be eroded by high levels of unemployment, especially when it comprises a substantial element of firm specific skills.

If rising unemployment is associated with a lengthening of unemployment spells, we may also observe factors associated with negative duration dependence. The effectiveness of search may be impaired by demoralization,[23] or the loss of informal contacts within the labour market. Furthermore, if long spells of unemployment are identified by firms as negative credentials, this may reduce the probability of an individual unemployed worker finding new employment.

These considerations imply that the natural rate model in equations [2.7] and [2.8] does not adequately capture the variety of channels through which hysteresis may operate in the labour market. The model is itself a special case, which can be rationalized by arguing that, in the context of an

---

[23]. This is related to the notion of a "discouraged worker" effect.

insider-outsider model, wage setting behaviour is dominated by changes in the rate of unemployment rather than its level. The hypothesis here is that employed workers at the margin will perceive increases in unemployment rather than the level of unemployment as the chief threat to their future job prospects. Hence the wage setting strategy of marginal insiders will not be affected by a high but stable rate of unemployment, so that a once over increase in the rate of unemployment may eventually lead to the establishment of a new, higher, natural rate. This hypothesis is captured by equation [9] with $\eta = 1$.

What equations [7] and [8] formally capture, then, is the possibility of hysteresis in the natural rate due to the invariance of insider wage setting behaviour to the number of outsiders, in an insider-outsider model of the labour market. What these equations overlook is that this is only one way in which hysteresis in the labour market may occur; we need also take account of the effects of unemployment on the characteristics of "outsiders" (their human capital, search, tastes for market work and the nature of any quantity constraints that they face). The possibility that these variables are subject to adjustment asymmetries must also be taken into consideration. Suppose, for example, that depressed demand conditions induce both a level of unemployment above the natural rate, and capital scrapping (Soskice and Carlin, 1989). Even if demand subsequently expands, it is not obvious that lost capacity will automatically be replaced. Having scrapped physical plant and equipment, and lost established positions in input, capital and final product markets, firms may face start up costs which would not have been encountered had pre-existing capacity simply been under utilised rather than scrapped in response to the demand shock.[24] If this is the case, there will exist an adjustment asymmetry in the capital stock. Increasing unemployment above the natural rate will consequently induce a permanent change in the quantity constraints facing workers in the labour market, thus raising the value of the "natural rate" (or what may now be more appropriately referred to as the long run rate of unemployment).

What this illustrates is that even if we are able to write:

$$U^* = \gamma Z^*/(1 - \eta) \qquad [2.10]$$

where $\eta < 1$, if the vector of variables Z is affected by past values of U - that is, if Z is deeply endogenous - and if at least some components of Z are subject to adjustment asymmetries, then U^* in equation [2.10] will depend on the path taken towards it. The case where $\eta = 1$ is a sufficient but not a necessary condition for the existence of hysteresis

[24] Essentially, then, we assume that market entry is not costless.

in the "natural rate," and so rejection of the *specific* hypothesis that this assumption embodies does not constitute a rejection of the *general* hypothesis of hysteresis in the labour market.

IV Hysteresis as an alternative approach to macroeconomic modelling.

The preceding arguments illustrate the limits of unit root characterizations of hysteresis, and suggest that the distinction between hysteresis and persistence to which they give rise is not useful. In so doing, they show that hysteresis is not a special case, but must be treated as a potentially pervasive feature of macroeconomic systems.

Incorporating a generalized notion of hysteresis into macroeconomic models can therefore be identified as an important new research agenda for macroeconomics. Such a research agenda would avoid the ahistoricism of the traditional equilibrium approach, the results of which are ultimately determined by unexplained external data. It would also avoid the pitfall, inherent in current theorising, of permitting an illusory escape from hysteresis back into the traditional equilibrium approach via the notion of persistence. What remains to be seen, however, is precisely what the treatment of hysteresis as a general phenomenon implies for macroeconomic modelling.

The first implication is clearly negative. If hysteresis is a prevalent feature of economic activity, then the traditional equilibrium approach to macroeconomic modelling is guilty of overlooking an important aspect of macroeconomic dynamics. More specifically, hysteresis suggests that it is misleading to conceive the behaviour of an economic system as being governed by a determinate equilibrium. With hysteresis, any long run or final configuration that is reached will depend on the path taken towards it; the evolution of a system will be best described by "contemporaneous" values of variables expressed in terms of their own past history.

However, the second implication is positive - namely, that within the hysteresis critique of traditional equilibrium analysis lie the foundations of an alternative methodology for macroeconomic modelling. The key element of this alternative methodology is the notion that economies function in real historical time - economic events occur in a uni-directional sequence rather than instantaneously, and these sequential patterns of activity matter in the sense that they affect final outcomes. This idea is inherent in Keynes.[25] In the

---

[25]. Although Marshall is usually seen as the father of period analysis, Keynes provides a more appropriate starting point since Marshall's short and long runs are not sequential; the latter bears no strict intertemporal relationship to the former. Hence in Marshall, "there is no analysis of a process which could generate long period equilibrium as the (asymptotic) result of a sequence of short period equilibria, just as the short period equilibrium is not explicitly constructed as the outcome of output adjustments in a sequence of ultra short period equilibria. . . . The problem (of the historically consistent aggregation of time) is ingeniously side stepped by the use of ceteris paribus clauses in a nested set of equilibrium concepts." (Skott, 1983, p.10)

General Theory, the analysis of the general equilibrium effects of falling real wages is based explicitly on the fact that

> The hiring of workers and the marketing of the output produced by those workers are separated in time.
>
> (Skott, 1983, p.23)

i.e., that the hiring of workers, the implementation of a specific production plan, and the sale of final output comprises a distinct sequence of activities and that this sequence matters when we come to analyse the levels of output and employment in an economy.

a. Hysteresis and cumulative causation.

The insight that macroeconomic dynamics occur sequentially in real historical time is an inherent feature of the Keynesian tradition in macroeconomics. Recognition of the importance of sequential patterns of economic activity is not only implicit in the General Theory, but is also explicit in the work of authors such as Robinson (1974a, 1980) and Kaldor (1972, 1985). Particularly important in this respect is Kaldor's promotion of the concept of cumulative causation. Cumulative causation involves a circular interaction between economic variables, in such a way that an initial change in a variable x induces supporting changes in the vector of variables Z which reinforce the initial change in x. Hence with cumulative causation,

> the actual state of the economy during any one 'period' cannot be predicted except as a result of the sequence of events in previous periods that led up to it.
>
> (Kaldor, 1972, p.1244)

Comparing the definition and consequences of cumulative causation above with the definition and consequences of hysteresis discussed earlier, it is immediately apparent that a close correspondence exists between these concepts. This correspondence is not exact (that is, the concepts are not formally equivalent) because cumulative causation implies that successive changes in a variable are positively correlated - for example, an initial increase in x will induce subsequent increases in x. This is not necessarily true of hysteretic systems.[26] However, cumulative causation is based on the same methodological conception of the economy that we have already identified as the key organizing concept of a hysteretic research agenda - that of the economy as a historical entity,

---

[26] For example, in a model of institutional hysteresis, an improvement in economic institutions which facilitates an improvement in economic performance may subsequently lead to a worsening of institutions which in turn gives rise to a deterioration of economic performance. See Cornwall (1990), and chapter 5 below.

In fact, the notion that successive changes in a variable are positively correlated over time is a potentially serious shortcoming of cumulative causation in the long run. This issue is taken up in detail in chapters 4 and 5, in the context of the growth model developed in Part II.

which instead of adjusting inevitably towards some determinate equilibrium, comprises sequential patterns of activity which can, in and of themselves, affect the nature of final economic outcomes. In this way, Kaldor's treatment of the principle of cumulative causation can be interpreted as a preliminary attempt to understand the workings of a macroeconomy in hysteretic terms. It offers a concrete indication of how we might begin to conceive the process of modelling macroeconomic systems with hysteresis - a point to which we will subsequently return in chapter 3.

b. Hysteresis and lock in.

Hysteresis suggests an approach to macroeconomic modelling which, because of its attention to sequential patterns of activity, is compatible with real historical time. As has already been pointed out, this is in and of itself an important point of contrast with the traditional equilibrium approach. However, this central organizing concept of the hysteretic research agenda is also important because it brings other "new" concepts into macroeconomic debate, which have hitherto been overlooked because they have no place in traditional equilibrium analysis. Of particular importance in this respect is the concept of lock in. Lock in is a property of dynamic systems that arises when sequential patterns of activity form a "groove" from which it subsequently becomes difficult for a system to deviate. Lock in occurs when current behaviour is conditional on either past events, or the behaviour of other agents in the system.[27] These "frequency dependency effects" may give rise to situations where to deviate from the pursuit of a particular activity requires non-marginal adjustments, which it does not pay any one individual to undertake within a given period, or which are beyond the control of the individual decision making unit. Hence the system becomes "locked in" to this activity.

Lock in can occur in a variety of specific circumstances. Suppose, for example, that we are faced with a set of economic activities, each of which is designed to fulfil the same objective, that are such that the returns to these activities are directly proportional to the extent that they are practised - for example, they may involve learning by doing. Then it can be demonstrated that at each point in time, the rational choice of activity will be the one first chosen - even if, in some long run sense, it is inferior to one or all of the alternative activities (Arthur, 1988.) Hence the system becomes "locked in" to this initial activity.

A simple example serves to illustrate this point. Consider a person faced with a simple career choice - say law or medicine - where the rewards to each activity increase with experience.[28] Suppose that the initial rewards to law are

---

[27] Note that to the extent that lock in is conditional on the events of the past, it can be identified as a path dependent outcome.

[28]. This example is discussed by Arthur (1988).

highest, but are relatively quickly surpassed by those of medicine. Then a rational agent with a high discount rate may choose law as his/her career, and will continue to practise it at all times thereafter (since at all future points in time, the short run returns to *continuing* in law will exceed the short run rewards to *beginning* a career in medicine.) However, after a brief period of time in law, the person may come to realise that an equivalent time spent in medicine would have ensured higher earnings at all future dates, and therefore decide that medicine was in fact the career with the greater long term potential. But at no point in time will it pay the individual to switch to medicine; he/she is effectively locked in to a career in law.[29]

A second, and particularly important phenomenon which may generate lock in is interrelatedness (Frankel, 1955). Interrelatedness can be defined as interconnections between components of the production process (technologies, institutions, individual decision making units) which are usually thought of as being divisible and separable. There are two dimensions to interrelatedness, both of which can give rise to lock in. First, interrelatedness can occur *within* the individual decision making unit, in which case current behaviour is made conditional on the own past behaviour of the

---

[29] Note that this result relies upon the individual in question having a finite time horizon. Otherwise, the realisation that medicine is the superior option in the long run would presumably motivate a career change.

individual. Suppose, for example, that the components of the production process under the control of an individual decision maker are interrelated. Then changing any one of these components may require complementary changes in other components with which it is interrelated. In this scenario, the problem which interrelatedness poses is that change may become prohibitively costly. As a result, the current choices of the individual decision maker are constrained in such a way that it becomes locked in to its own past practices.

Second, interrelatedness can occur *between* individual decision making units, in which case the current behaviour of any one individual is conditional on the behaviour of others in the system. Suppose now that the components of the production process under the control of *different* individual decision making units are interrelated. Then a change in any one of these components by a particular individual decision maker may require accompanying changes in other components which are beyond the range of his/her control. For example, no individual firm can adopt broad gauge rolling stock if railway companies remain committed to narrow gauge tracks. The problem which interrelatedness now poses is that the decision making unit is small relative to the size of the unit of change required. This results in the system (i.e., all individual units) becoming locked in to current practices.

It should be clear from the nature of lock in as described above that this phenomenon may constitute an important source of adjustment asymmetries. If it is possible for certain features of the economy - such as its technological or institutional structures, for example - to become locked in to certain forms from which it is subsequently difficult for them to deviate, then simply restoring the external conditions that existed prior to the creation of these technological and institutional forms may not be sufficient to eliminate them. Lock in may, therefore, play an important role in generating hysteretic outcomes in the economy as a whole. In chapters 4 and 5, we will further explore this possibility, with the aim of demonstrating its potential implications for the evolution of an economy. For the time being, however, note that what lock in illustrates is that even when agents are making rational choices, they may yet experience "regret" (Arthur, 1988) i.e., the awareness that the type of activities that they are undertaking are demonstrably inferior to an available alternative.

c. Towards a hysteretic model of the macroeconomy.

It is important to realise that hysteresis, cumulative causation, and lock in are not equivalent concepts. Nor, indeed, does the existence of, say, hysteresis necessarily imply the existence of phenomena such as lock in. Rather, it is simply the case that all three of these concepts are mutually compatible, and this is so because of a methodological feature common to each of them; emphasis on the notion that economic activity occurs in historically sequential patterns. Cumulative causation and lock in have been discussed because they offer potential insights into the likely behaviour of a hysteretic economy. The challenge that is presented to us, then, is to construct a model of macroeconomic activity based on sequential patterns which is hysteretic, and which allows us to discuss the implications for macroeconomic evolution of phenomena such as lock in. A necessary first step in this procedure is to identify what constitutes the basic "building block" of such a model, and it is to this task that we now turn.

V Hysteresis and different equilibrium organising concepts.

As intimated at the beginning of chapter 1, the term "equilibrium" is used in a variety of different contexts in economics. It may be used in one sense in reference to the balance of payments, and quite another in reference to consumer theory. Moreover, what we have identified as the "orthodox notion" of equilibrium varies in important ways in its role as an organising concept, between what may be broadly defined as "Keynesian" and "neoclassical" approaches to macroeconomics.

a. Keynesian short run under employment equilibria.

In Keynesian macroeconomics, equilibrium is defined in terms of aggregates interacting in the circular flow of income. It defines a situation where either individuals do not desire to change their behaviour (because for example, planned saving is equal to planned investment), or do wish to change their behaviour, but are unable to influence the relevant variables (for example, individual workers cannot relieve the quantity constraints they face in the labour market). In Keynesian macroeconomics, therefore, we arrive at the notion of the under employment equilibrium. This is designed to capture the possibility that an economy may be in equilibrium, but that this may be suboptimal in the precisely defined sense that there exists involuntary unemployment.[30]

b. Equilibrium in neoclassical models.

In neoclassical macroeconomics, equilibrium is defined in terms of the same methodological individualism that advises neoclassical microeconomics. "Agents" with well defined

---

[30]. The Keynesian underemployment equilibrium is not a general equilibrium in the conventional sense of this term, since it does not describe a situation where supply is equal to demand in all markets (specifically, there exists excess supply in the labour market.) Conventionally, however, it has been thought useful to characterise Keynesian situations as "aggregate equilibria" (see Henin, 1986.) In connection with this, note that Keynesian situations may be thought of as equilibria in the broad sense of the term, as states of rest where no individual desires to change his/her current behaviour given the range of variables they are able to influence. Keynes himself argued that individual labour supply was influenced by relative wages as well as the own real wage, and that this may give rise to downward wage stickiness (i.e., an unwillingness to change current behaviour) even in the presence of involuntary unemployment. Similar results are generated by information theoretic models of the labour market (see, for example, Stiglitz, 1987.)

objective functions are seen as interacting in an impersonal market place in response to price signals. Equilibrium is achieved as a result of optimization by rational maximizing individuals. Although suboptimality may occur as a result of market failure, the underlying theme of this approach is that as long as such failures can be overcome, a decentralised market system will automatically organize individual maximizing activities into an aggregate optimum. Any debate over the nature of equilibrium is thus concentrated on how much more or less frequent and/or intransigent instances of market failure are.

c. Keynesian and neoclassical equilibria and the treatment of time.

The nature of equilibrium as an organizing concept is thus qualitatively different between Keynesian and neoclassical macroeconomic theory. Further illustration of this point is provided by considering the treatment of the distinction between short run and long run equilibrium in the two frameworks. In neoclassical theory, the temporal treatment of equilibrium is traditionally based on the quasi-relatedness of Marshall's periods.[31] The most important feature of this analysis is the relationship between the short run and the long run - or more precisely, the lack of any such relationship, since long run equilibrium is determinate and is

---

[31]. See footnote 21 above.

achieved regardless of the nature of any preceding short run equilibria. In other words, the long run equilibrium configuration is assumed to be independent of the short run equilibria established in prior periods, which led up to it.

Keynesian theory, on the other hand, is explicitly short run, and its concept of equilibrium is studied entirely in this context. No attempt is made to analyse the long run, although we are made aware that there will be no automatic tendency in subsequent periods for the economy to adjust to an optimal (i.e., full employment) equilibrium. However, one thing that is abundantly clear in this short run theory is that since equilibrium is determined on the demand side, there will be an inherent tendency for it to change over time in response to changes in autonomous demand. Hence a "Keynesian long run" would appear to be made up of a series of (potentially different) short run equilibria. It should be noted, of course, that a neoclassical long run equilibrium may also change over time due to changes in "data" such as tastes and technologies. However, changes in the equilibrium configuration due to such factors as these are not usually regarded as being of particular interest. On the contrary (and this is the key point of contrast) in Keynesian theory, examining ways of changing the short run equilibrium from period to period (i.e., redefining the equilibrium of the economy in the long run) is a central part of the analysis, not least because this is seen as being the key to eliminating involuntary unemployment.

d. "Reclaiming" the short run under employment equilibrium as a pedagogic device in a hysteretic model of the macroeconomy.

These deliberations suggest that it may be appropriate to retain the Keynesian short run under employment equilibrium as a pedagogic device in a hysteretic model of macroeconomic performance. There are two basic reasons for this. First, as is clear from the discussion of different equilibrium organizing concepts in the short and long runs, the Keynesian notion of under employment equilibrium is imbued with a "natural" intertemporal flexibility. Unlike its neoclassical counterpart, emphasising changes in the configuration of the short run under employment equilibrium over time and the welfare significance of these changes is an intrinsic feature of Keynesian macroeconomics. Such emphasis follows naturally from consideration of situations in which some individuals are constrained from making what would otherwise be desirable changes in their current behaviour (for example. when they are involuntarily unemployed). This lends the central concept of Keynesian macroeconomics to modelling with hysteresis. Starting from the position of regarding the short run under employment equilibrium as subject to changes over time, hysteresis can be interpreted as suggesting mechanisms that may generate these changes endogenously as a result of dynamic properties of the economy. For example, suppose an increase in government expenditure increases national income to a new, higher, short run under employment equilibrium. If this simultaneously raises the optimism of long run expectations, the resulting shift in the marginal efficiency of capital schedule may increase investment, hence further increasing the equilibrium level of income and so on.[32] The marriage of these two concepts, then, suggests a way in which we may model the aggregation of successive short runs to form a picture of the long run which is broadly consistent with historical time, and which in the process provides us with a long run theory of effective demand.

The second reason for reclaiming the short run under employment equilibrium is somewhat more practical. Note that what is essentially being proposed above is that we "lock up" any short run effects of hysteresis in the hope of modelling its longer term implications. We are allowing the short run under employment equilibrium to exist in any given period as a temporary equilibrium, which will not be sustained owing to subsequent hysteresis effects acting so as to redefine its position in the long run. In this sense, the model that is being proposed is not entirely faithful to all the

---

[32]. This notion of hysteresis in the level of national income operating via changes in the state of long run expectations bears close resemblance to Kalecki's analysis of the dynamics of investment. See, for example, Targetti and Kinda Hass (1982). The fact that early Keynesian authors such as Kalecki incorporated processes similar to hysteresis into their models further reinforces the idea that hysteresis and Keynesian macroeconomics are highly complementary.

implications of real historical time. However, as Samuelson (1965) notes:

> When the equilibrium of a system depends on (and is dictated) by its path toward equilibrium, the scientist has an uncomfortable feeling.
>
> (Samuelson, 1965, p. 112)

And well he/she might, since if history replaces equilibrium entirely, the "scientist" has lost one of his/her central pieces of analytical apparatus. Indeed, so serious would this loss appear to be, that for practical modelling purposes, we must seek to reclaim part of the equilibrium framework.

e. The short run under employment equilibrium as a traditional element of Keynesian pedagogy.

The view of the short run under employment equilibrium developed above - that of its being a necessary pedagogic device and probably little more - appears to be justified in terms of the history of economic thought. In fact, it seems to coincide closely with the role that Keynes himself intended for it. Hence Kregel (1976) argues that at the time of the General Theory, Keynes' thinking was divided between three different types of models of the macroeconomy. The first two of these, which appear in the General Theory, share as an important common feature the assumption of constant long run expectations.[33] Kregel argues that Keynes used this assumption to

> 'lock up' (the) effect of general expectations and uncertainty without assuming that they did not exist.
>
> (Kregel, 1976, p.212.)

and hence to demonstrate the principle of effective demand via the use of the short run under employment equilibrium. However, Kregel goes on to argue that Keynes' lectures of 1937 reveal his consideration of a third model in which long run expectations are subject to revision in the light of prior disappointment, implying that the underlying determinants of aggregate demand (the propensity to consume, liquidity preference, and the marginal efficiency of capital) are endogenous to the past history of the economy. The important feature of this model is that

> if ... realisation of error alters the state of expectations and shifts the independent behavioural functions, Keynes' model of shifting equilibrium will describe an actual path of an economy over time chasing an ever changing equilibrium ...
>
> (Kregel, 1976, p.217.)

i.e., the model displays hysteresis - its final or long run

---

[33]. They differ with respect to their treatment of short run expectations, as either realised on average, or disappointed.

configuration depends on the prior adjustment path of the economy. Kregel argues that

> the extreme complexity of such a situation explains why Keynes was willing to 'tame' this system, first making assumptions that allowed the definition of functions that he knew did not exist through their entire range , expressly in order to give force to the theory of effective demand.
>
> (Kregel, 1976, p.216.)

In other words, Keynes invoked the notion of the short run under employment equilibrium as a pedagogic device to illustrate the workings of his short run theory. It is for similar reasons that the concept may be retained in order to investigate the long run behaviour of a model akin to Keynes' shifting equilibrium system. To quote Kregel once more:

> For Post-Keynesian theory, the next step is not disequilibrium, but the model of shifting equilibrium in a monetary - production economy.
>
> (Kregel, 1976, p.220.)

Equipped with the short run under employment equilibrium and the complementary concepts of hysteresis, cumulative causation and lock in, and focusing attention on the long run, it is this step that we shall now attempt to take.

Part II.

Modelling Macroeconomic Systems With Hysteresis: Growth and Structural Change in Capitalist Economies.

Introduction.

In Part I, it was argued that Keynes' short run under employment equilibrium may provide an important theoretical foundation for a hysteretic model of macroeconomic performance. Adopting this notion as a starting point for our analysis will allow us to take as given the macroeconomic outcomes associated with short run under employment equilibria. We may then concentrate on modelling endogenous changes in these short run equilibrium configurations in the long run, due to hysteresis.[1] This will give rise to a moving short run under employment equilibrium model which captures

---

[1] The concepts of the "short" and "long" run are defined here in Marshallian terms. Hence the short run corresponds to a period during which the level of autonomous demand, the quantity and quality of the capital stock and the institutional structure of the economy are unchanging - i.e., to a given under employment equilibrium. The long run, meanwhile, constitutes a period during which the "data" underlying this short run equilibrium are subject to change.

The "long run" during which hysteresis effects prevail may represent a different period of actual calendar time, depending on what aspect of the economy we are considering as being subject to hysteresis. It will become apparent in Part III, and is consequently worth noting at the present time, that the effects of path dependence on economic growth and structural change take place over decades of actual calendar time.

medium and long term hysteretic relations in capitalist development.

In the course of this exercise, we will have succeeded in constructing a long run analog to the short run circular flow of income model developed by Keynes in the General Theory - a model which may be used to illustrate the fragile nature of successful macroeconomic performance in capitalist economies. Of course, this propensity of capitalist economies to malfunction - to experience prolonged periods of poor macroeconomic performance - is precisely what Keynes (1936) himself was concerned to illustrate. In this sense, the model developed may be thought of as representing a long run theory of effective demand.

Chapter 3 begins by identifying an important potential source of hysteresis - the existence of increasing returns to scale. This recalls Kaldor's contributions to non-equilibrium growth theory. Much of the chapter is devoted to the construction of a Kaldorian model of cumulative causation, which is interpreted as a preliminary attempt to understand macroeconomic dynamics under conditions of hysteresis. Chapters 4 and 5 then extend these beginnings by considering the impact of technological and institutional structures on growth. This extends the model of cumulative causation developed in chapter 3 into a more generally hysteretic long run form, and in so doing, illustrates the propensity of a capitalist economy to malfunction in the course of its growth and development.

Chapter 3.

A Model of Cumulative Causation.

In chapter 2, it was suggested that there exists a close correspondence between hysteresis and the principle of cumulative causation associated with authors such as Kaldor. This correspondence, coupled with the potential of the Keynesian macroeconomic framework for capturing hysteresis, suggests that the work of Kaldor may represent a useful point of departure for our attempt to construct a hysteretic macroeconomic model. The aim of this chapter is to explore this possibility.

I Increasing returns and the Kaldorian Heritage.

Central to Kaldor's models of cumulative causation (see, for example, Kaldor 1970, 1972 and 1985) is the concept of increasing returns to scale. Increasing returns may be of either a static or dynamic nature. Static economies of scale result within the firm from rising output in the presence of fixed costs. This will tend to reduce the level of fixed costs per unit of output and hence increase the allocative efficiency of the firm. Static economies of scale which are external to the firm may also exist. Hence when a firm locates in an area where industry is geographically concentrated, it will benefit from a developed industrial infrastructure, the local presence of producers of intermediate goods and so forth.

However, for authors such as Kaldor, the most important types of economies of scale are dynamic. One way in which dynamic economies of scale occur, first emphasised by Adam Smith and subsequently by Young (1928), is as a result of increases in the division of labour (i.e., increased specialization) in response to increases in the size of the market.[1] Hence a firm in an imperfectly competitive industry may hesitate to specialize its production process if it must win market share from competitors in order to achieve the level of output that will fully utilize the fixed capital required for specialization.[2] However, if the market as a whole is expanding, then even if market shares remain constant, the firm may achieve a level of output that will fully utilize a more specialized technique of production (Scitovsky, 1956; Lamfalussy, 1961, 1963). Dynamic returns to scale therefore accrue as increases in output in response to growing demand facilitate specialization in the production

---

[1] Note that increases in the scale of production over time can also enhance efficiency if accumulation increases costs in a linear fashion, but increases capacity geometrically. Hence if a 10m x 10m x 10m storage facility costs only twice as much as a 5m x 5m x 5m storage facility, it will nevertheless lead to an eight fold increase in storage capacity, thus reducing storage costs per cubic unit by 75%.

[2] Note that an important assumption underlying this argument is that there exist indivisibilities in production which make the adoption of more capital intensive techniques of production a discontinuous process.

process which raises firms' productivity.

Similar processes may occur on an industry wide basis within an imperfectly competitive economy, if market and hence industry expansion facilitate increasing product differentiation, and the emergence of specialized firms at intermediate stages of production.

Dynamic economies of scale may also arise from "learning by doing." These type of scale economies derive from the notion that the more production occurs over a given period of time, the more is learnt about the production process through its repeated use, and so the more this production process is likely to be improved and refined. Arrow (1962) argues that learning is a product of experience - that it takes place as a consequence of engaging in production activities which are imperfectly understood. Arrow argues that learning gives rise to both improved knowledge of existing processes, and improvement in processes themselves i.e., technological progress. By increasing the amount of learning by doing, therefore, increasing the amount of production per time period can have a positive effect on the technology set of a firm/economy, raising the efficiency with which existing techniques of production are used, and providing impetus for the process of innovation.

Learning by doing may also give rise to increasing returns by virtue of the very nature of knowledge as a commodity. One important attribute of knowledge is that it is (at least partly) non-excludable. This can give rise to externalities in its creation if new knowledge contributes to a general "pool" of information from which all producers are able to draw (Grossman and Helpman, 1991). This externality is especially important since it should accrue evenly to *all* firms, as a result of the fact that the use of a specific piece of knowledge by one party does not preclude its use by others.

To see how the accumulation of knowledge may therefore give rise to increasing returns, suppose that we double current inputs into the production process in the context of a given state of knowledge. Because this knowledge is common property, and can be used simultaneously by any number of producers, it should therefore be possible to double output by a process of replication. If we now double the extent of knowledge, output should increase still further - so that doubling all inputs (including knowledge) more than doubles final output (Romer, 1990).

a. The existence of large scale increasing returns.

An important assumption underlying Kaldor's discussion of increasing returns is that they are large relative to the size of the economy.[3] Two important dynamic properties of modern capitalist economies suggest the existence of large scale

[3]. Hahn (1974) discusses the importance of this assumption.

increasing returns. The first of these, emphasised by Kaldor (1970) and Cornwall (1977), stems from the conception of capitalism as a "dual economy" in which a basic dichotomy exists between "land based" and "processing" activities. The distinguishing feature of the latter is their reliance on "produced means of production" (i.e., capital goods), which are intimately related to increasing returns.[4] Furthermore, the use of produced means of production implies that the "scarcity of resources" in processing activities cannot be thought of as being independent of the level of activity in the economy. What is chiefly important in processing activities is the dynamic propensity of the economy to create resources (i.e., to deepen and/or widen its stock of capital) rather than the static problem of resource allocation. The endogeneity of this resource creation process gives rise to the potential for continuous dynamic economies of scale related to increases in the quantity and quality of the capital stock - i.e., the potential for increasing returns which, at any given point in time, are large relative to the size of the economy.[5]

However, it is not only the ability of capitalism to

---

[4]. For example, capital goods embody technological progress, which is affected by the process of learning by doing.

[5]A similar point is made by Schumpeter (1942), who argues that there exists a potentially boundless scope for technological advance and hence increased productivity in production activities which rely on produced means of production.

manufacture an important source of increasing returns that accounts for their large scale nature. As Cornwall (1972) illustrates, this is also due to the inherently unbalanced nature of the growth of capitalist economies, or more precisely, the nature of growth as a process of structural transformation. According to Cornwall, growth can be characterized as a movement through a hierarchy of commodities with different goods becoming "necessities" (i.e., having high income elasticities of demand) and "luxuries" (low income elasticities of demand) at different levels of per capita income. In this way,

> All that really needs to be argued to support the economies of scale position is that the growth of productivity in some firm or industry is more responsive to demand conditions when the rate of growth of demand is high and rising. When the market for the good becomes relatively saturated and demand shifts to other industries, a slowing down in the rate of productivity and the exhaustion of scale economies can be expected. (But this) comes at a time when it is less consequential for the overall growth of productivity.
>
> (Cornwall, 1972, p.64)

In other words, as demand transfers through the commodity hierarchy towards new products, the new industries producing them are able to reap dynamic economies of scale. Hence the growth of productivity in these industries will be rising just as resources are being transferred towards them i.e., as their weight in the economy is becoming larger. Thus the process of growth as transformation - the continuous and overlapping rise and decline of industries, and the local (i.e., intra-industry) exploitation and then exhaustion of increasing returns - implies that in aggregate, economies of scale are always present and hence may be considered large relative to the size of the economy. Whilst economies of scale may not always be present in a given industry, they may yet continuously exist in the economy as a whole because movement through the commodity hierarchy will enable the continual exploitation of new sources of economies of scale.[6]

b. Dynamic increasing returns as a potential source of hysteresis.

The importance of dynamic increasing returns is their implication that the technical possibilities for economic activity in any period depend crucially on the nature and extent of economic activity in the past, and that the ways in which these opportunities are currently being exploited will,

---

[6]. A formal analysis of this argument is provided in appendix A.

Note that the arguments expressed above implicitly assume that there exist commodities *throughout* the commodity hierarchy which can be produced under conditions of dynamic increasing returns, and that each successive commodity can be produced with a higher average productivity than its predecessor. These assumptions, which will be retained throughout the following analysis, attach special importance to the role of manufacturing activities in economic growth (see also Kaldor, 1970 and Cornwall, 1977). For an analysis of the limit effects of an economy whose commodity hierarchy becomes increasingly dominated by the output of the service sector, see Sundrum (1990) chapter 3.

in turn, create opportunities for the future which previously did not exist. Hence as the economy expands and in so doing accumulates capital, experiences learning by doing, and moves through the commodity hierarchy, it will realise dynamic increasing returns which may subsequently facilitate the further expansion of its output. Under these conditions, some part of the "state of technology" in an economy is deeply endogenous. More importantly, growth in any period will be influenced by the extent to which growth has been realised in the past - that is, growth will become a path dependent process. In light of this, we can identify dynamic increasing returns as an important potential source of hysteresis.

II Modelling cumulative causation.

The question remains, however, as to precisely how the dynamics of an economy with large scale increasing returns will operate, and how they may be modelled.

An early attempt to address this issue was Young's (1928) combination of Adam Smith's insights into the division of labour with Say's Law. Young argues that, at a very aggregate level, economic activity can essentially be seen as a process which involves the exchange of goods for goods. Hence every increase in the supply of a commodity enlarges, at least potentially (a vital qualification), the market for other commodities. This leads Young to conclude that whilst Smith was correct to point out that the division of labour depends on the extent of the market, the extent of the market also depends on the division of labour, in the sense that rising output accruing from an increase in the latter will potentially increase the former. Hence Young argues that the division of labour depends on the division of labour - not just in a simple tautological sense, but because the division of labour and the extent of the market are involved in a "joint interaction" over time, or what may be defined as a process of cumulative causation.

However, this combination of Say's law and Smith's theorem does not in itself provide a complete dynamic theory. As Ricoy (1987) notes,

> ... there is no 'degree of freedom', no independent leading element in the system. In this sense, the analysis remains, as it were, 'hanging in the air.'
>
> (Ricoy, 1987, p.732.)

From a Keynesian perspective, what is missing is a link between the effects of a change in the level of production and the level of aggregate demand, in order to ensure that a given change in production ultimately stimulates rather than depresses production in the economy as a whole. One of the central contributions of Kaldor (1970, 1972, 1985) was to realise this, and show that the gap may be filled by Keynes' theory of effective demand.

In his 1985 model, Kaldor emphasises the importance of external trade as the leading source of aggregate demand. For Kaldor, the process of economic development and the accompanying division of labour imply a continuous subdivision of industries, and the growth of specialist firms at intermediate stages of production.[7] He argues that there is an obvious tendency for this process to become concentrated in and around geographical centres, since its success depends in part on such factors as the local presence of skilled manpower and the frequent transfer of unfinished goods between specialised firms. Hence if success in industrial development is at all uneven between different centres, then there will exist scope for mutually advantageous exchanges to take place - i.e., we would expect interregional trade to occur.

a. A formal model.

Following Cornwall (1977) and Thirlwall (1980), the relationship between interregional trade, aggregate demand, dynamic increasing returns and growth in a Kaldorian model of cumulative causation can be illustrated by the dynamic interaction of the following equations:

---

[7]. Note that this vision of production as a process of extending social co-operation is similar to Frankel's (1955) argument that the degree of interrelatedness between the components of the productive process (machines, firms, institutions etc.) increases as the economy grows. This idea is explored more fully in chapter 4.

$$\dot{q}_{jt} = r_t + \alpha_j \dot{Y}_{jt-1} \qquad [3.1]$$

$$\dot{p}_{jt} = \dot{w}_{jt} - \dot{q}_{jt} \qquad [3.2]$$

$$\dot{X}_{jt} = \beta_j (\dot{p}_{wt} - \dot{p}_{jt}) + \gamma_j \dot{Y}_{wt} \qquad [3.3]$$

$$\dot{Y}_{jt} = \lambda_j \dot{X}_{jt} \qquad [3.4]$$

where for any region j, $\dot{q}_j$ is the rate of growth of productivity, r represents exogenous influences on the rate of productivity growth, $\dot{Y}_j$ is the rate of growth of output, $\dot{p}_j$ and $\dot{w}_j$ are domestic price and wage inflation respectively, $\dot{X}_j$ is the rate of growth of exports, $\dot{p}_w$ and $\dot{Y}_w$ represent price inflation and output growth in the "rest of the world" respectively, and subscripts denote time periods.[8] The model is intended to be couched in terms of traded goods which are subject to dynamic increasing returns.

Equation [3.1] expresses productivity growth as a function of exogenous influences and the "Verdoorn component," $\alpha_j \dot{Y}_{jt-1}$.[9] The coefficient α_j essentially reflects the extent to which output growth facilitates the realisation of dynamic increasing returns, and hence productivity growth in region j.

---

[8] Although this model is usually presented without time subscripts (see, for example, Cornwall (1977) and Thirlwall (1980), the lag structure in equations [3.1] - [3.4] is imposed in order to ease exposition of the dynamic interaction of these equations.

[9] Verdoorn's Law states that there exists a strong positive relationship between the growth of manufacturing output and productivity growth. See Thirlwall (1983) for further discussion of this law.

Equation [3.2], meanwhile, follows from the mark up price equation:

$$p_t = \left(\frac{w_t}{q_t}\right).\tau$$

where p is price, w is the nominal wage, q is the level of labour productivity and τ = 1 + the (fixed) percentage mark up of prices over labour costs.

In equation [3.3], the coefficient γ_j is the income elasticity of region j's exports, whilst β_j represents their own price elasticity. β_j therefore captures the extent to which an increase in the price competitiveness of region j will expand the market for its exports. Equation [3.3] is derived from Thirlwall's (1980) export function:

$$X_{jt}=\left(\frac{p_{jt}}{e.p_{wt}}\right)^{\phi}.Y_{wt}^{\gamma}$$

where $-\phi = \beta_j$, and e represents the (fixed) exchange rate. Finally, equation [3.4] relates export growth to the subsequent growth of income. In keeping with Kaldor's emphasis on the demand side, it makes the Hicksian assumption that the growth of autonomous demand determines the long run rate of output growth (Thirlwall, 1980). In this case, exports are assumed, for simplicity, to be the only source of autonomous demand that is non-constant, in order to highlight Kaldor's particular emphasis on the importance of export led growth.

The structural model outlined above is of a simple form, and is by no means intended to be a complete structural specification of a capitalist macroeconomy. Notice, for example, that the model does not involve a balance of payments equilibrium condition. It assumes that trade surpluses and deficits can be accumulated indefinitely, overlooking the fact that this would require deficit regions to be able to attract a permanent net inflow of capital (Thirlwall, 1980).[10]

There are also problems specific to the formal model in equation [3.1] - [3.4] which question its suitability as a characterisation of Kaldor's (1970, 1985) arguments. For example, taking $\dot{Y}_{wt}$ as given, $\dot{X}_{jt}$ in equation [3.3] depends entirely on the inflation differential ($\dot{p}_w - \dot{p}_j$), making no allowance for non-price forms of competition to affect export growth. It is not clear that Kaldor considered price differentials to be the only, or even the most important factor in interregional competition.

More importantly, note that by substituting equations [3.3], [3.2] and then [3.1] into equation [3.4] and imposing

[10] Thirlwall (1980, pp. 270-272) shows how this particular problem can be overcome. By introducing an import function and balance of payments equilibrium condition, he derives an expression for the rate of growth in region j which is consistent with long run balance of payments equilibrium in this region.

steady state conditions, we arrive at the expression:

$$\dot{Y}_j = \frac{\lambda_j\beta_j(\dot{p}_w - \dot{w}_j + r) + \lambda_j\gamma_j\dot{Y}_w}{1 - \lambda_j\alpha_j\beta_j} \qquad [3.5]$$

From the analogous conditions to [3.1] and [3.2] for region w, we have (in the steady state):

$$\dot{q}_w = r + \alpha_w\dot{Y}_w \qquad [3.1a]$$

$$\dot{p}_w = \dot{w}_w - \dot{q}_w \qquad [3.2a]$$

$$\Rightarrow \quad \dot{p}_w = \dot{w}_w - r - \alpha_w\dot{Y}_w$$

Substituting into [3.5] we arrive at:

$$\dot{Y}_j = \frac{\lambda_j\beta_j(\dot{w}_w - r - \alpha_w\dot{Y}_w - \dot{w}_j + r) + \lambda_j\gamma_j\dot{Y}_w}{1 - \lambda_j\alpha_j\beta_j}$$

$$\Rightarrow \dot{Y}_j = \frac{\lambda_j\gamma_j\dot{Y}w - \lambda_j\alpha_w\beta_j\dot{Y}_w}{1 - \lambda_j\alpha_j\beta_j}$$

given that $\dot{w}_w = \dot{w}_j$.[11] Hence we can now write:

---

[11] The assumption of constant wage relativities (i.e., w_j/w_w constant, which implies that $\dot{w}_j = \dot{w}_w$) is a key feature of Kaldor's exposition of cumulative causation. It is discussed in more detail below.

$$(1 - \lambda_j\alpha_j\beta_j)\dot{Y}_j = \lambda_j(\gamma_j - \alpha_w\beta_j)\dot{Y}_w$$

$$\Rightarrow \frac{\dot{Y}_j}{\dot{Y}_w} = \frac{\lambda_j(\gamma_j - \alpha_w\beta_j}{1 - \lambda_j\alpha_j\beta_j} \qquad [3.6]$$

which suggests that in the model in equations [3.1] - [3.4], long run relative growth rates are determinate - that is, they are independent of the prior growth paths of regions j and w.[12] This seems quite unlike the result intended by Kaldor, who envisaged initial relative growth rates having a permanent effect on their subsequent values in a system without a determinate equilibrium (see, for example, Kaldor 1985, pp. 61-3).

However, recall from section IIIa of chapter 2 that even in a system where a unique, determinate equilibrium exists, long run outcomes may nevertheless be path dependent if the speed of adjustment towards this equilibrium is slow relative to the speed at which the data underlying it are changing. This possibility would seem especially likely in the current context, not least because the "data" in equation [3.6] may themselves vary with the growth paths $\dot{Y}_{jt}$ and $\dot{Y}_{wt}$. We will pursue this likelihood in extensive detail in chapters 4 and

---

[12] This result depends, of course, on the assumption that $|\lambda_j\alpha_j\beta_j| < 1$. If $|\lambda_j\alpha_j\beta_j| > 1$, the model will explode in response to an initial perturbation. If $|\lambda_j\alpha_j\beta_j| = 1$, no determinate equilibrium exists in [3.6].

5 below. At present, however, simply note that the coefficient γ_j, the income elasticity of demand of region j's exports, will depend on j's progress through the commodity hierarchy, which is itself a function of j's stage of development and hence prior growth path (Cornwall, 1977).

Making the "slow relative adjustment" assumption allows us to use the model in equations [3.1] - [3.4] to illustrate the dynamics of cumulative causation without, at present, making assumptions about the effects of growth on the "data" in equation [3.6]. Under this assumption, the equations can be treated as a recurring sequence in order to demonstrate their interaction in a process of cumulative causation.

b. The process of cumulative causation.

According to Kaldor, an important feature of the process of economic development is that differences in productivity growth are likely to emerge between regions. This claim may appear curious given the conventional notion that rational entrepreneurs, regardless of their location, will always adopt the least cost (i.e., most efficient) method of production currently available. However, in Kaldorian models, efficiency gains result from dynamic increasing returns, and may therefore accrue to regions unevenly as a result of the fact that economic development does not occur simultaneously in all regions. It is clear from equation [3.1] that an economy which experiences an initially high rate of output growth will also experience a high rate of productivity growth.

The consequences of this are clearly demonstrated by equations [3.2] - [3.4]. From equations [3.2] and [3.3], we can see that, other things being equal, a region with a high rate of productivity growth will experience a high rate of growth of exports. With $\dot{w}_{jt}$ given in equation [3.2], efficiency wages (the ratio of money wages to productivity), and hence unit costs of production will fall in an economy with a high rate of productivity growth. Kaldor argues that this will occur because whilst productivity growth varies with output growth due to increasing returns, wage relativities tend to remain constant between regions over time. Hence differences between regions in the rate of productivity growth will not be offset by equivalent differences in the rate of growth of money wages.

In this way, a region with a high rate of productivity growth should be able to exploit its low efficiency wages by realising a relatively low domestic rate of inflation. The region will thus gain a competitive advantage over other regions which, as indicated by equation [3.3], will enable it to win market share from them and so improve its rate of growth of exports. This, in turn, will raise the level of economic growth, as captured by equation [3.4].

However, in the presence of dynamic increasing returns, the process will not end here. Returning to equation [3.1], we can see that the growth of income in [3.4] will leads to the subsequent realisation of dynamic increasing returns. This will act so as to reinforce the initial productivity growth advantage of region j, further enhancing its trade position and hence the growth of its income and so on. In this way, the interaction of equations [3.1] - [3.4] describes a process of cumulative causation applied to interregional trade, which can be summarised by the causal chain:

$$\dot{X}_{jt} \rightarrow \dot{Y}_{jt} \rightarrow \dot{q}_{jt} \rightarrow \dot{X}_{jt+1} \rightarrow \text{etc.}$$

c. The consequences of cumulative causation - a two region example.

An important feature of the model developed above is that it involves a successful region winning market share from its competitors in interregional trade. This means that the growth of one region is not neutral with respect to its effects on the other regions with which it competes. This can be illustrated by considering an interregional trading system comprising just two competing regions, A and B.

Suppose, then, that $\dot{Y}_{Bt-1} > \dot{Y}_{At-1}$, due to the legacy of early industrialisation in region B.[13] Then from equation [3.1], we will have $\dot{q}_{Bt} > \dot{q}_{At}$, and hence $\dot{p}_{Bt} < \dot{p}_{At}$ in equation [3.2] assuming that $\dot{w}_{Bt} = \dot{w}_{At}$ (constant wage relativities).

[13] Note that the conclusions which follow from this assumption can also be motivated by assuming an initially favourable export shock in B (Cornwall, 1977) or an initially higher income elasticity of demand for region B's exports (Thirlwall, 1980).

Since this implies that $\dot{p}_{At} - \dot{p}_{Bt} > 0$, and conversely that $\dot{p}_{Bt} - \dot{p}_{At} < 0$, it is clear from equation [3.3] that B's competitive advantage will result in $\dot{X}_{Bt} > \dot{X}_{At}$. This means that $\dot{Y}_{Bt} > \dot{Y}_{At}$ in equation [3.4], and so $\dot{q}_{Bt+1} > \dot{q}_{At+1}$ in equation [3.1] and so on as the growth dynamics of the model become self reinforcing. These patterns of export, output and productivity growth can be summarised by the causal chains:

Region A: low $\dot{X}_{At} \rightarrow$ low $\dot{Y}_{At} \rightarrow$ low $\dot{q}_{At} \rightarrow$ low $\dot{X}_{At+1} \rightarrow$ etc.
Region B: high $\dot{X}_{Bt} \rightarrow$ high $\dot{Y}_{Bt} \rightarrow$ high $\dot{q}_{Bt} \rightarrow$ high $\dot{X}_{Bt+1} \rightarrow$ etc.

As a result of the initial differences in productivity growth, which are themselves explained by interregional differences in the historical timing of development, the joint interaction of $\dot{X}$, $\dot{Y}$ and $\dot{q}$ in region A gives rise to a vicious circle of cumulative causation, whilst region B experiences a virtuous circle of cumulative causation. The important point is that the effect of the initial disparity in trade performance is cumulative due to the presence of dynamic increasing returns. Hence region B's initial advantage in trade is translated into a long run advantage as it succeeds in continually growing at the expense of region A. Note in connection with this claim that since $\dot{X}_B > \dot{X}_A$ in all periods, region B's share of total trade between the two regions will be increasing over time in a cumulative fashion.[14]

Furthermore, suppose that $M_A = X_A$ and $M_B = X_B$ initially, where M_A and M_B represent the level of imports in regions A and B respectively. Then other things being equal, since the trade flows between the regions in all periods are such that $\dot{M}_A = \dot{X}_B > \dot{X}_A = \dot{M}_B$, the dynamics of the model will involve the cumulative improvement of region B's balance of trade at the expense of the cumulative worsening of region A's. Finally, note that although region A is growing positively, it is doing so at a slower rate than region B ($\dot{Y}_B > \dot{Y}_A$ in all periods). Since $Y_B > Y_A$ initially by virtue of B's early start to development, this will result in the *levels* of income in the

---

[14] To see this formally, note that

$$\dot{X}_B > \dot{X}_A$$

$$\Rightarrow \dot{X}_B > \dot{X}_T$$

where $X_T = X_A + X_B$. Hence

$$\dot{X}_B - \dot{X}_T > 0$$

$$\Rightarrow \left(\frac{X_B}{X_T}\right) > 0$$

$$\Rightarrow \frac{d(X_B/X_T)}{dt} > 0 \; since \; \frac{X_B}{X_T} > 0$$

two regions diverging over time.[15] Hence in a model of this nature, we would expect increasing inequalities to emerge over time between initially prosperous and initially poor regions as a result of interregional trade.

III A note on wage and price assumptions in the model of cumulative causation.

a. Flexible wages and competitiveness.

In order to motivate the process of cumulative causation in equations [3.1] - [3.4], a key assumption is that wage relativities between regions tend to remain constant over time. It may therefore appear that cumulative causation depends on a "special case" sticky wage assumption, and that its quantity dynamics will be offset if wages are flexible, and a vicious circle region experiences declining relative money wages.

Note, however, that the adjustment of relative wages necessary to restore the competitiveness of a vicious circle region would have to be instantaneous in order to completely

[15] The perpetual divergence between regional levels of exports, output and productivity that is implied by the process of cumulative causation suggests a potential limit to the applicability of this process in the long run. This problem, together with a proposed solution, is discussed extensively in chapters 4 and 5 below.

mitigate the quantity effects of cumulative causation.[16] To see this, we need only examine equations [3.2] and [3.4] as they appear for regions A and B in the two region example developed earlier. Hence write:

$$\dot{p}_{At} = \dot{w}_{At} - \dot{q}_{At} \qquad [3.2A]$$

$$\dot{p}_{Bt} = \dot{w}_{Bt} - \dot{q}_{Bt} \qquad [3.2B]$$

$$\dot{X}_{At} = \beta_A(\dot{p}_{Bt} - \dot{p}_{At}) + \gamma_j\dot{Y}_{Bt} \qquad [3.3A]$$

$$\dot{X}_{Bt} = \beta_B(\dot{p}_{At} - \dot{p}_{Bt}) + \gamma_j\dot{Y}_{At} \qquad [3.3B]$$

If, in any "round" of the process of cumulative causation, we have $\dot{q}_{Bt} - \dot{q}_{At} > \dot{w}_{Bt} - \dot{w}_{At}$ (i.e., if $\dot{w}_{At}$ does not adjust instantaneously to offset the productivity growth differential between the two regions) then we will have:

$$\dot{w}_{At} - \dot{q}_{At} > \dot{w}_{Bt} - \dot{q}_{Bt}$$

$$\Rightarrow \quad \dot{p}_{At} > \dot{p}_{Bt}$$

in equations [3.2A] and [3.2B], which, ceteris paribus,

---

[16] Such instantaneous adjustment could occur if, for example, the economy was regulated by an auctioneer, or if economic agents were forming strong form rational expectations on the basis of complete information in the context of a perfectly competitive labour market. As illustrated in Part I, however, no form of instantaneous adjustment endears itself to models based on the notion that adjustment is an intertemporal process. What the following analysis suggests is that without instantaneous relative wage (and in section IIIb, exchange rate) adjustment, wage (exchange rate) flexibility *per se* is not sufficient to offset the quantity dynamics of cumulative causation.

implies $\dot{X}_{Bt} > \dot{X}_{At}$ in equations [3.3A] and [3.3B].

If instantaneous relative wage adjustment does not occur, then, the competitive advantage realised in any round of the cumulative causation process by a region experiencing high productivity growth may be *reduced* by changes in relative money wages, but it will not be *entirely offset*. Hence the quantity dynamics of the model will continue to work in the manner described in section II.

However, regardless of the speed of wage adjustments, it is not obvious that we would expect a vicious circle region to experience declining relative money wages. Since region A's growth rate is positive, employment may be increasing in this region in spite of its *relative* decline. If wage adjustments do not occur unless the level of employment is falling, and hence a region is experiencing *absolute* decline, region A will not experience a downward pressure on money wages, and wage relativities may remain unchanged (Skott, 1985).

b. The exchange rate and competitiveness.

A second important assumption made by the model discussed earlier is that of a fixed exchange rate, e. If this assumption is relaxed, equation [3.3] can be rewritten as:

$$\dot{X}_{jt} = \beta_j(\dot{p}_{wt} + \dot{e}_t - \dot{p}_{jt}) + \gamma_j\dot{Y}_{wt} \qquad [3.7]$$

Is it the case, then, that changes in e may prevent a region from gaining and maintaining the type of competitive advantage necessary to generate cumulative causation?

In order to completely mitigate the quantity effects of cumulative causation, it would again be necessary for the exchange rate to adjust instantaneously to completely negate changes in relative competitiveness within any given round of the process of cumulative causation. To see this, we need only consider the workings of equation [3.7]. Suppose, then, that $\dot{q}_{jt} > \dot{q}_{wt}$, so that $\dot{p}_{wt} > \dot{p}_{jt}$ in equation [3.7]. Suppose further that $\dot{e}_t < 0$ - that is, that the exchange rate depreciates in favour of region w. If, in spite this, it is nevertheless the case that $\dot{p}_{wt} + \dot{e}_t > \dot{p}_{jt}$, there will still be a positive net effect on $\dot{X}_{jt}$ in equation [3.7]. If the exchange rate does not adjust instantaneously, then as in the case of sluggish wage adjustment, quantity effects will be realised and cumulative causation will continue as previously described.

The effects of a flexible exchange rate on the quantity dynamics of cumulative causation may also be impaired if the exchange rate has only a limited impact on relative price competitiveness, fails to adjust appropriately in response to changes in interregional trade flows, or simply fails to affect interregional trade flows. For example, note that in the model developed in section II, whilst the *producer* real wage in region j within any period is given by $\omega_j = (w_j/p_j)$, the *consumer* real wage in the same period can be written:

$$\omega_{cj} = \frac{w_j}{p_j^{\delta}.(e.p_w)^{\varepsilon}} \qquad [3.8]$$

where δ and ε are weights such that $0 < \delta < 1$, $0 < \varepsilon < 1$ and $\delta + \varepsilon = 1$. It follows from equation [3.8] that the rate of growth of the consumer real wage is given by:

$$\dot{\omega}_{cj} = \dot{w}_j - \delta\dot{p}_j - \varepsilon(\dot{e} + \dot{p}_w) \qquad [3.9]$$

Equation [3.9] suggests that consumer real wage growth varies inversely with fluctuations in the exchange rate, e. Suppose, then, that during some period we have $\dot{e} > 0$ - i.e., that the exchange rate depreciates in favour of region j.[17] Other things being equal, this will reduce the rate of growth of the consumer real wage during this period. However, if region j now experiences real wage resistance - that is, if $\dot{w}_j$ is bid up in an attempt to offset the decline in $\dot{\omega}_{cj}$ - then this will reduce the effects on competitiveness of the original devaluation (Skott, 1985). Equation [3.2] will now take the form:

$$\dot{p}_j + \varepsilon\dot{e} = \dot{w}_j - \dot{q}_j \qquad [3.2a]$$

[17] Note that e is defined throughout the present analysis as the number of units of region j's currency per unit of region w's currency. Hence $\dot{e} > 0$ implies a depreciation in the value of region j's currency.

where $\varepsilon\dot{e}$ represents the change in $\dot{w}_j$ necessary to offset the effects of exchange devaluation on the consumer real wage in region j. This in turn implies that equation [3.7] can be re-written as:

$$\dot{X}_j = \beta_j(\dot{p}_w + \dot{e} - \dot{p}_j - \varepsilon\dot{e}) + \gamma_j\dot{Y}_w$$
$$\Rightarrow \quad \dot{X}_j = \beta_j(\dot{p}_w - \dot{p}_j + \dot{e}(1-\varepsilon)) + \gamma_j\dot{Y}_w \qquad [3.7a]$$

Since $(1 - \varepsilon) > 0$ by assumption, this illustrates that real wage resistance will reduce the effects of a given exchange depreciation on relative competitiveness and hence export growth in region j.

Aside from real wage resistance, the exchange rate may be deliberately manipulated in response to political economy considerations, or its movements may be dominated by speculative capital flows rather than trade in goods and services (Dornbusch, 1987). If the exchange rate consequently fails to adjust appropriately in response to changes in interregional trade flows, it will not systematically offset the quantity dynamics of cumulative causation.

Finally, changes in the exchange rate simply may not have large effects on trade flows between regions in the short run. If entry into foreign markets involves sunk costs, there will exist a range over which the exchange rate may fluctuate without encouraging exit from a particular regional market (Dixit, 1989; Krugman, 1989). Current incumbents may practice a policy of "pricing to the market" in order to offset currency fluctuations and so avoid the loss of their hard won (through expenditure on sunk costs) market share. If trade flows are thus invariant to exchange rate movements, changes in currency values will have little short run effect on the quantity dynamics of cumulative causation.

c. Summary

Without formally specifying an appropriate theory of wage and exchange rate dynamics under conditions of hysteresis, the preceding comments illustrate the possibility that relaxing the wage and price assumptions implicit in the model developed in section II need not systematically negate the implications of this model. This suggests that the assumptions of a fixed exchange rate and constant wage relativities can be retained as simplifying abstractions throughout the remainder of our analysis in Part II.

IV A long run analog of the General Theory.

In chapter 2, it was suggested that hysteresis may provide a way of generalising traditional short run Keynesian macroeconomics into a long run form, enabling us to conceive the economy as a "moving short run under employment equilibrium." How far towards this goal does the model of cumulative causation developed in this chapter take us?

Recall the relationship between output and export growth in equation [3.4]:

$$\dot{Y}_{jt} = \lambda_j \dot{X}_{jt} \qquad [3.4]$$

This equation can be broadly interpreted as capturing the multiplier effect of increased net exports on income:

$$\Delta Y = \Delta X + k.\Delta X + k^2.\Delta X + \ldots \quad , \qquad 0 < k < 1$$

$$\Rightarrow \qquad \Delta Y = (1/1-k).\Delta X \qquad [3.10]$$

The expression in equation [3.10], which is analogous to Harrod's trade multiplier, can be thought of as capturing the adjustment between short run under employment equilibria in a conventional Keynesian macroeconomic model. Note, however, that in light of equations [3.1] - [3.3], the act of "getting intc" a new short run under employment equilibrium described in equation [3.10] will subsequently affect the productivity of the economy due to the Verdoorn Law. This will in turn affect the economy's competitiveness, inducing further changes in its exports and hence output. In other words, the process of "getting into" the equilibrium described in equation [3.10] will endogenously induce further changes in the configuration of the short run under employment equilibrium itself. Equations [3.1] - [3.4] simply capture this process in continuous time. Essentially, then, they model a moving short run under employment equilibrium, and hence constitute a long run theory of effective demand.

The Keynesian features of the model developed in this chapter are by no means limited to aspects of pedagogy. In equations [3.1] - [3.4], as in Keynes (1936), the moving short run under employment equilibrium need display no automatic tendencies towards an optimal outcome associated with the full employment of resources (i.e., a full employment growth path). Instead, macroeconomic performance will depend crucially on the interplay of sequential patterns of activity. Indeed, if combined in the form of a vicious circle, these sequential patterns may lead an economy to fluctuate for long periods of time at levels of activity below full employment.

V Summary.

The model developed in this chapter and summarised in equations [3.1] - [3.4] serves as a useful tool to illustrate the dynamics of cumulative causation. As intimated earlier, it by no means represents a complete specification of the dynamics of a capitalist economy. Even within the context of models of cumulative causation, it suffers the obvious drawback of emphasising price competitiveness to the exclusion of non-price competition, which is an increasingly important feature of international product markets.

However, as illustrated in section III, the model is not unduly compromised by its simplicity and consequent "neglect" of factors such as wage and exchange rate flexibility. Furthermore, section IV suggests that it invites comparisons to the traditional short run Keynesian macroeconomic model, of which it is essentially a long run analog. As illustrated in section IV, the model allows us to begin to conceive the economy as a moving short run under employment equilibrium. In view of the nature of this conception (described in chapter 2), this in turn suggests that the model provides basic insights into the dynamics of a hysteretic economy. The close correspondence between cumulative causation and hysteresis enables us to begin to grasp the potential consequences of the latter for macroeconomic performance. Indeed, it seems possible to interpret Kaldor's analysis of cumulative causation as nothing less than a preliminary attempt to understand the workings of the economy under conditions of hysteresis.

Chapter 4.

The Supply Side and Macroeconomic Performance I: Technological Evolution.

In chapter 3, it was suggested that Kaldor's model of cumulative causation can be interpreted as a first step towards understanding macroeconomic dynamics under conditions of hysteresis. Recall, however, that our characterisation of this model in equations [3.1] - [3.4] yielded the determinate long run relative growth rates:

$$\frac{\dot{Y}_j}{\dot{Y}_w} = \frac{\lambda_j(\gamma_j - \alpha_w\beta_j)}{1 - \lambda_j\alpha_j\beta_j} \qquad [3.6]$$

In chapter 3, it was assumed that adjustment towards this determinate long run outcome is slow relative to the speed at which the data underlying it are changing. This assumption was justified by the suggestion that these data may themselves ultimately be sensitive to the growth path of a region. In this and the following chapter, we will explicitly study mechanisms which may give rise to changes in the coefficients α and β, which play a key role in defining the outcome in equation [3.6]. This will be shown to transform the model of cumulative causation developed in chapter 3 into a more generally hysteretic model of long run growth and development.

I Problems with the Kaldorian model of cumulative causation.

a. Initial conditions and cumulative success/failure.

There exists an important potential caveat in Kaldor's concept of cumulative causation as discussed in chapter 3. In this model, it appears that success breeds success and failure begets failure indefinitely. The initial conditions that exist in a region relative to those of its competitors determine that region's long run competitive fate. An initially fast growing economy will permanently experience fast growth, whilst its less successful trading partners will endure slower growth in perpetuity. In sum, once the initial success or failure of a region is known, so too is the rest of its history.

The extent of this problem should not, however, be overstated. Consider the case of a virtuous circle, which can be summarised by the causal chain:

$$\text{high } \dot{X}_t \rightarrow \text{high } \dot{Y}_t \rightarrow \text{high } \underbrace{\dot{q}_t \rightarrow}_{a} \text{high } \dot{X}_{t+1} \rightarrow \text{etc.}$$

In any "round" of the process of cumulative causation, the dynamics of a virtuous circle depend on sufficient demand (i.e., export growth) being forthcoming in response to the realization of scale economies to support the next "round" of cumulative causation. In other words,

> whether the process keeps its momentum.....or gets stopped (and probably reversed) depends on the 'next round' of demand, on the response of demand to the inducement to further growth provided by the rise in efficiency.
>
> (Ricoy, 1987, p.733.)

Hence there exists a potential weak link in the causal chain of a virtuous circle, if the response of exports to a productivity stimulus is uncertain. This weak link is marked at point a in the virtuous circle above. In any "round" of this causal chain, if insufficient demand is forthcoming to give rise to new efficiency gains, the cumulative process on which a virtuous circle is based will break down.

In order to explicitly motivate this possibility, consider the augmented formal model of cumulative causation below:

$$\dot{q}_{jt} = r_t + \alpha_j \dot{Y}_{jt-1} \qquad [3.1]$$

$$\dot{p}_{jt} = \dot{w}_{jt} - \dot{q}_{jt} \qquad [3.2]$$

$$\dot{X}_{jt} = \beta_j (\dot{p}_{wt} - \dot{p}_{jt}) + \gamma_j \dot{Y}_{wt} \qquad [3.3]$$

$$\dot{Y}_{jt} = \lambda_j \dot{X}_{jt} \qquad [3.4]$$

$$\beta_{jt} = f(\varepsilon_t) \quad , \quad \varepsilon_t \sim (0, \sigma_\varepsilon^2) \qquad [4.1]$$

where equations [3.1] - [3.4] simply restate the model of chapter 3, and equation [4.1] treats the coefficient β_j as a function of the stochastic variable ε. One possible motivation for this latter equation is if competition for market share involves non-price as well as price variables. For example, suppose that consumers view goods as imperfect substitutes and hence develop brand loyalties which may vary over time in response to strategic advertising, marketing etc. on the part of firms. One possible consequence of this is that the extent to which price advantages afforded by dynamic increasing returns will subsequently generate increased export sales will not be fixed over time - i.e., β_j may vary over time.

It is clear from equation [3.3] that the response of $\dot{x}_{jt}$ to $(\dot{p}_{wt} - \dot{p}_{jt})$ depends on the size of β_j in any round of the process of cumulative causation. If $d\beta_{jt}/dt < 0$ in any given round, then as a result of equation [4.1], price advantages arising from the efficiency gains induced by previous high growth in equation [3.1] may not generate sufficient new demand in equation [3.3] to perpetuate high growth in equation [3.4]. Notice that even if the reduction in β_j is temporary, the effect will be akin to a once over reduction in $\dot{Y}_j$ in equations [3.1] - [3.4]. In terms of the cumulative interaction of these equations, this will shift region j onto a permanently lower growth path. Under the conditions described in equations [3.1] - [3.4] and equation [4.1], then, the long run relative growth rate of a region will not be determined entirely by its initial success or failure.

b. Growth and endogenously induced supply side changes.

One possible objection to the preceding argument is that it relies on unexplained changes in competitive behaviour to generate the breakdown of a virtuous circle. This is not unlike the reliance of traditional equilibrium models on external "shocks" to explain changes in macroeconomic performance. The changes in competitive behaviour postulated above have not been shown to arise endogenously in the course of growth and development. Hence the weak link on the demand side of the model does not arise from *within* the process of cumulative causation itself. However, there are a number of ways in which the dynamics of cumulative causation may give rise to endogenously induced supply side changes in the course of growth. These induced supply side changes suggest the possibility of endogenously induced weak links on the supply side of the model developed in chapter 3, if the efficiency gains arising from a high rate of growth (i.e., the ability of the economy to realise dynamic increasing returns) are insufficient to perpetuate a virtuous circle.

To see this, consider the augmented model of cumulative causation below:

$$\dot{q}_{jt} = r_t + \alpha_{jt}\dot{Y}_{jt-1} \quad [3.1]$$

$$\dot{p}_{jt} = \dot{w}_{jt} - \dot{q}_{jt} \quad [3.2]$$

$$\dot{X}_{jt} = \beta_j(\dot{p}_{wt} - \dot{p}_{jt}) + \gamma_j\dot{Y}_{wt} \quad [3.3]$$

$$\dot{Y}_{jt} = \lambda_j\dot{X}_{jt} \quad [3.4]$$

$$\alpha_{jt} = f_j(\dot{Y}_{j0}, \ldots, \dot{Y}_{jt-1})$$

$$f_{ji}' \neq 0 \text{ some } i = 0, \ldots, t-1$$

$$\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} \neq 0 \quad [4.2]$$

where equations [3.1] - [3.4] again restate the model of chapter 3. This time equation [4.2] follows the methodology outlined in section IIId of chapter 2 for introducing hysteresis into linear models.[1] It characterises α_j as deeply endogenous and subject to adjustment asymmetries. That is, α_j is taken to vary with the prior growth path of region j, and in such a way that these variations do not "cancel out" over time. Hence the long run value of α_j depends on the prior growth path of region j.

In chapter 3, the parameter α_j was interpreted as indicating the extent to which output growth leads to the realisation of dynamic increasing returns, and hence productivity growth. The rationale underlying the treatment of

[1] Note that the function f is also assigned the subscript j. This is intended to capture the possibility that the precise evolution of α in response to a region's prior growth path may vary even between regions with like growth experiences.

α_j in equation [4.2] is that α can be thought of as depending on the technological and institutional context in which growth is taking place. More specifically, α depends on the ability of firms to undergo technological change (to the extent that capital embodies technological progress and hence accumulation is vital for the realisation of dynamic increasing returns) and the extent to which the institutional structures of an economy are conducive to increasing the efficiency of production.

Both technological change and institutions are conventionally regarded as exogenous in economic analysis. However, if they are in fact deeply endogenous to the process of growth and development - that is, if they depend on the historical growth path of the economy - then the coefficient α in equation [3.1] must in turn be thought of as deeply endogenous, as in equation [4.2]. What this implies is that the realisation of dynamic increasing returns in [3.1], and hence the maintenance of a virtuous circle depends on the technological and/or institutional evolution of the economy over time. If an economy suffers inefficient technological and/or institutional evolution - that is, if its technological and/or institutional structures cease to be functional to the maintenance of high growth dynamics - a virtuous circle may endogenously break down. In terms of equation [4.2], initially higher growth may cause $\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} < 0$, so that $d\alpha_{jt}/dt < 0$.

This will reduce the efficiency gains arising from a given rate of growth in equation [3.1], hence reducing the relative inflation advantage and so the rate of export growth in equations [3.2] and [3.3]. These developments will ultimately lower the rate of growth realised in equation [3.4]. In terms of the cumulative interaction of equations [3.1] - [3.4] which follows thereafter, the effects of this will again be to permanently shift region j onto a lower relative growth path. Equation [4.2] therefore captures the possibility of an endogenously induced weak link on the supply side of the cumulative causation process, illustrated at the point b in the causal chain below:

$$\text{high } \dot{X}_t \rightarrow \text{high } \dot{Y}_t \underbrace{\rightarrow}_{b} \text{high } \dot{q}_t \rightarrow \text{high } \dot{X}_{t+1} \rightarrow \text{etc.}$$

The question remains, however, as to how technological and institutional evolution can be modelled in order to illustrate the nature and consequences of this endogenous weak link. In what follows below, we will analyse the nature of technological change, returning to the topic of institutional evolution in chapter 5. In analysing both technological and institutional evolution, we will make use of similar concepts which indicate potential asymmetries in the adjustment of technology and institutions in the course of their evolution. Particular use will be made of the concept of lock in discussed in chapter 2. In each case, the treatment of α as a deeply endogenous variable subject to adjustment asymmetries will be shown to extend the Kaldorian model of cumulative causation into a more generally hysteretic representation of the growth process.

II Technological interrelatedness and the problem of lock in.

In chapter 2, it was suggested that the concept of lock in is compatible with hysteresis, and that it may provide important insights into the workings of a hysteretic economy. One way in which this is so is if lock in affects the technological evolution of an economy. Lock in may have an adverse effect on the rate of productivity growth in an economy, by limiting its capacity to undergo technological change and hence to realise dynamic increasing returns associated with technological change. The technological evolution of an economy may thus become inefficient - it will cease to be functional to the maintenance of high growth dynamics, with potentially adverse consequences for the region's relative rate of a growth.

a. Technological interrelatedness and lock in.

Frankel (1955) argues that given the historical fact that regions do not all industrialize simultaneously, older economies may come to suffer a disadvantage relative to newer economies with respect to their ability to assimilate new and more modern techniques of production. Hence productivity growth will wane in established industrial economies, as plant and equipment age and become obsolete.

This argument is based on the central idea that as an economy industrializes and develops, interconnections arise among its technological components - machines, plant, raw material supplies, transport network, etc..[2] For example, plant specifications, such as size or the strength of flooring, may be deliberately contrived to accommodate machinery of an existing vintage. Alternatively, haulage facilities may be designed to complement existing features of the infrastructure, such as the weight capacity or the height of bridges, the width of roads, or the gauge of railway tracks. This growing *interrelatedness* in the process of production may limit the ability of more developed economies to undergo technological change, locking them in to practices inherited as a legacy of the past. As was discussed in chapter 2, there are two dimensions to interrelatedness that can give rise to lock in - interrelatedness *within* and interrelatedness *between* individual decision making units.

---

[2]. Frankel also argues that similar interconnections will likely develop between an economy's institutions. The possibility of lock in to interrelated institutions is explored in chapter 5.

b. Technological interrelatedness within decision making units.

Technological interrelatedness within individual decision making units - or what may be referred to as intra firm technological interrelatedness - occurs when interconnections arise between the components of a firm's production process. Under these circumstances, replacing an individual component of the production process may require complementary changes in other components with which it is interrelated, and this may hinder technological change by making it prohibitively costly. To see this, consider an example involving two firms, one of which is a new entrant to the market, and the other of which is an established, incumbent firm. Assume that these firms have identical time horizons, and identical rates of time preference.

Consider now a technological innovation - say, the development of a new machine - which, when combined with complementary components, gives rise to a technique of production which is more efficient than that currently employed by the incumbent firm. How will the two firms respond to this technological advance? For the new entrant with no previous production history, the choice between the two techniques is unambiguous since:

$$\int_0^T (R_{nt} - C_{nt}) . e^{-rt}\, dt = \pi_n > \pi_o = \int_0^T (R_{ot} - C_{ot}) . e^{-rt}\, dt$$

by assumption, due to the greater efficiency of the new technique, where R represents revenue, C capital and variable costs, π denotes profits, T is the time horizon of the firm and r its rate of time preference, and o and n denote the old and new techniques respectively. However, for the incumbent firm, the relative profitability of the two techniques is less definite, since:

$$\int_0^T (R_{nt} - C_{nt}) . e^{-rt}\, dt - S = \pi_n \quad \pi_o = \int_0^T (R_{ot} - C_{ot}) . e^{-rt}\, dt$$

Here, S represents costs associated solely with the act of *changing* techniques, and hence abandoning elements of a *previously existing* technique. Because they correspond to past production decisions, these costs are unique to incumbents in the market - they do not influence the choice between techniques of new entrants. Costs of this nature fall into two main categories. First, there exist direct costs of abandonment. These include severance payments to redundant employees, and expenditures associated with the scrapping of plant and equipment. Essentially, these are costs which negate the standard assumption of "free disposal" of inputs into the production process.

Second, there exist sunk costs of production. As defined in chapter 2, a sunk cost is any expenditure that cannot be recovered if the action that led to its being incurred is subsequently reversed. Sunk costs of production are past outlays that can only be recovered from future revenues. An example of a sunk cost is the depreciation of capital values that may occur due to asymmetric information in the market for used plant and equipment (Akerlof, 1970). This accrues as a sunk cost since it ensures that the original purchase value of plant and equipment (net of an appropriate discount for its physical depreciation) cannot be recovered through its subsequent resale. If a firm abandons its existing technique of production, the sunk costs associated with this existing technique will "carry over," to be recovered from future revenues accruing to the new technique of production.

The important thing to note is that the greater the degree of technological interrelatedness, the more changes to its existing production process the incumbent firm will have to make in order to adopt the new technology. Hence the higher will be new fixed capital costs (a component of C_{nt}), the higher will be the costs, S, associated with the act of changing the production process, and so the lower will be π_n. Other things being equal, then, the greater the degree of technological interrelatedness, the less likely is the incumbent firm to adopt the new technology.

What this demonstrates is that with interrelatedness in the production process, technological change ceases to be the process of change at the margin that it is considered to be in traditional equilibrium economics. With technological interrelatedness, it may be that the relevant comparison between the profitability of old and new methods of production is not on, say, a machine for machine basis, since existing plant may need to be overhauled if a new machine is to be utilized. In this way, the best technique available to a new industrial region (with a low degree of technological interrelatedness) may not be profitable in an established industrial region. The latter will thus become locked in to certain techniques of production, inherited as a legacy of the past.

A final point, with respect to the time horizon over which lock in can occur, remains to be made. It may be argued that the problem outlined above should only be temporary, simply calling for a delay in replacement in the established region until existing plant wears out.[3] However, because investment is a continuous process, in the sense that different elements of the capital stock need to be replaced at different points in time, it need never be the case that existing plant will simultaneously wear out, and thus require

[3] In the example used above, the investment decision for the incumbent firm would then be akin to that of the new entrant, and the new technology would unambiguously be adopted.

wholesale replacement. The sequential nature of replacement investment may make it rational for an established region to continually replace elements of its original capital stock, which will entail continued use of an obsolete method of production. Hence the region may remain locked in to existing methods even in the long run.[4]

c. Technological interrelatedness between decision making units.

Technological interrelatedness can also occur between individual decision making units. This inter firm technological interrelatedness arises when there exist interconnections between the production processes of different firms. Such interconnections are likely to proliferate in the presence of uniform technical standards, which exist throughout an industry/economy regardless of how much more or less well suited they are to the production plans of any particular, individual firm. Uniform technical standards are pervasive in advanced industrial economies, and are often efficiency enhancing.[5] At the same time, however, they create circumstances where changing some element of the production process of one firm would require complementary changes in the

[4] In this case, of course, the relative efficiency and competitiveness of the region will be reduced. These issues are taken up in sections III and IV below.

[5] For example, if firms know that computer software is "IBM compatible," this may reduce the transactions costs incurred in the process of purchasing such software.

production processes of other firms, with which the former is interrelated by virtue of the technical standard. This may hinder technological change by making the size of the individual decision making unit small relative to the size of the "technological unit" that requires change. For example, as noted in chapter 2, an individual firm may be constrained from upgrading its rolling stock if this would require complementary changes to the uniform standard gauge of railway tracks, which are in turn owned by different firms, and which are consequently beyond the range of its control. If this type of problem prevails, we may once again observe the phenomenon of lock in to existing methods of production.

III Technological interrelatedness, lock in and productivity growth.

Frankel's argument, that technological change may be impeded by technological interrelatedness, is summarised in Figure 1 below, which shows the speed of diffusion of some new method of production in two regions A and B, where region A can be thought of as an initially less developed region suffering a lower degree of technological interrelatedness.

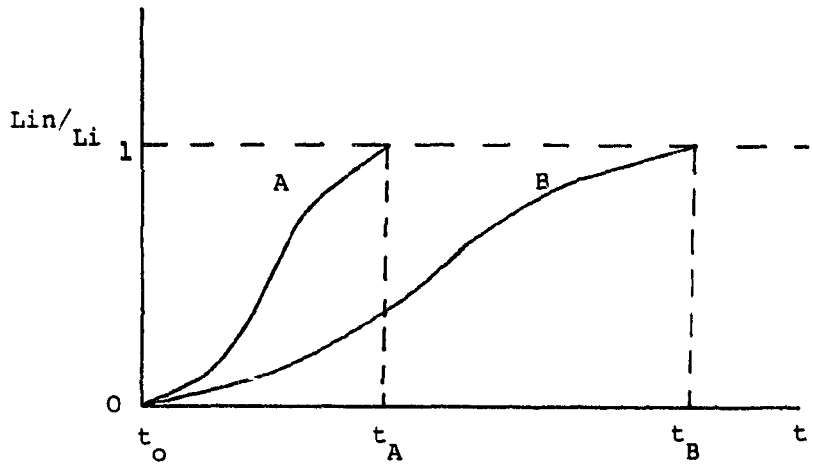

Figure 1. Rates of diffusion of a new technique of production in two regions.

Source: Cornwall (1977) p.115, with amendments.

The vertical axis in Figure 1 measures the percentage of the labour force in each region i employed in using the new technique, n. It captures the extent to which resources are being devoted to the new technique and hence, over time, the rate at which the latter is being adopted.

Following Cornwall (1977), a logistic pattern of diffusion of the new technique is assumed for each economy.[6] Following Frankel, the logistic S-curve is more "squeezed" for region A than for region B, indicating a faster rate of diffusion in the former region due to an initially lower degree of technological interrelatedness.

---

[6]. See Cornwall (1977) p.115-6 for the rationale behind this assumption.

What are the implications of Figure 1 for the rate of productivity growth in regions A and B? Consider the expression:

$$q_i = \lambda_{in}.q_{in} + \lambda_{io}.q_{io} \ , \qquad i = A,B \qquad [4.3]$$

where for any region i = A,B, q_i is the aggregate level of labour productivity, q_{in} and q_{io} are the levels of labour productivity corresponding to the new (n) and the old (o) techniques of production respectively, $\lambda_{in} = L_{in}/L_i$ is the share of labour resources devoted to the new technique, and $\lambda_{io} = L_{io}/L_i$ is the share of labour resources devoted to the old technique. Assume for simplicity that $q_{Bn} = q_{An} = q_n > q_{Bo} = q_{Ao} = q_o$. Then it is clear from equation [1] that the higher is the labour share of the new technique n, the more weight is attached to $q_n > q_o$, and so the higher will be q_i. As a result, other things being equal, the faster λ_{in} is increasing in region i, the faster will be the rate of increase of q_i. Hence if $d\lambda_{An}/dt > d\lambda_{Bn}/dt$ - that is, if the rate at which labour is being transferred towards the new technique is higher in region A than in region B as in Figure 1, then we will have $\dot{q}_A > \dot{q}_B$. As a result of its faster adoption of a new, higher productivity technique of production, region A can gain a productivity growth advantage over region B.[7]

[7] This result is illustrated in Appendix B, which following Sundrum (1990), simulates the dynamic analog of equation [4.3] to illustrate the implications for aggregate

IV Technological evolution and its implications for cumulative causation.

a. Technological interrelatedness, lock in and "growth in one region."

What the preceding section demonstrates is that the rate of productivity growth in a region depends in part on the rate at which it is able to adopt new, high productivity, techniques of production. This rate of adoption is in turn influenced by the degree of technological interrelatedness experienced by the economy, which is a function of its level of development.[8] As a region grows, then, increases in its degree of technological interrelatedness may reduce the opportunities for the exploitation of dynamic economies of scale related to technological change. This will have an adverse effect on the rate of productivity growth, which may in turn impair the growth dynamics of cumulative causation. It appears that we have identified an endogenous supply side mechanism that may cause the dynamics of a virtuous circle to break down.

---

productivity growth of the rate at which labour is being transferred between techniques of production with unequal levels and rates of growth of productivity.

[8] Recall that technological interrelatedness depends on interconnections arising between components of the production process. These interconnections can be expected to multiply as the economy develops and the division of labour increases - i.e., as production becomes more specialised and technically advanced. See Frankel (1955).

In order to demonstrate this argument formally, consider again the following augmented model of cumulative causation, in which α_j is treated as deeply endogenous:

$$\dot{q}_{jt} = r_t + \alpha_{jt}\dot{Y}_{jt-1} \qquad [3.1]$$

$$\dot{p}_{jt} = \dot{w}_{jt} - \dot{q}_{jt} \qquad [3.2]$$

$$\dot{X}_{jt} = \beta_j(\dot{p}_{wt} - \dot{p}_{jt}) + \gamma_j\dot{Y}_{wt} \qquad [3.3]$$

$$\dot{Y}_{jt} = \lambda_j\dot{X}_{jt} \qquad [3.4]$$

$$\alpha_{jt} = f_j(\dot{Y}_{j0}, \ \ldots, \ \dot{Y}_{jt-1})$$

$$f_{ji}' \neq 0 \text{ some } i = 0, \ldots, t-1$$

$$\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} \neq 0 \qquad [4.2]$$

The key relation in this augmented model is equation [4.2]. This equation captures the sensitivity of α_j - the ability of region j to realise dynamic increasing returns from a given rate of output growth - to the prior growth path of region j. The question is whether a higher prior growth path will enhance or reduce the long run value of α_j (i.e., whether $\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} \gtrless 0$ in equation [4.2]). On one hand, faster economic growth may increase opportunities for the realisation of dynamic scale economies ($f_{ji}' > 0$ for some i in equation [4.2]), as activities associated with learning and innovation become more specialised at higher levels of development. For example, separate research and development departments may be set up within firms. This will result in a higher proportional change in q_j in response to a given proportional change in Y_j - i.e., an increase in α_j.

However, what section II and III above illustrate is that as an economy grows, it may begin to experience technological interrelatedness and hence lock in to a prior technological structure. If faster economic growth increases the degree of interrelatedness in an economy, this will eventually impede the process of technological change and reduce the ability of the economy to realise dynamic increasing returns associated with technological change ($f_{ji}' < 0$ for some i in equation [4.2]).As a result, the economy will experience a decline in the value of α_j.

Suppose, then, that this second effect dominates, so that faster growth in a region such as B in chapter 3 leads to $\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} < 0$ in equation [4.2], and the value of α_B falls. Other things being equal, the effect of this will be to reduce the rate of productivity growth realised in equation [3.1]. This will, in turn, reduce region B's relative inflation advantage and hence rate of growth of exports in equations [3.2] and [3.3], and so ultimately lower the rate of output growth in equation [3.4]. In other words, region B will experience a climacteric - a deceleration of its trend rate of growth. In terms of the causal chain:

$$\text{high } \dot{X}_t \rightarrow \text{high } \dot{Y}_t \underbrace{\rightarrow}_{b} \text{high } \dot{q}_t \rightarrow \text{high } \dot{X}_{t+1} \rightarrow \text{etc.}$$

we have identified a weak link at point b, on the supply side of the model. This illustrates the possibility of an economy experiencing inefficient technological evolution in the course of growth, and hence the possibility that high growth dynamics may ultimately sow the seeds of their own long run destruction.

b. Technological interrelatedness, lock in and relative growth performance.

Suppose now that region A, B's trading competitor in the two region model developed in chapter 3, has experienced relatively more efficient technological evolution. Specifically, suppose that A has encountered a lower degree of technological interrelatedness owing to its lower relative growth rate. As illustrated in Figure 1, Region A may thus be able to gain a relative productivity growth advantage by virtue of its faster assimilation of new techniques of production. In terms of equation [3.1], if A's relatively more efficient technological evolution implies that $\alpha_{At} > \alpha_{Bt}$, then we may observe:

$$\dot{q}_{At} = r_t + \alpha_{At}\dot{Y}_{At-1} > r_t + \alpha_{Bt}\dot{Y}_{Bt-1} = \dot{q}_{Bt}$$

If this occurs, region B's growth performance will not only be slowed - it may also be relatively reversed. If region A is able to use its productivity growth advantage to win market share from B, so that $\dot{X}_{At} > \dot{X}_{Bt}$ in equation [3.3], the relatively faster rate of output growth that A realises in equation [3.4] will subsequently enable it to reap dynamic scale economies which will, in turn, reinforce the initial disparity in the relative efficiency of the two regions. As its export, output and productivity growth become virtuously cumulative, region A will take over as the export, output and productivity growth leader, whilst the dynamic macroeconomic performance of region B will enter a cumulative relative decline. The long run competitive fate of the two regions will now be summarised by the following causal chains:

Region B: low $\dot{X}_{Bt} \rightarrow$ low $\dot{Y}_{Bt} \rightarrow$ low $\dot{q}_{Bt} \rightarrow$ low $\dot{X}_{Bt+1} \rightarrow$ etc.

Region A: high $\dot{X}_{At} \rightarrow$ high $\dot{Y}_{At} \rightarrow$ high $\dot{q}_{At} \rightarrow$ high $\dot{X}_{At+1} \rightarrow$ etc.

This "dynamic leapfrogging" may be accompanied by a second effect. Since $\dot{Y}_A > \dot{Y}_B$ in all future periods, region A may begin catching up with region B's *level* of income.[9] Indeed, if A can maintain its competitive advantage and avoid a breakdown of its own virtuous circle, it may even *overtake* B's income level. In this case, we will observe a change in

---

[9] Recall that $Y_A < Y_B$ initially, since region A is at a lower stage of development. Therefore "catch up" is only a *possibility* when $\dot{Y}_A > \dot{Y}_B$, because although B is growing relatively slowly, it is doing so from a higher initial base level of income.

the international ranking of economies by income.

c. Technological interrelatedness, lock in and the endogenous break down of virtuous circles.

The relative reversal of B's fortunes outlined above is induced by the virtuous circle in this region ultimately creating conditions which are dysfunctional to the maintenance of high growth dynamics. Because the ability of an economy to undergo technological change depends on the degree of technological interrelatedness it experiences, fast growth which induces increases in interrelatedness will reduce an economy's capacity to realise dynamic increasing returns associated with technological change. The technological evolution of the economy will therefore be inefficient, in the sense that it will not be functional to the maintenance of the high growth dynamics associated with a virtuous circle. This may also provide an initially less developed region (such as A) with the chance to break into fast growth and dynamic leadership by winning a technology "adoption race" against (and subsequently market share from) the initially more developed region B. Instead of success breeding success indefinitely in region B, allowing for the possibility of an induced supply side effect such as technological interrelatedness and lock in suggests a way in which the dynamics of a virtuous circle of cumulative causation may endogenously break down, and even be relatively reversed.

V Summary.

It is important to note that the analysis of this chapter, based on the augmented model of equations [3.1] - [3.4] and equation [4.2], is thoroughly in keeping with the methodology of hysteresis on which we are basing our modelling exercise. The supply side effects that are held to pose potential problems for the dynamics of cumulative causation arise endogenously, as a result of the historical time path of the economy. Essentially, then, the analysis in this chapter extends our initial frame of reference to generate a broader conception of the process of cumulative causation, which takes account of the significance of induced supply side changes which an economy undergoes in the course of growth and development. This exercise is valuable for two main reasons.

a. "Dynamic leapfrogging" and hysteresis in the augmented model.

The first valuable feature of the augmented model is that it introduces "dynamic leapfrogging" into the process of cumulative causation. No longer is it the case that success breeds success and failure begets failure indefinitely; the dynamics of cumulative causation may be affected by induced supply side changes. These arise in the long run as an endogenous response to the process of growth as cumulative causation itself. The technological evolution of an economy is influenced by technological interrelatedness and lock in arising in the course of growth. Hence the ability of an economy to realise dynamic increasing returns, which is sensitive to the "technological environment" in which growth is occurring, must be treated as deeply endogenous - it will vary with the precise growth path of the economy over time. As a result, long run competitive advantage no longer depends only on the initial conditions prevailing in competing regions. Instead, a successful region may sow the seeds of its own future decline, and come to experience not only a diminution of its virtuous circle, but even a relative reversal of its growth performance.

In connection with this, note also that the long run relative growth outcome associated with equations [3.1] - [3.4], which may be written as:

$$\frac{\dot{Y}_j}{\dot{Y}_w} = \frac{\lambda_j(\gamma_j - \alpha_w\beta_j)}{1 - \lambda_j\alpha_j\beta_j} \qquad [3.6]$$

is also affected by equation [4.2]. More specifically, if α_j is deeply endogenous and subject to adjustment asymmetries, this long run relative growth outcome is no longer determinate. Note, for example, that if $\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} < 0$ in equation [4.2] so that $d\alpha_j/dt < 0$, a decline in α_j will reduce the value of $\lambda_j\alpha_j\beta_j$ thus increasing $(1 - \lambda_j\alpha_j\beta_j)$ and so lowering $\dot{Y}_j/\dot{Y}_w$. The outcome in [3.6] therefore depends on the path taken towards it, so that when the model in equations [3.1] - [3.4] is augmented by equation [4.2], it becomes hysteretic *regardless* of assumptions, such as those made in chapter 3, about disequilibrium adjustment speeds.

The model developed in this chapter thus generalizes Kaldor's model of cumulative causation into a more generally hysteretic formulation, where successive changes in the variables of interest (export, output and productivity growth) need not be positively correlated over time. It also extends the Thirlwall (1980) representation of this model (captured in equations [3.1] - [3.4]) into a more generally hysteretic form, in which long run relative growth outcomes are strictly indeterminate, depending in all circumstances on the prior growth paths taken towards them.

b. Polarisation versus emulation.

The second valuable feature of the augmented model of cumulative causation developed in this chapter relates to Kaldor's (1987) distinction between "polarisation" and "emulation." Kaldor argues that the type of model developed in chapter 3 involves a "polarization process," in which growth is inhibited in some regions and highly concentrated in others. However, he argues that throughout economic history, this polarization process has, to some extent, been counteracted by the spread of industrialization between regions. This point is illustrated by Cornwall's (1977, pp.97-

121) analysis of technology transfers between regions and their impact on the rate of growth. It has also become a central feature of the literature that seeks to explain relative growth rates in terms of technological "catch up" (see, for example, Baumol, Blackman and Wolff, 1989). Kaldor is therefore led to conclude that

> The interaction between these forces - i.e., that of polarization which leads to concentration of development in successful areas, and of imitation or emulation which leads to the spread of industrialization into a wider range of areas - has never...been properly explored.
>
> (Kaldor, 1987, p.206)

It is precisely this interaction that the model developed in this chapter seeks to capture. Analysing the growth of technological interrelatedness within the context of a virtuous circle integrates the consequences of both the polarization and emulation processes referred to above within a single, long run model. This illustrates both how these processes work in opposition to each other, and the possible long run consequences of their opposition. As indicated above, the advantage of this model over the traditional Kaldor model of cumulative causation is that it does not predict that initial high/low growth will propagate subsequent high/low growth indefinitely. Its advantage over models based on the notion of technological "catch up," meanwhile, is that it allows for the possibility of changes in international economic rankings by income, rather than simply convergence towards some technological leader (see also Cornwall and Cornwall, 1992).

Chapter 5.

The Supply Side and Macroeconomic Performance II: Institutional Evolution.

In recent years, renewed interest has been shown in both the importance of institutions for the functioning of capitalism, and the impact of capitalism's dynamics on the nature of institutions.[1] However, the common interests of this literature conceal profound methodological differences, which in turn give rise to substantially different conclusions with regard to the origins, nature and role of institutions in a capitalist economy.

This chapter considers the nature of institutional evolution, and its consequences for the growth dynamics of a capitalist economy. By illustrating the potential for the emergence and persistence of inefficient institutions, its development parallels that of chapter 4, insofar as inefficient institutional evolution is identified as a potential weak link on the supply side of the economy that may cause a virtuous circle of cumulative causation to break down.

I Defining institutions.

Before proceeding, it is important to define what we mean

[1] For a survey of the different branches of this literature, see The Review of Political Economy, Vol. 1, No. 3, 1989.

by the term "institutions" in the context of a capitalist economy. Economic institutions constitute "any correlated behavior of agents ... that reoccurs under the same or similar conditions" (Dopfer 1991, p.536). Institutions are "proceduralist" rather than "consequentialist", influencing the type of behaviour that occurs in a particular situation independently of the goal orientation of the individual actor (Elster, 1989a).[2] As social relations that frame the activities of production, consumption and exchange, institutions can be thought of as a structure within which individual action in the economy takes place.

Institutions thus defined posses a number of salient features. First, they are distinct from organizations (Bromley, 1989). The existence and operation of organizations depends on institutions, but the two are not equivalent. For example, in the labour market both trade unions and firms are examples of organizations, whose internal structure depends on institutions (rules and norms concerning the appointment of officials, decision making processes, etc.) and whose relationship to each other is governed by a set of institutions (the industrial relations system.)

Second, following Basu, Jones and Schlicht (1987), the definition of institutions used above allows us to subsume under a single heading what are variously referred to in

---

[2] Whether or not institutions are functional to the goal orientation of individual actors will be discussed subsequently.

institutional analysis as institutions, norms, rules and customs. Finally, the definition is suggestive of arrangements and patterns of behaviour which do not originate from the actions of one individual alone, and/or which affect the actions of more than one individual i.e., institutions are collective rather than individualistic in nature.[3]

In the analysis that follows, we will implicitly distinguish between institutions on one hand, and basic features of capitalism such as the price mechanism and private property on the other. From a broader historical perspective, these "basic features of capitalism" may themselves be regarded as historically specific institutions, whose origins and role require analysis. Whilst the importance of this argument is not denied, its implications are held to be beyond the scope of the present analysis; our interest in the functioning of capitalist economies is assumed to justify a distinction between these intrinsic features of capitalism and what have been defined above as institutions.

II Is the price mechanism enough? Early views on the role of institutions.

Institutional analysis is by no means a new field of research. However, neither has there always been universal acceptance of the importance of institutions for the

[3] Although this does not, of course, preclude the possibility of an explanation of institutions based on the aggregate effects of individual action.

functioning of a capitalist economy. Hayek (1945) argues that the price mechanism alone is sufficient to ensure that an economy operates efficiently.[4] For Hayek, the basic economic problem facing society is how to organize the economy so that widely dispersed, individual specific information can be used in such a way that the economic system as a whole functions efficiently. He argues that whilst most events in an economy might be seen as having some bearing on the decisions a particular individual has to make, individuals do not need to be aware of all these events, nor all their effects. Instead,

> It does not matter for him (her) (the individual decision maker) *why* at the particular moment more screws of one size than another are wanted ... all that is significant for him (her) is *how much more or less* difficult to procure they have become compared with other things with which he (she) is also concerned.
>
> (Hayek, 1945, p.525, emphasis in original)

Hayek argues that the key feature of the price mechanism is that it embodies exactly this information; prices consist of "rates of equivalence" which attach an index to each scarce resource reflecting its significance in the economy as a whole. Thus for Hayek, all relevant information is

---

[4]. Efficiency here refers to allocative efficiency, i.e., a situation where "...the marginal rates of substitution between any two commodities or factors ... (are) the same in all their different uses." (Hayek, 1945, p.519)

It should be noted that Hayek has more recently expressed a greater interest in the role of institutions in a capitalist society. See Rutherford (1989) for a review.

concentrated in prices.[5] By adjusting to price signals, individuals are seen to be able to organize their behaviour without having to "solve the whole puzzle of the economy," and in such a way that their actions are consistent with the efficient functioning of the system in aggregate.

This view, that "the price mechanism is enough," is challenged by Richardson (1959), who argues that prices do not embody sufficient information to ensure that an economy with decentralized decision making will function efficiently. According to Richardson, the configuration of an economic system at a point in time depends on both a set of economic actions, and an accompanying set of beliefs or expectations concerning economic conditions. Two different types of economic conditions may be identified: "primary" conditions, which correspond to the data of traditional equilibrium analysis (preferences, technologies etc.) and "secondary" conditions, which refer to the projected activities of other agents in the system. Richardson argues that the price mechanism may embody sufficient information to facilitate the formation of expectations concerning primary conditions. However, he argues that expectations concerning secondary conditions are a different matter, since the price mechanism cannot tell us exactly how other agents in the system will

[5] Although this, of course, takes as given the definition of the commodities available on the market - i.e., what goods and services are provided through the market, and why any particular commodity exists in the precise form that it does.

react to its signals.[6]

These considerations lead Richardson to conclude that in practice, a variety of institutional factors usually considered as "frictions" or "imperfections" actually play a vital role in transmitting important secondary information to individual agents. For example, "goodwill" (i.e., consumer loyalty, which may be identified as a behavioural norm) may provide information to firms concerning the likely future behaviour of their customers in the event of price fluctuations. Other product market institutions such as formal or implicit market sharing and barriers to entry may provide an indication to firms as to how actual or potential competitors will react in a particular economic environment. In short, Richardson contends that institutions are sources of important secondary information which is not embodied in price signals. The procedural nature of institutions allows agents operating under conditions of uncertainty to anticipate the likely future behaviour of other agents in the economy.[7] As such, institutions are instrumental in the operations of a decentralized market economy.

---

[6] The one exception to this is if a single "true" model of economic activity is salient amongst all individuals. Then any one agent can form expectations of how others will react to a given signal simply by equating their response with his/her own response to that same signal.

[7] The term "uncertainty" is used here in its Keynesian context, to denote a situation in which individuals exist in a state of partial ignorance about the future.

III Competing conceptions of institutional evolution: the "old" and "new" institutional economics.

An increasing acceptance of Richardson's position - that "the price mechanism is not enough" - has played an important part in provoking much of the current interest in institutional economics. However, this emerging consensus that institutions matter belies the significant disagreement between different schools of thought as to why this is so. This conflict of opinions is illustrated by the contrast between two highly influential methodologies in institutional economics - what may be referred to as "old" institutional economics[8] (OIE) and the neoclassical "new" institutional economics (NIE.)

a. The OIE.

The OIE, associated with authors such as Commons (1961), Veblen (1975) and Ayres (1962), explains institutions in terms of a historical mode of analysis. By tracing institutions from one period to the next, their existence is accounted for on the basis of the principle that earlier states account for later ones. The OIE pursues a holistic methodology, postulating that the economy cannot be understood as a set of

---

[8]. Following Rutherford (1989), the term "old" institutional economics is used here "... (not) to imply that the tradition concerned is no longer vital, only that it represents the longer history of continuous and central concern with institutional questions." (Rutherford 1989, p.300.)

separable parts. Holism implies that individual phenomena cannot be explained without reference to the whole of which they form a part - that the characteristics and functioning of the part depend on its relations with other parts, and hence its place in the whole. For example, in the OIE the behaviour of individual agents in the economy is seen as a function of existing institutions, which form an environment to which individuals become socialised over time.

The holistic methodology of the OIE leads to an emphasis on structure (i.e., institutions) over action (i.e., the choices and activities of individuals) in the determination of economic outcomes. One potential shortcoming of this "structuralism" is that it can lead to a de-emphasis of the role played by goal oriented individuals, both in determining economic outcomes, and in shaping the institutional environment itself to meet their own ends (Brunner, 1987). This is particularly true of the Veblen/Ayres tradition, in which institutions are treated as constraints inhibiting technologically determined social and economic progress. In these circumstances, the individual is easily caricatured as a "plaything of social forces or the passive executor of inherited standards" (Elster 1989b, p.97).

A related problem, as Brunner (1987), Basu, Jones and Schlicht (1987) and Rutherford (1989) argue, is that the OIE often attributes the existence of current institutions to their history as if this in and of itself explains their origins and persistence. For example, Gordon (1980, p.17) asserts that a basic characteristic of institutions is that they are inherited as a legacy of the past. The problem with this argument is that it suggests that contemporary institutions can be taken as given from the past. This precludes a proper explanation of institutional evolution since, devoid of any theoretical or historical dynamics, it provides no indication of what types of institutions will survive over time, or why an institution's demise should occur when it eventually does.[9] Although Gordon argues that technological change provides the imperative for institutional change, it will be recognised in light of our earlier criticism that this simply replaces one autonomous, supraindividual force (institutions) with another (technology) (Rutherford, 1989).

b. The NIE.

It is perhaps ironic that in attempting to overcome the problems associated with the OIE, many authors have pursued the opposite methodological extreme and based their analysis of institutions entirely on the self interested behaviour of individuals. These theories, associated with authors such as Schotter (1981) and Williamson (1985), make up the NIE.

---

[9] This is particularly troublesome since, for many institutionalists, the notion that institutions are not immutable but subject to change over time is central to the OIE. See, for example, Miller (1978, pp.14-17).

The central contention of the NIE is that institutions arise spontaneously in the course of market activity.[10] The core of this analysis is a model of rational economic behaviour on the part of individuals, who are defined by psychologically given preference structures, and whose key feature is that they are evaluative i.e., they are willing to make trade offs between things that they value positively in order to achieve the greatest personal benefit. Proponents of the NIE argue that the application of this model is not limited to some narrowly defined "economic sphere" of activity.[11] Rather, they suggest that it characterises all social behaviour. Hence they arrive at the notion of a "generalized trade off" in which social arrangements such as institutions are seen to arise essentially because they are positively valued by rational maximizing individuals.

Of course, this still leaves us to explain precisely how social arrangements which are extraneous to the individual come about. The emphasis in NIE analysis is on "invisible hand" mechanisms; it is argued that interaction between rational individuals produces social patterns of behaviour without any one individual directing the results. The most common invisible hand mechanisms appealed to are repetition in

---

[10]. This is usually coupled with the further claim that these institutions will be efficient. This aspect of the NIE will be discussed in more detail subsequently.

[11] See, for example, Brunner (1987).

game theoretic models[12] (for example, Schotter, 1981) and evolutionary selection processes[13] (for example, Alchian, 1950.)

The NIE infers that the institutions of a capitalist economy are completely endogenous to market behaviour, in the sense that they are constantly evolving in response to rational individual actions. In other words, it pursues a reductionist methodology; the explanation of the institutional structure of an economy is entirely subsumed within an individualistic, choice theoretic framework. Causality is strictly uni-directional, with institutions arising solely in response to the current maximizing behaviour of rational individuals. In the NIE, individual action gives rise to structure (i.e., institutions) and any reverse line of causality is denied.[14]

However, if, as the OIE argues, institutions form a "social environment" to which individuals are socialized, the preferences and goals of individuals on which their actions

---

[12] This involves the same game being played repeatedly in consecutive periods. The repetition mechanism in games of strategy can be thought of as a metaphor for time in what is otherwise a static framework of analysis.

[13]. Chiefly the notion that competition will give rise to a process of "economic natural selection" amongst institutional forms.

[14] North (1985) departs somewhat from this methodology by allowing for the influence of ideology in the formation of institutions. However, NIE theorists generally treat individual maximising agents as asocial atoms, whose behaviour is independent of any form of social structure.

are based can no longer be treated as "psychological givens." Instead, they can only be interpreted within a particular institutional context (Hodgson, 1986). Rutherford (1989) suggests that this criticism leaves the reductionist with two options. The first is to deny that existing institutions can significantly affect current behaviour and hence the development of new institutions. However, Rutherford argues that it is hard to avoid the notion that existing institutions can be conservative, resisting change and hence influencing the subsequent development of the institutional structure. There may be vested interests in pre-existing institutions which will deliberately resist change.[15] Alternatively, institutional inertia may arise spontaneously if the cost of enforcing compliance to rules increases substantially when institutions are changed frequently.

The second option identified by Rutherford is for the reductionist to accept that pre-existing institutions do affect current behaviour and hence the development of new institutions, but to argue that it is possible to trace institutional evolution back to some original pre-institutional starting point. This approach is popular in game theoretic strands of the NIE. The caveat in this argument is that it does not avoid the necessity of taking some initial

[15] Bardhan (1989) distinguishes between efficiency enhancing and distributional institutional changes, where the latter are related to control over economic outcomes i.e., the exercise of power. Distributional claims of this nature may motivate the "vested interests" identified by Rutherford.

set of rules or norms of behaviour as exogenously given (Field, 1984). For example, the emergence of co-operative solutions in game theory depends vitally on games being repeated. This in turn requires that no party can simply exit the game or inflict catastrophic damage on others that prevents retaliation (Rutherford, 1989). In other words, the games have a prior structure that is not explained by individual self interested behaviour.

These criticisms suggest that it is not possible to conceive individual action as being independent of the institutional context within which it occurs. The methodclogical individualism of the NIE is flawed in that, by seeking to explain institutions wholly in terms of individual behaviour, it overlooks the reverse line of causality (attributable to the OIE) through which individual behaviour is influenced and constrained by institutions.

IV Towards a new conception of institutional evolution: hysteresis and institutional structures.

The criticisms outlined above suggest that neither the structuralist approach of the OIE nor the reductionism of the NIE offer truly satisfactory methodologies for institutional analysis. What is needed is an approach which captures the dynamic interaction between institutions and activity - one which recognizes the importance of current behaviour in shaping future institutions, but which at the same time takes account of the extent to which this behaviour is constrained by pre-existing institutional arrangements.

These considerations suggest that the institutional structure of an economy may be best conceived in terms of a process of hysteresis. In an institutional context, hysteresis will exist when current institutions influence the nature of current economic activity, which in turn influences subsequent institutional forms.

The joint interaction between institutions and economic activity in a process of hysteresis has previously been considered by Cornwall (1990) and Cornwall and Cornwall (1987). However, these authors are primarily concerned to illustrate the implications of this joint interaction for the unemployment - inflation trade off, rather than the nature of institutions themselves. To illustrate the process of "institutional hysteresis" when the latter is of primary concern, suppose initially that the choice sets of individuals and groups are defined by currently existing institutions.[16] These institutions will consequently exert a straightforward influence on current economic activity. This activity in turn

[16] Whilst existing institutions will define the *actual* choice sets faced by individuals and groups (for example, the law places bounds on consumption sets) it is important to note that even if they only define *perceived* choice sets, institutions may still influence the nature of economic activity. Hence social norms which bind some individuals to traditional lifestyles and communities may affect labour mobility, despite the fact that they place no literal constraint on an individual's choice of occupation and place of residence.

gives rise to economic outcomes, which may be judged by individuals and groups as either good or bad. Depending on this evaluation, individuals and groups will then attempt to either modify existing institutional arrangements or retain the status quo, which will in turn affect the form of future institutions and hence future choice sets and activities and so on. The essence of this model of institutional hysteresis is summarised by the following causal chain:

$$\begin{tikzcd} \text{Institutions} \arrow[r] & \begin{array}{l}\text{Definition of}\\ \text{individ./group}\\ \text{choice sets.}\end{array} \arrow[r] & \begin{array}{l}\text{Economic}\\ \text{activity}\end{array} \arrow[r] & \begin{array}{l}\text{Outcomes}\\ \text{and their}\\ \text{evaluation}\end{array} \arrow[lll, bend right] \end{tikzcd}$$

The central feature of the model proposed above is that institutions act as exogenous constraints in the short run, but in the long run, they are endogenous to the workings of the economy. The short run exogeneity derives from the OIE, where institutions are not assumed to display the "putty like" malleability necessary for their moment by moment renegotiation, as in the NIE. Instead, some degree of inertia is postulated, which will give rise to an institutional environment framing current economic activity.

However, the possibility that pressure for institutional change will arise from individual's and group's evaluations of current economic outcomes is not overlooked; hence the long run endogeneity of the institutional structure, which derives from the NIE, allows economic outcomes to have feedback effects on the form of institutions. Unlike, the reductionist NIE, however, there is no tendency towards some determinate long run "institutional equilibrium" in the current model. Rather, long run institutional changes can only be interpreted in terms of the sequence of short run patterns of economic activity leading up to them - patterns of activity which are themselves influenced by previously existing institutions. In other words, long run institutional changes are path dependent, deriving from the specific adjustment path the economy takes towards them. This is a fundamental difference between the way in which the endogeneity of the institutional structure is conceived in a model of institutional hysteresis, and the way it is conceived by the NIE.

The short run exogeneity/long run endogeneity of institutions in the model of institutional hysteresis gives rise to a second important feature of this model. Specifically, it is consistent with a view of institutional evolution which foresees periods of institutional stability (deriving from the short run exogeneity) punctuated by substantial institutional change (the manifestation of the long run endogeneity.)[17] According to this view, whilst economic outcomes, their evaluation, and subsequent pressures for institutional change may arise continuously, there is some

---

[17] Note that the model itself does not *predict* this form of institutional evolution, but is *compatible* with it due to the distinction made within the model between short and long run aspects of institutional evolution.

extent to which this is counterbalanced by forces of institutional inertia. Hence unlike the NIE, which suggests that institutional evolution is a continuous process, the model of institutional hysteresis allows for the possibility of different institutional forms persisting over time in an episodic fashion.

V Hysteresis and the efficiency of institutions.

The model of institutional hysteresis developed above combines exogenous and endogenous aspects of institutions in a historically consistent manner. It reconciles some of the most valid methodological contentions of the OIE and the NIE, whilst avoiding their worst excesses.

However, an important feature of the NIE that has so far been overlooked is its claim that institutions are functional to the dynamics of capitalism. According to the NIE, the invisible hand co-ordinates the activities of rational individuals in such a way as to give rise to institutions that are socially efficient.[18] Is it the case, then, that a hysteretic theory of institutions subscribes to the view that evolving institutional structures are socially efficient? The answer to this question is in the negative, for a number of reasons.

---

[18]. Institutions are assumed to be "socially efficient" in the sense of being Pareto optimal i.e., no alternative set of institutions exists that could make one individual better off whilst leaving no others worse off.

a. The persistence of inefficient institutions.

As previously intimated, the distinction between short run exogeneity and long run endogeneity in the model of institutional hysteresis permits a further distinction between the emergence of institutions, and their persistence over time once they have become the dominant form of procedure in a particular situation. This distinction is not immediately apparent from the moment by moment renegotiation of institutions in the NIE. Its importance, however, is that it gives rise to the possibility of institutions persisting over time even when they have ceased to be efficient.

As with the case of technology discussed in chapter 4, one way in which this can occur is if an economy experiences lock in to a pre-existing, inefficient institution. Once again, it is precisely this type of problem to which interrelatedness between and within institutions can give rise. By definition, institutions imply a degree of connectedness between individual behaviours in an economy - i.e., there exists interrelatedness *within* any institution. Because of this, an inefficient institution may persist by virtue of a "mutually sustaining structure of sanctions" (Basu, Jones and Schlicht, 1987; Akerlof, 1976) The self interested individual will comply with the behaviour implied by an institution, even if it is inefficient, in order to avoid being ostracised. Meanwhile, others in society will continue to ostracise deviants fearing that if they do not, they will in turn be socially sanctioned. The changes necessary to rectify this situation are of a non-marginal nature. They require a restructuring of the institution as a whole which, as illustrated above, cannot be brought about by incremental changes in the behaviour of the individuals who currently comply with the institution. On the contrary, no one individual can deviate from the institution given the behaviour of all others. As a result, the system will become locked in to this institution. The essence of the problem in this case is that the size of the decision making unit (the individual) is small relative to the scale of the unit of change required (the entire institution) due to the interrelatedness between individuals which is naturally implied by the institution.

Lock in arising in this manner may help explain the persistence of a variety of economic behavioural norms, such as the tendency of businessmen to wear ties even on hot days. A further example is provided by Akerlof (1976), who argues that continued racism in a firm's hiring practices may arise from the fear of ostracisation by others which would result if the firm hired a worker from the "wrong" racial group. At the same time, other firms would systematically punish any such deviance, for fear of the sanctions that would be consequent upon their failure to do so.

Interrelatedness may also exist *between* institutions. For example, Elbaum and Lazonick (1986) argue that Britain's industrial relations system and its pattern of industrial organization at the turn of the century were highly interrelated. To the extent that this is the case, and changes in one institution require complementary changes to a second with which it is interrelated, connectedness between institutions will clearly compound the type of problem, identified above, concerning the size of the decision making unit relative to that of the "institutional unit" which requires change. Again, the possibility arises that faced with non-marginal changes beyond the control of any one individual decision making unit, the economy will experience lock in to an inefficient institution.

When interrelatedness exists between institutions, a second potential problem arises. Even if the decision making unit is relatively large, changing an inefficient institution may be prohibitively costly due to the changes that would be required in other institutions with which it is interrelated. This is despite the fact that net benefits would accrue to changing the inefficient institution itself, and that such changes might actually occur in a competitor economy which suffers a lower degree of institutional interrelatedness. Even though the economy would be acting rationally on a cost benefit basis by not changing the inefficient institution, it may nevertheless experience "regret," resulting from the knowledge that it is locked in to an institution which is demonstrably inferior to an available alternative that is being adopted elsewhere.[19]

b. The dynamics of institutional hysteresis and the emergence of inefficient institutions.

There are two ways of conceiving the emergence of institutional structures: by considering a competitive struggle between existing institutions on the basis of the assumption that one will emerge as dominant, or by considering the processes leading up to the very genesis of a particular institutional form. We will examine each of these conceptions in turn, illustrating that in both cases, the dynamics of path dependence inherent in the model of institutional hysteresis suggest mechanisms which may lead to the emergence of inefficient institutions.

As noted earlier, an emphasis on the role of "natural selection" in the emergence of dominant institutional forms is common in some variants of the NIE. The natural selection mechanism conceives institutional outcomes as arising over time, as a result of a sequential, competitive struggle for survival between pre-existing institutions.[20] Usually, the

---

[19] This is to say nothing of the competitive decline the economy may experience if the inefficient institution subsequently affects its relative economic performance. This issue is taken up in more detail subsequently.

[20] This emphasis on a sequential, competitive struggle taking place *over time* makes the natural selection mechanism compatible with the methodology of hysteresis, discussed in chapter 2. This contrasts with the repetition mechanism in game theoretic variants of the NIE (see footnote 12 above), in which institutional outcomes are determined at a *point in time*

outcomes of this process are held to be efficient. However, as Langlois (1986) notes,

> ... organizations well designed to economize on transaction costs in the long run may find themselves selected out by a very inhospitable selection environment in the short run, leaving behind an observed population of relatively ill adapted forms.
>
> (Langlois, 1986, p.21)

The causal chain underlying this reasoning is of the form:

Nature of Economic Environment → Survival of Particular Organizational Forms → Nature of Economic Environment → etc.

This path dependent process contains no necessary implication that surviving organizations (and hence the institutions they embody) will be efficient. On the contrary, as Langlois effectively argues, a hostile short run selection environment may mean that institutional forms which are efficient in some long run sense are "selected out" in the short run, the bequest to the future of this process being a collection of inefficient institutional forms. If this creates a subsequent selection environment which is again hostile to efficient

---

as a result of a dynamic programming exercise by individual agents. The repetition mechanism is therefore fundamentally incompatible with the methodology of hysteresis, and is therefore not considered further. See also Setterfield (1992b).

institutions, then the process will become self reinforcing over time. As efficient institutions are repeatedly selected out, so inefficient institutional forms will emerge as dominant.

To illustrate this, consider the following example adapted from Hodgson (1991). Suppose there exist two types of firms, type I and type II, in an expanding market which is attracting new entrants. The profits, π, earned by these firms are described as follows:

Type I: π = percentage of type I firms in market

Type II: π = 40 + (1/2 * percentage of type I firms in market)

The difference in the profit structures of these firms is intended to capture differences in their internal organization. Hence type I firms engage in knowledge sharing, and their profit structure reflects the fact that such firms are highly interdependent. Meanwhile, type II firms do not share knowledge, although they benefit partially from the existence of type I firms due to knowledge "spillovers" from the latter. Suppose we choose an initial environment in which type I firms account for less than 80% of all firms in the market. Then there exists an incentive for each new entrant to be of type II. As the proportion of type II firms in the market thus increases, so the selection environment will become less and less favourable to type I firms over time.

Eventually, type II firms will entirely dominate the market, despite the fact that the average profitability of these firms will be just 40, compared to average profits of 100 that would be realised if type I firms gained total market domination. The selection path emerging from the chosen initial conditions will therefore routinely select out the efficient practice of knowledge sharing, and form a groove dominated by an inferior institution. If the economy subsequently becomes locked in to this self selecting inefficient institution, then as was demonstrated earlier, it will persist over time.

The preceding arguments suggest that the process of "natural selection" amongst competing alternatives envisaged by authors such as Alchian (1950) and Friedman (1953) may not give rise to the emergence of efficient institutions if the process of selection is path dependent. This relationship between hysteresis and inefficient outcomes, which is analogous to a process of selection between exogenously given, Pareto rankable multiple equilibria, is summarised in Figure 1 below. Note that the axes in Figure 1 are not calibrated in any continuous scales. Rather, the diagram is merely intended to indicate 1) that the behaviour associated with institution B differs in the degree of its social efficiency relative to the behaviour corresponding to institution C, and 2) that there exists an intervening range of behaviours whose social (in)efficiency is such that non-marginal changes in behaviour would be required in order to replace institution B with institution C.

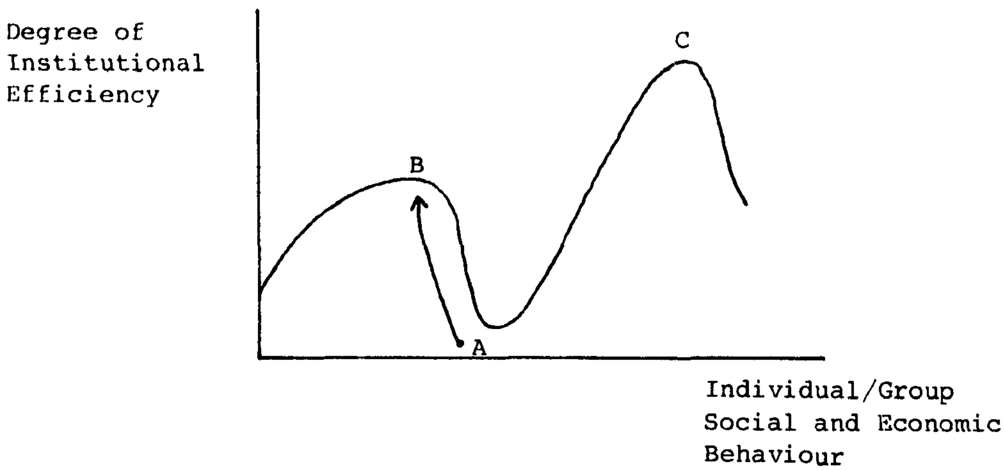

Figure 1. ***"Selecting out and grooving in:" hysteretic "natural selection" and lock in.***

Initial conditions at A give rise to the path dependent selection of the sub-optimal institutional form denoted by B. The arrow represents the selection path taken by the system in the course of this process. The system may become locked in to B if the distance to the optimal institution C and/or the depth of the intervening "valley" between B and C is too great, making the behavioural changes necessary for the replacement of institution B by institution C of a non-marginal nature.

However, the type of path dependence inherent in the model of institutional hysteresis affects more than simply the process of selection between existing institutional forms. Implicit in this model is the notion that the set of existing institutional forms is not static, but is itself path dependent. This property derives from the principle that past institutions are instrumental in creating a subsequent economic environment from which new institutions originate. In other words, the model involves an endogenous process of institutional procreation; it endogenously creates its own set of future institutional possibilities.

The notion of institutional procreation is omitted from the Alchian/Friedman concept of evolutionary natural selection, which, as noted earlier, concentrates on the competition for survival amongst pre-existing institutional forms. However, institutional procreation is vital to the process of institutional evolution, not least because it may, in and of itself, be responsible for the emergence of inefficient institutions. Hence note that in the model of institutional hysteresis, prior institutions create an environment in which economic behaviour takes place, and from which new institutions may subsequently emerge. The question that arises, then, is whether this environment is conducive to the procreation of efficient institutions. Just as the *selection* environment may be hostile to the selection of efficient institutions, so now the *creation* environment may be hostile to their very procreation. For example, the type of institutions that are created in a particular economic environment may depend more on the ease with which they can be created than on their efficiency properties (Hodgson, 1991). To see this, reconsider the earlier example involving type I and type II firms. Suppose now that a new industry is emerging in an initial economic environment characterized by decentralized decision making and adversarialism. Under these conditions, which are antithetical to the cooperation required for knowledge sharing, it may be easier for individuals to create "closed" firms than ones which engage in knowledge sharing. Hence only type II firms may be created in the new industry, regardless of the existence of any efficiency gains that would result from externalities associated with knowledge sharing.

It is even the case that initially efficient institutions may bequeath an evolutionary time path which ultimately results in the procreation of inefficient institutions. This point can be illustrated by the following example. Both Cornwall (1990) and Bowles, Gordon and Weisskopf (1990) argue that the postwar (1945–c.1966) institutions of social cooperation in advanced capitalist economies, which included the rise of the welfare state and greater acceptance of the labour movement, were functional to the dynamics of the postwar "golden age" of macroeconomic performance in these economies. However, they argue that this economic environment fostered new demands from labour which were beyond the domain of the original capital-labour consensus.[21] Hence the latter was replaced by a newly created (path specific) set of labour market institutions based on adversarial industrial relations, which both Cornwall and Bowles, Gordon and Weisskopf identify

[21] These "new demands" were related to concerns including control of the work process, hiring and firing practices, and the overall level of employment. They were accompanied by enhanced distributional claims by a labour movement buoyed by conditions of full employment.

as having been dysfunctional to advanced capitalist macroeconomic performance since the late 1960s. The "efficient" capital-labour consensus therefore ultimately gave rise to the subsequent procreation of inefficient labour market institutions.[22]

The process of path dependent institutional procreation is illustrated in Figure 2 below.

[22] Note that in this example, the economy is engaged not only in the creation of inefficient institutions, but also the destruction of previously efficient institutions. The latter theme is discussed by Schumpeter (1942), and also by Hirsch (1976), who argues that capitalism inexorably erodes values inherited from the pre-capitalist past (for example, values associated with religion). This is despite the fact that these pre-capitalist values may be instrumental in forming the "social prerequisites of markets" - shared values on such issues as the appropriate sphere of market activity - and may therefore, as authors such as Adam Smith and J.S. Mill believed, be functional to the successful operation of capitalism.

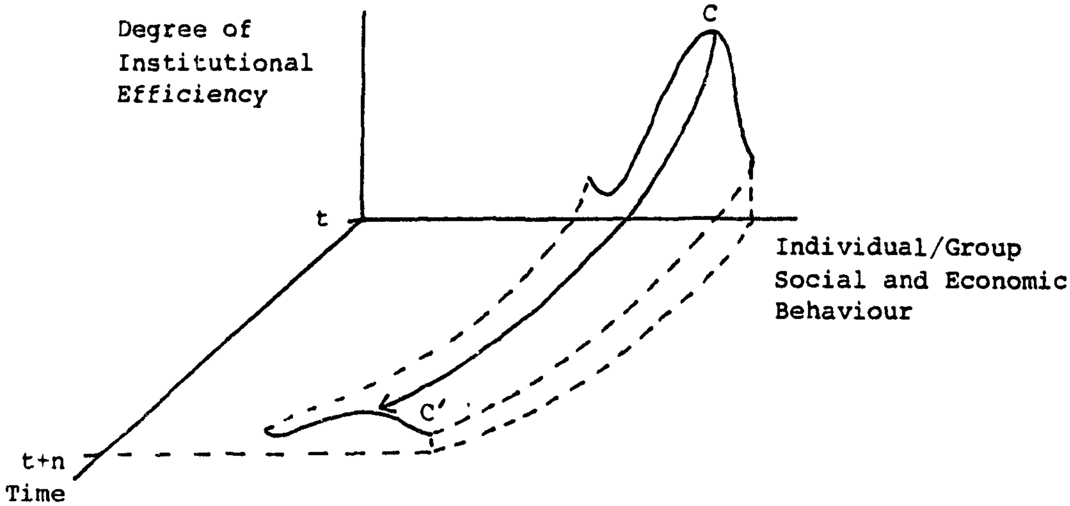

Figure 2: ***Hysteresis and the process of endogenous institutional procreation.***

In this case, the time path of the economy affects not only the process of selection between existing institutional forms; it also endogenously affects the nature of any institutional form created in the course of the economy's evolution. This is captured by the transformation of the institutional efficiency surface over time, in response to movement along the timepath denoted by the arrow between C and C′.

Figure 2 depicts the Cornwall/Bowles, Gordon and Weisskopf case where, starting with the institution C, an intertemporal adjustment path unfolds which changes the contours of the institutional efficiency surface so that the outcome C′ is inefficient by comparison with C by period t+n.

Figure 2 encompasses Figure 1, although this time it is assumed that the outcome of a selection process has resulted, at some initial time t, in the dominance of the relatively efficient institutional form C.[23] It then illustrates the

[23] Note also that for purposes of diagrammatic clarity, the institutional efficiency surface from Figure 1 has been truncated to exclude the institution B. This allows us to

dynamic (long run) consequences of this institutional structure, as it affects economic behaviour and outcomes and hence the creation of subsequent institutional forms between periods t and t+n. This is captured by the (endogenously) changing shape of the institutional efficiency surface over time.

The consequences of path dependence for the processes of institutional selection and institutional procreation illustrate that the dynamics of the model of institutional hysteresis are fundamentally incompatible with the notion that emerging institutions are always efficient.

VI Institutions in an evolving hysteretic economy.

The preceding analysis creates a picture of capitalism as a system which is dependent on institutions for its successful operation, but which contains no mechanism to ensure that efficient institutions will automatically emerge and persist. On the contrary, we have seen that conceiving institutional evolution in terms of a model of institutional hysteresis is immediately suggestive of mechanisms which will give rise to the emergence and persistence of inefficient institutions. These mechanisms arise from the sequential nature of the joint interaction between institutions (structure) and economic activity (action). They suggest that in the course of its

---

focus attention on the evolutionary timepath CC′, and the relative efficiency of the institutions C and C′.

development, an economy will endogenously generate institutional changes and/or institutional inertia which may be more or less functional to its macroeconomic performance.

There are obvious parallels between these considerations and those developed in chapter 4 - especially insofar as we have now seen that institutions, as with technology, can be inefficient by virtue of the influence of lock in. Bearing these parallels in mind, it is not surprising that the potential consequences for growth of institutional evolution in the presence of lock in mirror the consequences of technological evolution in the presence of lock in discussed in chapter 4. In short, the maintenance of high growth dynamics depends in part on the maintenance of an institutional environment conducive to these dynamics. Inefficient institutional evolution may corrupt and even relatively reverse a virtuous circle of cumulative causation. In order to see why this is so, it is instructive to concentrate on the institutions of the labour market.

a. Labour market institutions and their effects on growth.

In the model of chapter 3, an important basic assumption is that there exists a stable wage distribution and a variable productivity growth distribution across regions, which gives rise to a varying efficiency wage distribution. It is the self reinforcing changes in the efficiency wage distribution due to increasing returns which underlie the dynamics of this model.

However, suppose we now allow institutional changes in the labour market induced in the course of a virtuous circle to affect the wage distribution between regions. Specifically, suppose that these institutional changes can give rise to upwardly flexible wages in a virtuous circle economy, which capture an increasing share of the productivity gains arising from dynamic increasing returns. Then the ability of this region to maintain a relatively favourable rate of inflation, on which the subsequent growth of exports and hence the furtherance of the virtuous circle itself depend, will be impaired. This "Kaldor effect," so named because it emphasises the impact of institutional changes on nominal wage relativities and hence the efficiency wage distribution central to the model of chapter 3, suggests that induced institutional changes may cause the dynamics of a virtuous circle to break down.

Alternatively, suppose that institutional changes induced in the course of a virtuous circle lead to money wage inflation, but this time policy authorities respond by deflating the domestic economy in an attempt to "discipline labour" and purge the economy of cost inflation. This "Cornwall effect," so named because following Cornwall (1990), it emphasises the impact of institutional change on policy, may, by depressing domestic demand, reduce the extent to which dynamic increasing returns are realised. This will adversely affect the economy's subsequent ability to sustain high rates of growth of exports and hence output. Once again the economy will experience a growth climacteric - a deceleration of its trend rate of growth - as the dynamics of its virtuous circle break down.

Both the Kaldor and the Cornwall effects outlined above can be described as "indirect effects" of institutions on the dynamics of a virtuous circle. They affect the cumulative interaction of export, output and productivity growth via the prior effect of institutions on either nominal wage relativities (and hence the efficiency wage distribution) or on macroeconomic policy. However, following Weisskopf, Bowles and Gordon (1983) and Bowles, Gordon and Weisskopf (1990), it is possible to identify mechanisms through which institutions may have a *direct* effect on the dynamics of the model developed in chapter 3. What may be identified as the "Bowles, Gordon and Weisskopf (BGW) effect" suggests that labour market institutions may be more or less functional to the rate of productivity growth.

According to the BGW effect, productivity is influenced by the way in which labour resources are utilized at the point of production. Hence an industrial relations system which results in less capital-labour co-operation may affect worker morale and hence work effort, and may also result in an increase in the incidence of disruptions to the production process, such as absenteeism or strikes. This will in turn have an adverse affect on the intensity with which labour resources are utilized over time, and hence retard the rate of productivity growth.

Furthermore, the BGW effect also recognises that productivity growth is vitally dependent on the ability of firms to undergo changes in their processes of production which allow them to reap the advantages of embodied technological progress. Because they have a substantial impact on issues such as control over the labour process, the de- and/or re-skilling of labour, and labour displacement, the ease with which such changes can be made depends significantly on the compliance of labour in the production process. The less compliant is labour, the slower will be the pace of technological change within the firm, which will again adversely affect the rate of productivity growth.

b. Inefficient institutional evolution in the labour market and its consequences for cumulative causation.

In order to illustrate the possible implications of the BGW effect for cumulative causation, consider once again the augmented model of this process developed in chapter 4:

$$\dot{q}_{jt} = r_t + \alpha_{jt}\dot{Y}_{jt-1} \qquad [3.1]$$

$$\dot{p}_{jt} = \dot{w}_{jt} - \dot{q}_{jt} \qquad [3.2]$$

$$\dot{X}_{jt} = \beta_j(\dot{p}_{wt} - \dot{p}_{jt}) + \gamma_j\dot{Y}_{wt} \qquad [3.3]$$

$$\dot{Y}_{jt} = \lambda_j\dot{X}_{jt} \qquad [3.4]$$

$$\alpha_{jt} = f_j(\dot{Y}_{j0}, \ \ldots, \ \dot{Y}_{jt-1})$$

$$f_{ji}' \neq 0 \text{ some } i = 0, \ \ldots, \ t-1$$

$$\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} \neq 0 \qquad [4.2]$$

Once again, the key relation is equation [4.2], which captures the sensitivity of α_j, and hence the ability of region j to realise dynamic increasing returns, to the prior growth path of region j. In order to parallel the argument developed in chapter 4, suppose that higher growth in a virtuous circle region such as B in chapter 3 induces institutional interrelatedness, and consequently lock in to an industrial relations system which has ceased to be functional to technological change, and hence the realisation of dynamic increasing returns and productivity growth. The consequences of this will be picked up in equation [4.2] with $\sum_{i=0}^{t-1} f_{ji}'.d\dot{Y}_{ji} < 0$, so that $d\alpha_{jt}/dt < 0$. Other things being equal, this will have implications for the dynamics of region B's virtuous circle similar to those identified in chapter 4. First, the rate of productivity growth realised in equation [3.1] will fall due to $d\alpha_{jt}/dt < 0$. This will, in turn, reduce region B's relative inflation advantage and hence rate of export growth in equations [3.2] and [3.3], and so ultimately lower the rate of output growth in equation [3.4]. In other words, the economy will experience a growth climacteric as a result of its institutional structure becoming dysfunctional to the maintenance of high growth dynamics.

Furthermore, if region B's competitor, region A, has experienced relatively more efficient institutional evolution by virtue of its relatively lower prior growth rate, there exists the possibility of dynamic leapfrogging similar to that described in chapter 4. The "institutionally efficient" region A may reap a higher rate of productivity growth than region B, by virtue of its faster assimilation of new techniques of production. In terms of equation [3.1], if A's relatively more efficient institutional evolution implies that $\alpha_{At} > \alpha_{Bt}$, then we may observe

$$\dot{q}_{At} = r_t + \alpha_{At}\dot{Y}_{At-1} > r_t + \alpha_{Bt}\dot{Y}_{Bt-1} = \dot{q}_{Bt}$$

If region A's productivity growth advantage allows it to win market share from B, so that $\dot{X}_{At} > \dot{X}_{Bt}$ in equation [3.3], the relatively faster rate of output growth that A realises in equation [3.4] will become self reinforcing in subsequent rounds of the process of cumulative causation. As it loses export share, and possibly even sees its absolute advantage in income levels eroded over time, the formerly virtuous circle region will experience a period of relative economic decline as a result of its relatively inefficient institutional structure.

In sum, the hysteretic nature of institutions, coupled with the fact that hysteretically evolving institutions need not be efficient, suggests a further mechanism whereby supply side changes endogenous to the process of growth and development can cause the break down of a virtuous circle. In terms of the causal chain

$$\text{high } \dot{X}_t \rightarrow \text{high } \dot{Y}_t \underbrace{\rightarrow}_{b} \text{high } \dot{q}_t \rightarrow \text{high } \dot{X}_{t+1} \rightarrow \text{etc.}$$

the possibility of inefficient institutional evolution once again implies a weak link at point b, on the supply side of the model of cumulative causation. This weak link implies that long run relative growth performance no longer depend only on the initial conditions prevailing in competing regions. Furthermore, it implies that the outcome:

$$\frac{\dot{Y}_j}{\dot{Y}_w} = \frac{\lambda_j(\gamma_j - \alpha_w\beta_j)}{1 - \lambda_j\alpha_j\beta_j} \qquad [3.6]$$

associated with equations [3.1] - [3.4] is not determinate, but instead depends on the path taken towards it.[24] Once again, treating α as a deeply endogenous variable (this time, in response to induced changes in the institutional structure of the economy) extends the model of cumulative causation developed in chapter 3 into a more generally hysteretic formulation.

VII Conclusion.

Chapters 4 and 5 have shown that by augmenting a simple model of cumulative causation with induced supply side effects, we can develop a more general model of hysteresis which links Kaldorian dynamics based on increasing returns to longer term dynamics which affect the technological and institutional structure of the economy. In this model, success need not generate success indefinitely - it may bequeath the conditions for subsequent relative economic failure. Indeed, current success will only give rise to future success as long as the institutional and technological structure of the economy remains conducive to the continuance of a virtuous circle.[25]

---

[24] The reader is referred to section V of chapter 4 for a formal demonstration of this argument.

[25] Notice that the interpretation of capitalism that arises from this long run hysteretic model - that of a transforming system which encounters different epochs in its development characterised by different technological and institutional "structures of accumulation" - is similar to that arising from traditions such as the French Regulation School (for example, Boyer, 1990), and the American Social Structure of Accumulation school (for example, Bowles, Gordon and Weisskopf, 1990).

This "long run theory of effective demand" has a number of interesting implications for the evolution of a capitalist system characterised by interregional trade. In particular, it indicates that there are conflicting forces in development, which mean that the growth and accompanying structural change of an economy can at once both stimulate (through the realisation of dynamic increasing returns) and hinder (by inducing inefficient technological and institutional evolution) its future growth and development. The historical balance of these conflicting forces will govern the long run competitive fate of an initially fast growing region - a balance which, rather than being determinate, can be conceived as depending on a dialectical interaction of these forces.[26] Ultimately, then the (hysteretic) realised growth path of a capitalist economy can be thought of as just one history out of many possible histories that remain unrealised.

[26] Dialectics is used in this context to denote a theory of historical motion based on conflict between contradictory forces (see Elster 1985, pp. 37-48).

Part III

The evolution of a macroeconomy: growth and structural change in the British economy since 1780.

Introduction.

In Part II, it was argued that the rates of growth of exports, output and productivity in an economy jointly interact in a self reinforcing process of cumulative causation. However, it was suggested that success need not create success indefinitely in this framework; a virtuous circle of cumulative causation may induce supply side changes in the economy which lead to the breakdown of high growth dynamics. Augmenting the process of cumulative causation to take account of these induced supply side effects leads to a model with more generally hysteretic properties. It was argued that the pattern of relative success and failure in the same region to which this model can give rise is suggestive of the possibility of "leapfrogging" in the international ranking of economies by income.

The purpose of Part III is to illustrate the workings of the long run theory of capitalist development modelled in Part II by reference to the experience of the British economy since 1780. As the "first industrial nation," Britain enjoyed unprecedented prosperity in the nineteenth century, and dominated the global political economy. However, as early as 1870-1914, Britain's relative economic performance began to falter. In the twentieth century, this relative decline accelerated, and Britain surrendered its pre-eminence in the world economy. It will be argued that this pattern of relative success and failure is best understood as a path dependent outcome, brought about by self-reinforcing mechanisms at work in the British economy. In particular, we will argue that the period of relative success in the nineteenth century bequeathed the conditions for the subsequent period of relative economic failure.

I An empirical methodology based on economic history.

Throughout Part III, frequent use is made of historical evidence, which may also be qualitative in nature. Three important considerations recommend this empirical methodology over an approach based on econometrics. First, as was pointed out in chapter 3, the model we have used to characterize macroeconomic evolution is not a fully specified model of the macroeconomy, of the type which might be suitable for econometric estimation. Second, given the long run nature of the study, the lack of continuous (and consistent) data series forms a practical obstacle to econometric analysis. Finally, the emphasis of the theoretical model developed in Part II on the endogenous structural transformation of capitalist economies suggests that the macroeconomic data generating process underlying capitalism is endogenously changing, at least in the long run, or between what may be identified as different epochs in the history of an economy. This theoretical consideration implies that the classical statistical assumption of a stable population distribution will be violated in long run time series analyses of capitalism, invalidating econometric results based on this assumption. Furthermore, statistical modelling of the endogenously transforming macroeconomic population distribution underlying the British economy is held to be beyond the scope of the current thesis.

The preceding considerations demand that an alternative empirical methodology to the standard econometric approach be found. Hence throughout Part III, we will rely on the use of non-econometric evidence. We proceed on the basis of the assumption that such evidence provides a valid indication of the explanatory power of the theory of capitalist development expounded in Part II.

II Organization of Part III.

Part III is organized as follows. Chapter 6 provides a brief outline of the history of the British economy since 1780, highlighting the cumulative nature of Britain's rise and subsequent decline. Chapters 7 and 8 then concentrate on the years of relative failure after 1873, arguing that a combination of inefficient technological and institutional evolution help explain Britain's relative economic retardation in the twentieth century.

Chapter 6.

A brief history of the British economy since 1780.

The history of the British economy has drawn the attention of many leading scholars. Interpretations of distinct sub-periods of this history have been extensively debated - for example, the notion of a "Great Depression" 1873-96 (Saul, 1969) and the role of new industries in the 1930s economic "recovery" (Richardson, 1962; Dowie, 1968). Others have devoted attention to the formidable task of compiling accurate data to facilitate the economic interpretation of British history (for example, Feinstein 1972).

There is little that a brief survey of British economic history can add to debates over specific incidents such as the Great Depression. Any such survey is also necessarily limited by the adequacy of the available data. These qualifications aside, the purpose of this chapter is to outline the history of the British economy since 1780, showing how the relative success and decline of the British economy over this period conforms broadly to the pattern of the self reinforcing dynamics emphasised in Part II. In keeping with the theoretical analysis of Part II, we will chiefly emphasise the rates of growth of exports, output and productivity when evaluating macroeconomic performance. For expositional purposes, we will divide the period since 1780 into three sub-periods: 1780-1873, 1873-1950, and 1950-85. These correspond respectively to the period of Britain's initial industrial dominance, the period during which its relative decline first became apparent, and the period during which this decline became pronounced.

I "The first industrial nation," 1780-1873.

The period 1780-1873 was one of growth and prosperity for the British economy, which culminated in the Great Victorian Boom 1850-73 (Church, 1975), and saw Britain establish a dominant position in the global political economy.

Throughout this era, exports, output and productivity grew simultaneously. From 1801-60, the rate of growth of nominal GDP averaged 2.2% p.a. (Lee 1986, P.10), with peak rates of growth achieved during the Great Victorian Boom (Church 1975, p.20). Meanwhile, productivity growth averaged 1.9% p.a. between 1847 and 1874 (Church 1975, p.48), whilst the volume of exports of manufactured goods grew at an average rate of 5.6% p.a. 1830-57 and 3.1% p.a. 1857-73 (Lee 1986, p.117). Indeed, by the 1870s, Britain's share of total world exports of manufactured goods was 38%, and this had probably been higher in earlier periods (Harley and McCloskey 1981, p.51).[1]

---

[1] The decade 1876-85 is the earliest period for which reliable export share data is available. See Harley and McCloskey, 1981.

The expansion of overseas markets and relatively high (by contemporary international standards, and in the historical context of the British economy) trend rates of productivity and output growth were thus confluent in the British economy during the first three quarters of the nineteenth century. As the "first industrial nation," Britain achieved international economic pre-eminence, and established itself as technological leader in the world economy by 1820 (Maddison, 1982). However, the dynamic success of the British economy and its unrivalled dominance of world markets proved to be of limited duration.

II Emerging signs of relative decline, 1873-1950.

The period 1873-1950 itself comprises two distinct sub-periods: 1873-1914, and 1919-50. Between 1873 and 1914, Britain lost the unchallenged international position it had established as the "first industrial nation." It is important not to overstate the extent of Britain's relative decline during this period. Its export, output and productivity growth performance were not vastly inferior to that of the emerging new industrial nations (Floud, 1981), nor to Britain's own prior historical record.[2] Furthermore, with approximately a one third share of the world export market for manufactures,

---

[2] That Britain's economic performance 1873-1914 was not greatly inferior to that achieved during the preceding period may be explained in part by the fact that the British economy was never particularly fast growing, even during the period of its unchallenged dominance of the world economy (Maddison, 1982).

Britain continued to dominate world trade. However, Britain's grip on this position of dominance was clearly slipping.

After 1873, Britain was unable to maintain the rates of growth of output that it had achieved earlier in the nineteenth century (Floud, 1981). At a rate of 1.8% p.a., its annual average rate of real GDP growth 1873-1914 was lower than that in both the USA (4.5% p.a.) and Germany (2.8% p.a.), who were rapidly emerging as Britain's main industrial competitors (Floud 1981, p.8). Britain also experienced a productivity slowdown 1873-1914, the growth of total factor productivity averaging just 0.4% p.a.. Again, this compared unfavourably with the average rates of productivity growth in both the US (1.2% p.a.) and Germany (0.9% p.a.) (Floud 1981, p.22).

The geographical distribution of British overseas trade was also transformed after 1873, as Britain's traditional markets became self sufficient in goods such as textiles, which dominated British exports (Floud, 1981; Kaldor, 1981). In some instances (for example, the US) countries which had formerly been traditional markets for British goods actually began to compete with Britain as exporters in their own right. Accompanying the changing geographical pattern of its overseas trade, Britain therefore experienced a decline in its share of world exports of manufactured goods after 1873, a loss which was chiefly matched by the gains made by the US and Germany. This is illustrated in Table 1 below.

Table 1: *Shares of world exports of manufactures for the UK, US and Germany 1899-1950.*

	1899	1913	1929	1937	1950
UK	34.5	31.8	23.8	22.3	24.6
Germany	16.6	19.9	15.5	16.5	7.0
US	12.1	13.7	21.7	20.5	26.6

Source: Matthews, Feinstein and Odling Smee (1982).

In sum, Britain's export, output and productivity growth performance suffered a decline from 1873-1914, relative to both its own past performance, and the achievements of newly industrializing economies such as the US and Germany. This decline was exacerbated by the "Edwardian climacteric" - a deceleration in the trend rates of output and productivity growth[3] - whilst the loss of technological leadership to the USA by c.1890 was symptomatic of the weakening of Britain's dominant international position (Maddison, 1982).

Ironically, Britain's relative economic decline experienced something of a reversal during the period 1919-50. This is illustrated in Tables 3, 4 and 5 below, which indicate improvements in Britain's relative rates of growth of output,

[3] Although Britain's historical output and productivity growth slowdown dates from 1873, evidence suggests that this slowdown is most pronounced after 1899 (Feinstein, Matthews and Odling-Smee, 1982). This suggests that the notion of an Edwardian climacteric 1899-1914 is more useful than that of a "Great Depression" 1873-96 for the understanding of British economic history in the late nineteenth and early twentieth centuries.

per capita income and productivity between 1913 and 1950 when compared with the period 1870-1913.

Table 2: ***Rates of growth of total output for selected OECD economies 1870-1950 (percent per annum).***

	1870-1913	1913-50
United Kingdom	2.2	1.7
France	1.6	0.7
Germany	2.9[a]	1.2
Italy	1.4	1.3
United States	4.3	2.9

Note: a = 1871-1913

Source: Maddison (1964) p.28

Table 3: ***Rates of growth of output per capita for selected OECD economies 1870-1950 (percent per annum).***

	1870-1913	1913-50
United Kingdom	1.3	1.3
France	1.4	0.7
Germany	1.8[a]	0.4
Italy	0.7	0.6
United States	2.2	1.7

Note: a = 1871-1913

Source: Maddison (1964) p.30

Table 3: *Rates of growth of output per person hour for selected OECD economies 1870-1950 (percent per annum).*

	1870-1913	1913-50
United Kingdom	1.5	1.7
France	1.8	1.6
Germany	2.1[a]	0.9
Italy	1.2	1.9
United States	2.[illegible]	2.4

Note: a = 1870-1913

Source: Maddison (1964) p. 37

Trends in export share statistics also reflect a marked deceleration in Britain's relative economic decline between 1915 and 1950. During this period, the decline in Britain's share of world exports of manufactured goods was arrested, and even modestly reversed. As illustrated in Table 1, despite falling rapidly from 31.8% in 1913 to 23.8% in 1929, by 1937 Britain's export share was still as high as 22.3% and had actually risen to 24.6% by 1950 (Matthews, Feinstein and Odling Smee 1982, p.435).

The deceleration or even reversal of Britain's declining economic fortunes between 1913 and 1950 has received widespread attention in the economic history literature, and a number of theories intended to explain this phenomenon have been forwarded. Some authors concentrate on developments in the domestic economy, arguing, for example, that the new industries (motor vehicles, chemicals etc.) formed a "development block" which spurred a trend acceleration of British economic growth during the 1930s (Richardson, 1967). Others have emphasised developments in the external sector, arguing that Britain's improving fortunes were a response to changes in commercial policy after 1930, which became increasingly orientated towards protectionism (Kaldor, 1981; Kitson and Solomou, 1989).

It is important, however, not to exaggerate the extent of Britain's relative economic "recovery" after World War I. Britain's output and productivity growth performance 1913-50 was still worse than the average for the twelve European and North American economies considered by Maddison (1964),[4] and lagged considerably behind the achievements of at least one of the economies that had emerged as a major competitor for British overseas markets during the period 1870-1913 - the USA.[5] Furthermore, Britain's rate of productivity growth 1913-50 (1.7% p.a.) was little better than it had achieved 1870-1913 (1.5% p.a.) whilst its rate of growth of per capita income (1.3% p.a. in both periods) was no better at all (Maddison 1964, pp. 30, 37).

The interwar years were marred by unprecedented social

---

[4] Maddison's sample includes the US, Canada, UK, Switzerland, Sweden, Norway, the Netherlands, Italy, Germany, France, Denmark and Belgium.

[5] The years 1913-50 proved extremely turbulent for the German political economy, its average economic performance suffering accordingly.

strife which culminated in the equally unprecedented mass unemployment of the Great Depression. They also proved to be a period of "inter regnum" in international affairs. With the end of "Pax Britannica," the US became poised to replace Britain as the dominant power in the global political economy.

On the whole, then, the years 1873-1950 can be regarded as a period of transition for the British economy. Its macroeconomic performance was, on the whole, worse than it had been between 1780 and 1873, and worse than that of the economies which had begun to emerge as its chief rivals - but not greatly so. Britain experienced something of a "genteel decline;" it remained a major participant in the global economy, and suffered less from the ravages of the monetary instability of the 1920s and the Great Depression of the early 1930s than did some of the newer industrial economies (most notably, Germany and the US). However, whilst by no means disastrous, Britain's faltering economic performance and loss of international economic pre-eminence during this era appear, in retrospect, to have been signs of what was to come.

III The continuance and acceleration of relative decline, 1950-85.

The acceleration of Britain's relative economic decline between 1950 and 1985 is clearly illustrated by a comparison of its export, output and productivity growth performance to that of a number of its main OECD rivals. Table 5 below presents real GDP growth data for six OECD economies including the UK. It is clear from this table that although Britain's postwar rates of output growth were historically high (at least up until the late 1960s), they were relatively low compared with other OECD nations. Indeed, the years 1955-60 apart, Table 5 reveals that amongst this selection of economies, the UK consistently ranked last in terms of its growth performance throughout the postwar years.

Table 5: ***Rates of growth of real GDP for selected OECD economies 1950-85 (percent per annum).***

	1950-55	1955-60	1960-64	1964-69	1969-73	1973-79	1979-85
UK	2.9	2.5	3.4	2.5	2.8	1.4	1.1
Fra	4.4	4.8	6.0	5.9	6.1	3.1	1.1
W.Ger	9.1	6.4	5.1	4.6	4.5	2.5	1.2
Italy	6.3	5.4	5.5	5.6	4.1	2.7	1.4
Japan	7.1[b]	9.0	11.7	10.9	9.3	3.7	4.3
USA[a]	4.2	2.4	4.4	4.3	3.4	2.6	2.2

Notes: a GNP
b 1952-55

Sources: Cairncross (1981); Handbook of International Statistics (1986).

These trends are reflected in Britain's relative rate of per capita income growth, which as illustrated in Table 6 below, was consistently worse than that of any of its major competitors between 1950 and 1980, with the possible exception of the US.[6]

Table 6: *Rates of growth of real GDP per capita for selected OECD economies 1950-90 (percent per annum).*

	1950-60	1960-67	1967-73	1973-79	1979-90
UK	2.2	2.3	3.0	1.5	1.9
France	3.5	4.3	4.7	2.3	1.7
W. Ger	6.5	2.9	4.5	2.5	1.7
Italy	5.3	4.9	4.2	3.2	2.2
Japan	N/A	9.1	8.0	2.5	3.5
US	1.6	3.2	2.5	1.4	1.6

Sources: Maddison (1967); Handbook of International Statistics (1986); OECD Historical Statistics (1990).

Britain's relatively poor output growth after 1950 was matched by poor productivity growth and export performance. Table 7 below illustrates that between 1963 and 1985, the UK fared worse than any of its major competitors in terms of productivity growth performance, with the possible exception

[6] During the 1980s, Britain's relative rate of per capita income growth was more favourable. It was superior to that achieved by the US, West Germany and France between 1980 and 1990. This suggests that Britain's recovery from the recession of the early 1980s was relatively strong, although whether this is indicative of an improvement in its trend rate of relative per capita income growth is not possible to ascertain at the time of writing.

of the US.[7]

Table 7: *Rate of growth of output per person hour in manufacturing for selected OECD economies 1963-85 (percent per annum).*

	1963-73	1974-85	1963-85
UK	4.4	2.5	3.3
France	7.0	3.6	5.1
W. Germany	5.8	3.3	4.4
Italy	7.0	2.8	4.7
Japan	11.7	4.2	7.6
US	3.2	3.1	3.1

Source: Handbook of International Economic Statistics (1986).

Meanwhile, as is illustrated in Table 8, Britain's share of world trade in manufactured goods has declined monotonically since 1950. During the postwar years, Britain's share of world exports has been surpassed not only by the USA, But also by Germany, France and Japan. Indeed in 1983, Britain became a net importer of manufactures for the first time in its history (Rowthorn and Wells 1987, p.98).

[7] The possibility that since the 1960s, the US has entered a growth climacteric and is following a development path similar to that previously experienced by the UK is not investigated here. However, this possibility is beginning to draw the attention of a number of authors. See, for example, Elbaum, 1990; Elbaum and Lazonick, 1986.

Table 8: *Shares of world exports of manufactures for selected OECD economies 1950-85 (percent).*

	1950	1962	1970	1980	1985
UK	24.6	15.9	10.8	9.7	7.9
France	9.6	9.2	8.7	10.0	8.4
W. Ger	7.0	19.9	19.8	19.8	18.6
Italy	3.6	6.0	7.2	7.9	7.7
Japan	3.4	7.4	11.7	14.8	20.0
US	26.6	20.2	18.5	16.9	16.8

Sources: Matthews, Feinstein and Odling Smee (1982); Handbook of International Economic Statistics (1986).

Whilst the OECD economies as a whole experienced periods of both sustained growth and stagnation between 1950 and 1985, one constant feature of this era was the relatively poor performance of the British economy. In terms of its export, output and productivity growth performance, Britain was surpassed not only by the US (which finally assumed Britain's former position of world hegemonic power) but also by a number of its European and Asian competitors. In the course of the postwar years, Britain established itself within the OECD as a "growth laggard."

IV The significance of Britain's relative economic decline.

The figures presented above clearly illustrate Britain's declining relative macroeconomic performance since the late nineteenth century. On the other hand, becoming a "growth laggard" is of less consequence if a region's competitors achieve faster growth through the expansion of resources - especially labour, in which case faster output growth may be offset by faster population growth. Under these circumstances, variations in relative growth performance reveal little about variations in the relative rates of growth of efficiency and living standards. Furthermore, slow growth in and of itself need not imply a relative decline in absolute standards of living, if the output base from which fast growing regions are expanding is small (see footnote 9, chapter 4). Similarly, relatively slow export growth may give rise to a larger absolute increase in output than that achieved by faster growing competitors, if the initial export base of the latter is relatively small.

However, the relatively slow rate of growth experienced by the British economy since the late nineteenth century has been accompanied by relatively slow efficiency growth and relatively slow per capita income growth, as indicated in sections II and III above. It is also important to note from Table 8 that countries such as Germany and Japan, which have surpassed Britain's share of world exports of manufactures, have subsequently maintained or further increased their export share whilst Britain's has continued to decline. This illustrates that the decline in Britain's export share has not just occurred as a result of competitors experiencing faster export growth from a smaller initial export base; some of these competitors have surpassed the absolute level of British exports and continued to realise faster rates of export growth thereafter.

Slower growth has also acted to the detriment of Britain's relative standard of living. In 1950, output per capita in the UK was higher than in any EEC economy apart from Belgium.[8] By 1973, it was lower than in any EEC economy with the exception of Italy (Cairncross, 1981). Table 9 below further illustrates the decline of Britain's relative standard of living during the post war era.

Table 9: ***GDP per capita in selected OECD economies 1953-84 ($US 1975).***

	1953	1963	1973	1984
UK	3080	3834	5053	5616
France	2432	3476	5437	6362
W. Germany	2319	3866	5628	6473
Italy	1814	2903	4363	4721
Japan	1054	2245	4974	6314
US	4946	5503	7371	8511

Source: Rowthorn and Wells (1987) p.222

These data suggest that the superior growth performance of

[8] The six original members of the EEC from 1957-73 were West Germany, France, Italy, Belgium, the Netherlands and Luxembourg.

Britain's competitors has not merely reflected "catching up." It has relegated Britain to a lower position in the international ranking of economies by per capita income, as the living standards achieved in the "first industrial nation" have gradually been surpassed by its later industrialising competitors.

V British economic history since 1780: an interpretation.

The brief outline of British economic history provided above is congruent with the nature of the model of capitalist development outlined in Part II. The period of complementary export, output and productivity growth 1780-1873, which established Britain as a growth and productivity leader, was not sustained. Instead, what ensued were subsequent periods of relative economic decline which were at first moderate (1873-1914) and later, more dramatic (the post war years), and which saw Britain become an export, output and productivity growth laggard.

This pattern of development is consistent with an interpretation of British economic history in terms of the cumulative interaction of export, output and productivity growth, in an initial virtuous circle which subsequently gave rise to the conditions for its own breakdown, and hence Britain's relative economic failure. The notion that British competitive success in the nineteenth century resulted from a process of cumulative causation is supported by Elbaum (1990).

Elbaum argues that examining the structure of British industry and British trade patterns in the nineteenth and early twentieth centuries reveals that Britain's nineteenth century prosperity and dominance of the world economy are best explained by cumulative advantages related to its status as the "first industrial nation." First, he points to the concentration of British exports in a small number of traditional staple industries with strong interconnections (for example, textiles and textile machinery.)[9] Second, Elbaum points out that these sectors built and retained a dominant position in the British (and indeed world) economy despite lacking any distinctive "comparative advantage" in technology or factor endowments. For example, they were extensive employers of low skilled labour in what was essentially a high wage economy.[10]

This leads Elbaum to suggest that Britain's export and growth success in the staple industries was related to cumulative competitive advantages arising from economies of scale. For example, the cotton industry enjoyed external

---

[9] By the end of the nineteenth century, textiles still accounted for about 40% of the value of British exports (Harley and McCloskey, 1981) whilst as late as 1910, textiles, iron and steel, machinery and shipbuilding comprised 70% of British manufactured exports (Elbaum, 1990).

[10] 63% of value added in the British textile industry was labour cost, and wages were substantially higher than in the textile industries of competing nations, except the US (Elbaum, 1990). Clark (1987, p.14) documents the substantial labour cost advantages that European textile industries enjoyed relative to the British textile industry as early as the 1830s.

economies of scale arising from the combination of its vertical and horizontal specialization with its geographical concentration in Lancashire (Lazonick, 1983, 1986; Farnie, 1979). Individually, small and specialized British cotton firms achieved long production runs, whilst as a whole they produced a wide variety of products. The local development of an organized commodity market for raw cotton allowed spinning firms to keep small inventories and transferred the risks associated with the importing of raw cotton to specialized traders. Finally, the growth of the textile machinery industry in Lancashire reduced the cost of capital for the British cotton industry, and eliminated the need for spinning and weaving firms to hire their own maintenance staff.

Scale economies resulted in efficiency gains for British textile firms which offset their relatively high wages (Elbaum, 1990) and contributed to the low costs per efficiency unit of labour in the British cotton industry (Clark, 1987). This in turn contributed to the export success of the industry, which, if it could not always undersell its competitors (particularly those in the Far East), was consistently able to supply a more varied and higher quality product (Farnie, 1979). Furthermore, the growth of external markets during the nineteenth century encouraged entry into the cotton industry, which tended to reinforce the geographical concentration and specialization of the industry which were the sources of its external economies of scale

(Lazonick, 1983, 1986). The success of the British cotton industry, therefore, seems to have been directly influenced by precisely the cumulative process propagated by increasing returns to scale that was identified in the model of chapter 3.

Evidence also suggests that returns to scale, and hence the cumulative interaction of output growth and efficiency gains, were instrumental in the success of a number of other British industries which grew to dominate world markets during the nineteenth century, such as shipping and shipbuilding (Elbaum, 1990). In the textile machinery industry, an important branch of British engineering, Britain's initial technical superiority in the eighteenth century was subsequently reinforced by the expansion of the industry's markets, which gave rise to dynamic economies of scale within the firm based on learning by doing and specialization (Saul, 1968; Farnie, 1979). Hence

> (t)he market for textile machinery became large enough to permit the separation of the industry both from the cotton industry and from general engineering. (This) reduced the industry's costs by providing full employment for highly skilled and specialized workers and enabled it to supply machinery of the highest quality and the lowest price in the shortest time.
>
> (Farnie 1979, p.55)

As late as 1914, the largest of the six Lancashire firms that dominated the British textile machinery industry still produced an output equal to that of the entire US textile machinery industry (Saul 1968, pp.191-192).

The preceding evidence suggests that the competitive success of a number of key British industries in the nineteenth century was directly related to the cumulative mechanism based on increasing returns to scale that was identified in Part II. However, if Britain's nineteenth century prosperity can be understood in terms of a virtuous circle of cumulative causation, the question arises as to how this growth dynamic broke down, giving rise to Britain's subsequent relative economic failure. In-keeping with the model developed in Part II, the argument that we will pursue is that Britain's virtuous circle of export, output and productivity growth prior to 1873 was corrupted by mechanisms endogenous to the dynamics of this virtuous circle itself. It will be contended that Britain provides a striking example of the notion that success does not breed success indefinitely, but may instead sow the seeds of its own long run destruction.

In order to develop this argument, we will concentrate on two processes which have been instrumental in determining the evolution of the British economy since 1870. The first of these is the problem of technological interrelatedness and hence of lock in to certain techniques and lines of production. This phenomenon was especially evident during the period 1870-1930, having both intra and inter firm dimensions. Technological interrelatedness and lock in affected both the choice of techniques and hence the rate of productivity growth in British industry, and the sectoral composition of Britain's output, to which its export success was also linked in a growing world economy progressing through the commodity hierarchy.

The second process is institutional hysteresis, especially as it has affected the evolution of the British system of industrial relations. This factor has been particularly significant throughout the twentieth century,[11] and has again had two dimensions, related this time to the persistence and procreation of inefficient institutions. Together, these aspects of institutional evolution have tended to be dysfunctional to British economic performance during the twentieth century.

The remaining chapters of Part III are devoted to investigating these causes of Britain's relative economic decline in detail. Chapter 7 examines the prevalence of technological interrelatedness and lock-in in British industry at the turn of the century, and the extent to which this caused inefficient technological evolution in the British economy. Chapter 8, meanwhile, develops a causal chain designed to illustrate the process of institutional hysteresis as it has affected the British industrial relations system in

---

[11] See also Elbaum (1990) and Elbaum and Lazonick (1986) for the argument that institutional factors have been the most important elements in Britain's accelerated relative economic decline in the twentieth century.

the course of the twentieth century. Once again, emphasis is placed on the extent to which factors such as lock in generated inefficient institutional evolution, and how this contributed to Britain's long run relative economic decline.

Chapter 7

Technological Interrelatedness and Lock In in the British Economy, 1870-1930.

By the turn of the twentieth century, the resources of the British economy were concentrated in the production of a narrow range of output emanating from staple industries such as cotton and coal, and in traditional (i.e., mid nineteenth century) techniques of production. In 1907, 46% of Britain's net industrial was accounted for by just three industries - coal, iron and steel, and textiles (Aldcroft 1968, p.23).[1] As was noted in chapter 6, the same industries also dominated Britain's exports. For example, cotton textiles alone made up 25% of Britain's total visible exports in 1910, whilst textiles as a whole accounted for fully 37% (Elbaum 1990, p.1257). At the same time, Britain's manufacturers were slow to adopt new technology. By 1913, for example, new ring frame technology accounted for 87% of all spindles in the US cotton industry, compared to just 19% in the British cotton industry (Lazonick 1986, p.19).[2]

---

[1] Although Aldcroft does not explicitly define the composition of "industrial output," it is clear from the context in which he uses the term that this category comprises the output of the manufacturing and mineral extraction sectors.

[2] In terms of Figure 1 in chapter 4, it appears that in industries such as cotton textiles, Britain's logistic adoption curve was "less squeezed" than that of economies such

The explanation for and implications of this industrial structure for the British economy - what brought it about and how it affected the economy - are controversial subjects. One popular thesis concentrates on the failings of British entrepreneurship in the late Victorian era. It is argued that the maturing of British firms bred an inertia that accounts for the continued use of traditional technologies and the continued production of staple commodities after 1870 (see, for example, Aldcroft, 1964). However, the idea that Britain suffered an "entrepreneurial malaise" during the late nineteenth century, as a result of which the propensity of British firms to undergo changes in their methods and lines of production was significantly weakened, is by no means universally accepted (see, for example, Payne, 1974 and Sandberg, 1981).

An alternative to the theory of entrepreneurial failure is based on the notion that at the turn of the century, the British economy was "over-committed" to the industrial structure that had provided the wellsprings of its nineteenth century success (Richardson, 1965; Aldcroft, 1968. See also Harley and McCloskey, 1981). The term "over commitment" is intended "... to convey the impression that once resources are tied up in certain sectors (technologies) ... there are high costs involved in releasing them for other uses ...; ceteris

---

as the US, indicating a slower rate of adoption of new technology.

paribus, the greater the proportion of resources 'committed,' the more inflexible the economy's structure" (Richardson 1965, p.238).

As illustrated in chapter 3 and again in chapter 6, specialization can be advantageous if it yields increasing returns to scale. However, as illustrated in chapters 4 and 5, extensive growth within a particular institutional and technological context can become disadvantageous. Richardson's hypothesis suggests that by the early twentieth century, Britain's resources were *too* concentrated in a nineteenth century industrial structure. The commitment of British producers to a narrow range of staple industries and traditional techniques of production is seen to have been symptomatic of an inflexibility in the British economy,[3] which thwarted technological and structural change and was thus instrumental in Britain's relative economic decline around the turn of the century.

Richardson's concept of over commitment is clearly similar to the concept of lock in discussed earlier. Both are

---

[3] It is important to note that Richardson originally defined the term "over commitment" in connection with the sectoral composition of Britain's output. However, he acknowledges the importance of technological change within individual industries, and in fact describes over commitment to traditional lines of production as a problem relating to the adoption of new *technologies* associated with new industries. Therefore, it does not involve a great aberration of Richardson's original meaning to extend the use of the term "over commitment" to describe the concentration of Britain's capital stock in certain traditional technologies and associated techniques of production.

intended to capture the notion that structural change in an economy may be constrained by some prior or existing set of production activities. Furthermore, neither rely on the idea of "entrepreneurial failure" to account for phenomena associated with structural change, which, as reasonable as this idea might actually be, is especially difficult to demonstrate.[4]

Building on Richardson, then, in what follows, we will concentrate on the possibility that even if Britain's entrepreneurs were individually rational and well motivated, British firms were locked in to a prior industrial structure which was no longer functional to the maintenance of high growth dynamics. In keeping with the analysis of chapter 4, we will stress the role of technological interrelatedness in generating this lock in. Hence we proceed on the hypothesis that technological interrelatedness and lock in constrained the process of structural change in the British economy. This adversely affected the rates of productivity and export growth, causing the cumulative export, income and productivity growth dynamics which had characterised Britain's economic success prior to 1870 to break down, and giving rise to a period of relative economic decline thereafter.

---

[4] Richardson (1965, pp.240-241) stresses that the over commitment thesis is not contingent on the assumption of irrational behaviour on the part of firms.

I Technological interrelatedness, lock in and over-commitment in the British economy.

As intimated above, two aspects of over commitment or lock in affected structural change in the British economy after 1870. The first aspect was related to the continued use of certain *methods* of production, and the second to the continued manufacture of certain *lines* of production. The first of these two aspects involves the problem of lock in to outmoded technologies and techniques of production.

In chapter 4, it was demonstrated that technological interrelatedness within and between firms can give rise to this type of problem.[5] Hence if the technical components of the production process are interrelated, the introduction of a new technology may involve non-marginal changes which it is rational for the individual firm to avoid within any given period. As a result, firms will be locked in to pre-existing methods of production.

If lock in impairs the rate of adoption of a new technology, and hence the realisation of dynamic scale

---

[5] Notice that interrelatedness *within* firms cannot explain why new entrants in a region fail to adopt the latest technology, and hence why an industry as a whole should remain locked in to an obsolete technology. *Between* firm interrelatedness is therefore important in the explanation of the latter phenomenon, since this affects both new entrants and current incumbents alike.

Note also that between firm interrelatedness begs the question as to why consolidation does not occur in an industry. The issues of market entry and consolidation are addressed in the course of the historical case studies that follow - the latter in some detail in section IIc.

economies accruing to it, the rate of productivity growth within an economy may be arrested. This will corrupt the cumulative dynamics of export, output and productivity growth. Furthermore, if the new technology is adopted faster in competing regions, these may subsequently realise a relatively higher rate of productivity growth. Technological interrelatedness and lock in, which will be manifest in the form of an over-commitment to traditional and obsolete technologies, may then result in the replacement of a virtuous circle of export, output and productivity growth with a vicious circle of cumulative relative economic decline.

To what extent do the theoretical dynamics outlined above provide insights into Britain's relative economic decline after 1870? One way in which to pursue this question is to examine the structure of British industry 1870-1930 in search of examples of technological obsolescence, and identify the ways in which this was related to the phenomena of technological interrelatedness, lock in and poor productivity performance. This approach is not without difficulties, not least because technological interrelatedness is not itself a quantifiable variable. It is also difficult to procure comparative cost data to prove conclusively that the technological interrelatedness of existing plant and machinery was a source of cost disadvantage to British producers competing under conditions of technological change. However, a substantial body of historical evidence points to the importance of technological interrelatedness in British industry. Furthermore, it seems reasonable to assume that costs will increase monotonically with the scale of the changes to existing plant and machinery which are necessary to accommodate a new technology. This suggests that where technological interrelatedness proliferates, it will be a source of cost disadvantage to producers faced with the prospects of technological change.

The following examples provide a pertinent illustration of the extent to which technological interrelatedness and lock in at the industry level contributed to the relative decline of the British economy after 1870. Whilst the number of examples provided may at first appear quantitatively small, recall that the British economy was highly concentrated in a narrow range of industries. Hence it is not unreasonable to draw conclusions about the state of the British economy at this time by examining some of the small number of industries which so dominated its industrial structure.

II The cotton textile industry.

The experience of the cotton industry provides what is perhaps the best example of technological interrelatedness and lock in in the British economy. Although Britain continued to dominate world trade in cotton 1870-1914, the growth of the British cotton industry declined markedly after 1870, and this was coupled with the loss of traditional overseas markets and a decreasing share of world trade in cotton goods. (Tyson, 1968).[6] After 1914, the decline of the cotton industry continued apace, with Britain initially losing lower quality markets to Japan and India, and subsequently losing higher quality markets to the producers of Europe and the US (Lazonick, 1986).

One reason for the climacteric and decline experienced by the British cotton industry is related to the technical obsolescence of the industry dating from the late nineteenth/early twentieth centuries, and its subsequent loss of relative efficiency. Even as late as the 1950s, the dominant technologies in the British industry were the spinning mule and the plain or power loom. These technologies dominated in spite of the availability, by the mid-late nineteenth century, of technically superior alternatives - the ring frame and the automatic loom.

The spinning mule comprised a fixed bank of rollers and a moveable carriage.[7] Spinning took place as the carriage moved away from the rollers, and stopped as the carriage returned to its original position, winding the yarn onto a bobbin as it did so. On the ring frame, yarn passed through a bank of rollers and then around a ring fixed about a rotating bobbin. The ring frame stretched and twisted the yarn into its

---

[6] Britain's share of world trade in cotton goods fell from 82% in 1882-4 to 58% in 1910-13.

[7] These and other details of the spinning and weaving processes can be found in Sandberg (1974).

final form as it was being wound onto the bobbin at the centre of the ring. The advantage of the ring frame over the mule was that it spun continuously, whereas the latter spun only intermittently due to its separation of the spinning and winding processes.

The power loom was originally designed to mechanise the process of weaving,[8] but in fact required close manual supervision for two reasons. First, if a warp thread broke, the power loom continued working. It had to be stopped manually and the break repaired to prevent serious deterioration in the quality of the final cloth. Second, when the shuttle of a power loom exhausted its yarn, the loom had to be stopped and the yarn manually replaced and rethreaded. The automatic loom was purposely designed to overcome these problems. It stopped automatically in the event of warp breakage, and automatically changed and rethreaded the shuttle. These improvements not only meant that automatic looms were required to stop less frequently than power looms, but also that they required less intense manual supervision. Hence a single loom operator could tend more automatic looms than power looms.

The failure of the British cotton industry to adopt newly emerging ring frame and automatic loom technology in spite of its relative efficiency is illustrated by the fact that by

---

[8] Weaving involves passing a shuttle containing yarn (known as weft) between orthogonal strands of yarn (known as warp).

1913/14, only 19% of all spindles in Britain were ring spindles (compared to 87% in the US), and only 1-2% of all looms were automatic looms (compared to 40% in the US) (Lazonick 1986, p.19). Indeed, as late as 1954/5, ring spindles still accounted for less than half of Britain's spinning capacity, and automatic looms only 12% of its weaving capacity, whereas the spinning mule and power loom on which Britain continued to rely had completely disappeared from the US cotton industry (Lazonick 1986, p.20).

Britain's adherence to obsolete spinning and weaving technology involved substantial drawbacks. By 1900, the mule and power loom had essentially been perfected (limiting the extent to which learning by doing could subsequently raise the productivity of this technology), whereas the ring frame and automatic loom constituted an untapped source of technological improvement (Tyson, 1968). Consequently, deprived of the source of dynamic scale economies that were accruing to its competitors, Britain's rate of productivity growth was declining relative to that of countries such as the US even before 1914 (Tyson 1968, p.123). By 1960, labour productivity in the French, German and Dutch cotton industries was 30-60% higher than in the British industry (Lazonick 1986, p.20).

If one source of the competitive decline of the British cotton industry was thus its technical obsolescence and consequent inefficiency, the question arises as to why the industry continued to rely on obsolete technology? Why was the adoption of the ring frame and the automatic loom so slow in Britain relative to the rates of adoption achieved by its emerging competitors such as the US? Historical evidence suggests that one factor contributing to the continued use of the spinning mule and power loom was technological interrelatedness, which subsequently locked the British cotton industry in to its nineteenth century industrial structure.

a. Technological interrelatedness within cotton firms.

Both Frankel (1955) and Tyson (1968) suggest that technological interrelatedness within the firm were responsible for the slow adoption of the automatic loom in the British cotton industry. The substitution of automatic looms for power looms entailed not just the replacement of one machine with another, but usually involved the complete redesign of weaving sheds (Tyson, 1968). For example, the introduction of automatic looms required the strengthening of flooring, the elimination of pillars, and the respacing of other machinery (Frankel, 1955). This suggests that the changes necessary within established British weaving firms to accommodate the automatic loom were of a non-marginal nature, requiring complementary changes in the other components of nineteenth century plant and equipment in which the British cotton industry had invested heavily. Other things being equal, technological interrelatedness may thus have made automatic loom technology prohibitively expensive for British weaving firms to adopt. The consequent lock in of British firms to a prior (and obsolete) technology would help explain the slow rate of adoption of the automatic loom, and hence the declining relative efficiency of the British cotton industry.

b. Technological interrelatedness between cotton firms.

Even if *existing* British cotton firms were locked in to obsolete technology due to intra firm interrelatedness, this cannot explain why *new entrants* into the industry were not faster in the adoption of new technology. For this reason, it is important that it is also possible to identify *inter* firm dimensions of technological interrelatedness within the British cotton industry. This source of interrelatedness would have affected the choice of technology of both new and existing firms.

The nineteenth century industrial organization of the cotton industry was typified by a high degree of both horizontal and vertical specialization (Lazonick, 1983). This high degree of specialization appears to have hindered the introduction of new spinning and weaving technology in the late nineteenth and early twentieth centuries. For example, because mule yarn was wound on lightweight paper tubes whereas ring frame yarn was wound on heavier wooden bobbins, the latter was more expensive to ship between geographically separate spinning and weaving firms (Lazonick, 1986). Given the industrial organisation peculiar to the British cotton industry, then, ring frame technology may not have been the least cost method of production for individual British spinning firms.

More importantly, the introduction of new weaving technology required complementary changes in spinning firms (and vice versa) if the industry as a whole was to successfully convert from its traditional low throughput to the new high throughput technology (Frankel, 1955; Lazonick, 1986). For example, Automatic looms required weft yarn in larger packages than spinning mules, the dominant technology in spinning firms, were able to produce (Frankel, 1955).

Furthermore, Lazonick (1981) argues that vertical specialization may have deterred investment in high throughput weaving technology by creating uncertainty about the supply of and demand for yarn of differing qualities. An important feature of the Lancashire cotton industry was that spinning firms purchased raw cotton week by week as orders for finished yarn were placed. This, coupled with the fact that the spinning mule was capable of spinning cotton of a wide range of different qualities,[9] allowed British spinning firms to take advantage of even short term changes in the relative prices of cotton of different qualities. Uncertainty in the short term concerning the quality of yarn may therefore have deterred weavers from investing in the automatic loom, which

[9] The mule placed less strain on yarn during spinning, and was consequently capable of spinning a wider range of cotton qualities than the ring frame.

required yarn of a high quality for its successful operation. The consequent continued demand for mule yarn by British weavers may, in turn, have deterred spinners from investing in ring frame technology, due to the lack of any apparent market for yarn of a consistent high quality (Lazonick, 1981).

The technology of spinning firms was obviously beyond the control of individual weaving firms, and vice versa. Hence regardless of the cost efficiency of the new versus the old technology, the small size of the individual decision making unit (the individual spinning/weaving firm) relative to the size of the technological unit which required change (current weaving *and* spinning technology) hindered the adoption of new technology (ring frames and automatic looms). In this case, then, inter firm technological interrelatedness locked the British cotton industry into obsolete nineteenth century technology by virtue of the absence of a centralized decision making authority amongst vertically specialized spinning and weaving firms.

c. Inter firm interrelatedness and the issue of consolidation.

The disadvantages of horizontal and vertical specialisation in the cotton industry raises the question as to why spinning and weaving firms *were* so specialised. One reason for this was specific to the structure of the cotton industry itself. By the mid nineteenth century, the extensive degree of existing specialisation inherited as a legacy of

Britain's early industrialisation afforded considerable external economies of scale. These in turn encouraged new nineteenth century entrants into the cotton industry to replicate the pattern of specialisation (Lazonick, 1983). Hence the prior structure of the Lancashire cotton industry may have deterred the formation of integrated firms in the nineteenth century which would subsequently have been better suited to the adoption of new spinning and weaving technology.

Even so, this cannot explain why the cotton industry *remained* specialised after the advantages of integration became apparent. Why did British cotton firms not imitate trends in the German and US economies towards the consolidation and rationalisation of industry? One possible answer, which applies not just to the cotton industry but to the entire British economy, concerns the activities of The City in relation to British industry. Best and Humphries (1986) argue that whether or not the British financial sector differed from its US and German counterparts with respect to its attitudes towards risk and uncertainty, anti-industrial biases, or lack of information, it did differ by failing to participate directly in the restructuring of domestic industry in the late nineteenth century. Whilst German credit banks and US investment banks promoted combination and rationalisation in the German and US manufacturing sectors, British financial developments in the late nineteenth century provided no such spur for integration. On the contrary, Best and Humphries argue that by providing easy access to funds for new capacity, the British financial sector encouraged entry, thus exacerbating competitive pressures in product markets. This forced existing firms to increase dividends in the short term in order to attract new external capital, thus preventing them from amassing retained earnings which might otherwise have accumulated into a viable source of internal financing for acts of consolidation.

Whilst it is beyond the scope of this thesis to fully examine the controversial relationship between British finance and industry, the arguments above suggest that the failure of the British financial sector to play as active a role in the re-organisation of industry as its US and German counterparts may have been a key factor inhibiting widespread consolidation in British industry. This may, in turn, explain the continued high degree of specialisation exhibited by specific industries such as cotton, and hence the continued problems of inter firm technological interrelatedness suffered by these industries.

III The coal industry.

Although an extraction rather than a manufacturing industry, evidence suggests that the application of new technology played an important role in the late nineteenth/early twentieth century development of the coal industry. Furthermore, the adoption of this new technology in Britain again appears to have been impaired by technological interrelatedness and lock.

Despite unprecedented growth in the exports of British coal between 1875 and 1913, the factors which would subsequently lead to the diminution of Britain's share of world coal exports in the face of US and German competition were already evident during this period. Taylor (1968) argues that Britain's initial dominance of world trade in coal was based largely on geographical chance,[10] and that even as it enjoyed its greatest successes as an exporter, Britain's relative efficiency as a coal producer was being eroded. This trend is most evident after 1885, when the British coal industry experienced a negative trend rate of productivity growth (Taylor 1968, p.46). This loss of efficiency cannot be explained by the underemployment of mining capacity, as the years in which it was chiefly sustained correspond to periods of heavy demand for coal (Taylor, 1968).

However, it may be explained by the failure of the industry to adopt new technology. Coal mining comprises four main operations: the hewing of coal at the face; the conveying of coal to the main roadways; the carrying of coal to the pit bottom; and the raising of coal above ground. In the mid nineteenth century, the chief technological advances made by the industry were in the carrying of coal to the pit bottom,

---

[10] The ease of sea communication between the coalfields of South Wales and North East England and the markets of Europe and South America gave Britain a substantial cost advantage in haulage at a time when transporting coal over land was expensive. See Taylor (1968, p.41).

and British mines were amongst the foremost innovators during this period (Taylor, 1961). However, in the late nineteenth century, the main technological advance took place at (or near) the coal face - the development of the mechanical cutter and conveyor. British mines were slow to adopt this technology. Indeed by 1913, only 7.7% of Britain's total coal output was mechanically cut (Taylor 1968, p.57).[11]

Why was the British coal industry slow to adopt mechanical coal cutting technology? Evidence suggests that part of the explanation is that technological interrelatedness in the coal extraction process locked the industry in to its existing technique of production. Hence

> the efficient large scale use of cutters and conveyors demanded a reorganization of activity not only where the coal was hewn but throughout the cycle of operations from face to surface.
>
> (Taylor 1968, p.58)

For example, the mechanisation of coal cutting required the retraining of labour (Taylor, 1968), and so the abandonment of workers existing skills in which mine owners had previously invested. Furthermore, there existed problems with the supply of power. Up to 1905, compressed air was the most common form of power employed underground in British coal mines (Taylor,

---

[11] By way of comparison, it may be noted that by 1900, 20% of the output of the US coal industry was machine cut (Taylor 1968, p.56).

1961). This created problems for mechanisation of any kind, since transmission of this source of power was difficult, and the pressure achieved was often below necessary requirements (Taylor, 1961). The adoption of mechanical coal cutters and conveyors in British mines therefore required not only investment in the relevant machinery, but also an accompanying change in power generating technology from compressed air to electricity.

These details suggest that the technical changes required to accommodate mechanical cutters and conveyors were non-marginal in nature, necessitating accompanying changes in other components of existing plant and equipment, and in the human capital of workers. The costs associated with these changes may have made them appear unprofitable compared with continued use of the existing (obsolete) technique of production. This is despite the fact that the adoption of mechanised technologies appear to have been the most profitable option in less developed economies such as the US, which would have suffered less technological interrelatedness as a legacy of past mining activity.

It is important not to overlook factors such as geological difficulties in the twentieth century decline of the British coal industry. In the Lancashire coalfield, for example, where thinning seams were otherwise appropriate for the application of the mechanical coal cutter, geological faults and the pitch of the seams made mechanisation at the coal face impossible (Taylor, 1968). However, there is evidence to suggest that lock in to an outmoded technique of production, which was a legacy of Britain's early start as an industrial nation, contributed to the decline of British coal mining.

IV The chemical industry.

The problems of technological interrelatedness in the British economy emerged not only in traditional staple industries such as cotton and coal, but also in new industries such as chemicals.[12] Again, the legacy of an initial technology selection appears to have been a crucial factor in the subsequent development of certain of these industries.

Richardson (1968, p.278) argues that whereas before 1880, Britain had been the dominant producer of chemicals, by 1913 it ranked only third in terms of its share of world production, well behind the US and Germany.[13] Britain's declining fortunes as a chemical producer in the late nineteenth/early twentieth centuries are illustrated by the loss of its pre-eminence as an alkali producer. This branch of the industry was characterised by competition between the Leblanc and Solvay processes. The Leblanc process involved the

---

[12] The term "new industries" is used to denote those industries, such as the chemical and motor vehicle industries, which emerged in the late nineteenth century.

[13] The relative positions of these nations did, however, vary from product to product.

oxidisation of hydrochloric acid into chlorine using manganese dioxide. The Solvay process, meanwhile, involved the passing of carbon dioxide into salt and amonia in order to precipitate sodium bicarbonate, which was subsequently converted into soda.

Throughout the late nineteenth and early twentieth centuries, British alkali firms relied largely on the Leblanc process, despite the fact that as early as the 1870s, alkali produced by the newly emerging Solvay process could undersell Leblanc output by up to 20% (Richardson 1968, p. 284). Once again, the competitive decline of a British industry can be linked to its continued use of an obsolete technology; and once again, this may in turn be explained by technological interrelatedness and lock in. The Leblanc process emerged historically prior to the Solvay process, so that by the time the latter did emerge, Britain already had a sizeable alkali industry based on Leblanc technology. In contrast, the German and US industries developed largely after the Solvay process became available. Switching from the Leblanc to the Solvay process would have necessitated the complete reconstruction of existing British chemical plant. Such non-marginal changes for British chemical producers would, in turn, have necessitated the large scale scrapping of existing (and relatively new) plant. This would likely have made the costs of switching between processes prohibitively high for existing British manufacturers, effectively locking them in to an obsolete technology inherited as the legacy of an early start, and which constituted an obstacle to technological change that did not exist for late starting US and German producers.

The lock in of existing British chemical producers to the Leblanc process was partially offset by the adoption of the Solvay process by "new entrants" into the industry. For example, in 1893 the United Alkali Company, a Leblanc combine, opened a new plant utilising the Solvay process (Richardson 1968, p.282). Nevertheless, adoption of the Solvay process in Britain was extremely slow. In 1883, Solvay output accounted for just 12% of UK alkali production, compared with 44% in Germany and 100% in the US. British Leblanc producers prolonged their survival by improving their production process to reduce costs,[14] and most importantly, by specialising more in the production of bleaching powder and caustic soda.[15] However, with the rise of the electrochemical industry, the loss of its markets for caustic soda sent the British Leblanc industry into terminal decline. The industry had closed down completely by 1920.

[14] By 1890, the price of Leblanc alkali was just one third of its level in the early 1870s (Richardson 1968, p.284).

[15] This allowed Leblanc producers to differentiate their product from that of their Solvay competitors, who produced mainly pure soda.

V Technological interrelatedness and lock in to the production of certain commodities.

The second aspect of over commitment, and the one given primary emphasis by Richardson (1965), relates to the narrow range of output produced by the British economy and its concentration in traditional staple industries such as cotton and coal.

Lock in to the production of certain commodities can occur if the initial commitment of resources to specific interrelated lines of production subsequently retards the re-allocation of resources towards newly emerging industries. For example, if a region's transport system and/or its geographical distribution of labour have evolved in such a way as to cater for the needs of a specific block of existing industries (i.e., if its industry, physical infrastructure and stock of human capital are technologically interrelated), re-allocating resources towards new industries with different transport and labour requirements will involve both investment in these industries, and changes to the existing infrastructure and stock of human capital (which may in turn, of course, affect the existing industrial base). As such, this process of resource re-allocation will require non-marginal changes in the structure of a region's economy, which may make it prohibitively expensive compared to the alternative of continuing to work existing lines of production. Alternatively, the type of resource reallocation necessary for the growth of new industries may be beyond the control of individual decision making units within these industries. Either way, a region may become locked in to its existing industrial structure, committing resources to new industries at a slower pace than regions with a less developed prior industrial structure, where any necessary adjustments to the infrastructure and/or the technologies of existing traditional industries are easier to accommodate. This time, lock in will manifest itself in the form of an over-commitment to "traditional" industries within a region.

Lock in to traditional industries may have adverse effects on the growth dynamics of a region. Recall that in chapter 3, the growth of a capitalist economy was described as a process of structural transformation, involving progression through a "commodity hierarchy." If a region becomes locked in to certain traditional industries, its progress through this commodity hierarchy will be hindered. The composition of its output and hence exports will become dominated by "necessities," and tend not to reflect those commodities with the highest income elasticities of demand. The region will effectively be producing the "wrong" goods, in the sense that the demand for what it does produce will not be growing as rapidly as the demand for the output of newer industries to which it is failing to commit resources.[16] As the rate of

[16] Although this possible consequence of interrelatedness was not formally analysed in Part II, note that it will affect the causal chain of a virtuous circle by reducing the value of

growth of output subsequently falls, and the realisation of dynamic scale economies is retarded, the aggregate rate of productivity growth in the region will decline relative to that achieved in other regions. Hence lock in to traditional lines of production in a world economy which is moving through the commodity hierarchy may corrupt and even relatively reverse the high growth dynamics of a virtuous circle.

a. Over-commitment to traditional industries in the British economy.

In order to determine whether these mechanisms influenced Britain's relative economic decline after 1870, we need to identify to what extent Britain was producing the "wrong" goods - i.e., to what extent it was over-committed or locked in to the production of goods emanating from a narrow range of traditional (staple) industries.

The commitment of the British economy to a narrow range of output was illustrated at the beginning of this chapter. It was noted that in the early twentieth century, three industries (coal, iron and steel, and textiles) accounted for almost half of Britain's net industrial output, whilst textiles alone made up 37% of total visible exports (Aldcroft, 1968; Elbaum, 1990). Whilst industrial concentration is not, in and of itself, indicative of structural weakness in an economy, evidence suggests that the narrow range of industries

---

γ_j in equation [3.3] over time.

which dominated the output of the British economy were concentrated in declining sectors of the international economy (Aldcroft, 1968). Hence in 1913, contracting export industries accounted for 62.5% of Britain's total manufactured exports, compared to just 49.5% in Germany and 37.3% in the US (Svennilson 1983, p.295).[17]

Furthermore, there is evidence to suggest that over-commitment to declining industries played a direct role in diminishing Britain's share of world trade in manufactures. Tyszynski (1951) argues that a country's share of world trade is determined by both the relative importance in world trade of the commodity bundle that it exports, and its competitiveness in individual product markets. Hence it is possible to decompose a change in a country's share of world trade into that part which is due to changes in the structure of world trade, and that part which is due to changes in the country's international competitiveness. Utilizing this "export accounting" framework, Tyszynski (1951, p.289) finds that between 1899 and 1937, declining international competitiveness in all industries provides the chief explanation for Britain's declining share of world trade. However, he finds that the composition of British output during this period (reflecting the concentration of Britain's

---

[17] Svennilson applies the term "contracting industry" to those industries whose share of total world trade in manufactures exhibited a declining trend over the period 1913-50.

productive resources in traditional industries) reduced Britain's share of world trade in manufactures by 1.5 percentage points.[18] This suggests that the composition of Britain's export commodity bundle contributed directly to its declining importance in world trade in the late nineteenth/early twentieth centuries - i.e., that Britain was effectively producing the "wrong" goods.

b. Lock in to traditional industries: interrelatedness between industry and infrastructure.

Evidence also suggests that lock in provides at least part of the explanation for Britain's continued devotion to traditional industries in the late nineteenth/early twentieth centuries. Richardson (1965) argues that there existed technological interrelatedness between Britain's staple industries, and its power and transport facilities. The latter, as a result of Britain's early start to industrialisation, chiefly comprised coal, steam and gas, and railways respectively. Hence the development of the electricity industry in Britain was retarded by the commitment of its existing industrial base to more traditional sources of power. This in turn impeded the development of new industries such as aluminium, which were dependent on electricity. Hence the non-marginal changes necessary not only to Britain's

---

[18] The total decline in Britain's share of manufactured exports between 1899 and 1937 was 10.1 percentage points.

infrastructure, but also to the plant of existing staple industries in order to make the new industries, the traditional industries, and the infrastructure of the British economy mutually compatible formed a barrier to the development of new industries. Not only is it possible that changes of this magnitude would have been prohibitively costly for the British economy to undertake; it is also clear that they would have been extremely large relative to the size of the individual decision making unit in the British economy at that time. Meanwhile, Britain's late starting competitors developed staple and new industries simultaneously and, free of the legacy of an early start, an infrastructure well suited to both.[19]

c. Lock in to traditional industries: interrelatedness between industry and labour force.

The transfer of resources from old to new industries in Britain was also complicated by interrelatedness between the location of the staple industries and the location of Britain's industrial labour force. The regional concentration of new industries in the Midlands and the South East meant that their development required not just the occupational transfer of labour, but also its geographical reallocation.

---

[19] Many of the traditional industries which were steam powered in Britain could be powered (often more efficiently) by electricity, making their power demands compatible with those of new industries such as aluminium.

This would not only have increased the costs associated with setting up new industries in Britain relative to continued investment in traditional industries. It would also have impeded the transfer of resources towards new industries in Britain by virtue of the fact that by the early twentieth century, labour was a far less geographically mobile factor of production than it had been when the staple industries were first expanding during Britain's initial period of industrialisation.[20] Richardson (1965) argues that despite extensive north-south migration, the location of the labour force hindered the physical transfer of labour resources from old to new industries in Britain. Once again this suggests that a form of technological interrelatedness - this time between the location of industry and the location of a key input (labour) in the process of production - may have resulted in Britain becoming locked in to its prior industrial structure. British industry remained concentrated in a narrow range of traditional lines of production at a time when the growth of world income, and hence movement through the commodity hierarchy, demanded structural change in the composition of final output as a necessary condition for competitive survival.

---

[20] The loss of common land due to the enclosure movement, and the subsequent loss of livelihood experienced by rural workers was a key factor promoting labour mobility during Britain's initial industrialisation.

VI Conclusion.

This chapter has emphasised the failure of the British economy in the late nineteenth/early twentieth century to undergo structural changes related to methods of production, and the composition of final output. We have argued that the phenomena of technological interrelatedness and resultant lock in help explain this failure. Furthermore, it has been suggested that structural weaknesses related to the methods and lines of production in which British productive resources were concentrated played a role in causing the dynamics of Britain's mid-Victorian prosperity to break down. This, in turn, suggests that these weaknesses contributed to Britain's relative economic decline after 1870.

This hypothesis is of a complicated nature, and the preceding analysis cannot claim to have tested it exhaustively. Neither the descriptive statistics nor the qualitative historical evidence which have been called upon can be regarded as conclusive. Indeed, it cannot be overlooked that some authors have argued that the failures resulting from the slow pace of technological change in the British economy were of a modest nature. Hence Sandberg (1981) suggests that British firms' continued use of nineteenth century technology was rationally motivated, and that the resulting loss of efficiency in British industry relative to its US and German competitors was modest.[21]

However, none of the arguments presented in this chapter are necessarily contrary to the postulate that British firms acted in an individually rational manner. On the contrary, the chief determinant of British technological obsolescence is taken to be the *constraints* placed on firms by the prior industrial history of the British economy, rather than any shortcomings in the *motivation* of these firms. Furthermore, the evidence presented suggests that efficiency impairing failures related to the slow adoption of new methods of production were in evidence in the British economy after 1870.[22] Indeed, even if Sandberg is correct to point out that the extent of these failures and their subsequent impact on British economic performance was modest, it will be recalled from chapter 6 that so, too, was the extent of Britain's relative economic decline during this period. With only a modest reversal of Britain's relative economic position to explain, it is not unreasonable to assert that modest failings relating to technological change may help account for this reversal.

Finally, Sandberg does not address the issue of the

---

[21] This conclusion draws particularly on his extensive study of the decline of the British cotton industry. See Sandberg (1974).

[22] See also Lazonick (1990, pp.157-62) for the argument that Sandberg significantly underestimates the extent to which productivity growth declined in British industry from the late nineteenth century onwards.

composition of Britain's output, which has been discussed above, and to which authors such as Richardson (1965) attach importance. We may therefore conclude that the evidence presented in this chapter lends support to the idea that after 1870, Britain became over committed or locked in to methods and lines of production inherited as a legacy of the past, and that this had adverse implications for its relative economic performance.

Chapter 8.

Institutional Hysteresis and the Performance of the British Economy After 1914.

Chapter 5 stressed that in a model of institutional hysteresis, institutions are best regarded as evolutionary, path dependent phenomena which may not be efficient in the sense of being functional to the dynamics of a capitalist economy. It was shown that inefficiency may arise due to the persistence of inefficient institutions over time, or by virtue of the procreation of inefficient institutions.

In this chapter, we will argue that the evolution of institutions in the British economy has been a path dependent process which has displayed both of the sources of inefficiency mentioned above. This in turn has contributed to the continued and eventually accelerating relative decline of the British economy in the twentieth century.[1] In order to support this hypothesis, we will concentrate on the evolution of a set of institutions which, as was shown in chapter 5, has a direct effect on the performance of a capitalist economy - the industrial relations system.

---

[1] Emphasis is placed on the twentieth century because, as noted earlier, it has been the concern of a number of authors to suggest that institutional factors were the single most important source of Britain's competitive failure during this period. See, for example, Elbaum (1990) and Elbaum and Lazonick (1986).

Throughout the following analysis it is important to bear in mind the distinction made in Chapter 5 between *organizations* (such as firms and trade unions) and the set of *institutions* that defines the relationship between them (the industrial relations system). Following Dunlop (1958), the industrial relations system can be identified as comprising three sets of actors - the state, employers and their associations, and employees and their associations. The interaction of these actors produces the "web of rules" governing the employment relationship that is the essence of the industrial relations system. Our interest, then, is in the evolution of a set of *institutions* (the industrial relations system) which may be influenced by, but is nevertheless distinct from, the existence of certain labour market *organizations* (such as trade unions). Comments on the relative efficiency of the former should not, therefore, be misconstrued as comments on the desirability (or otherwise) of the latter.

What follows is a history (including industry level case studies) and analysis of twentieth century British industrial relations. This reveals that changes in the pattern of industrial relations in Britain during the this period have been largely superficial. Systemic change has been lacking; instead, what we observe is the persistence of nineteenth century industrial relations practices, which have not been conducive to technological and organisational change, and hence the dynamic efficiency of British industry in the twentieth century. Evidence suggests that institutional interrelatedness and lock in provide part of the explanation for this inefficient institutional evolution.

I A Brief History of British Industrial Relations in the Twentieth Century.

In order to analyse the evolution of British industrial relations, it is necessary to first outline their history. This can be broken down into three broad epochs: the prewar and interwar years; the post war years up to the late 1960s; and the period since the late 1960s.

a. The prewar and inter war years.

The first two decades of the twentieth century were part of a period of consolidation for the British labour movement. This followed attacks on its very legitimacy at the turn of the century, such as the Taff Vale Decision (Pelling, 1972).[2] Between 1914 and 1920, trade union membership doubled. However, the evolution of the industrial relations system at this time was influenced less by the growing strength of labour than by the state of national emergency caused by World War I. The conditions emanating from the War heralded a brief

[2] The Taff Vale Decision ruled that trade unions were liable for damages brought about by their sponsorship of industrial action. It effectively undermined the financial feasibility of strikes.

phase of close co-operation between the state, labour and capital (Pelling, 1972). For example, Labour Party politicians held positions in the coalition governments during the war, including positions in the War Cabinet after 1916. In the industrial arena, tripartite agreements limited the number of trades disputes in key industries. In the "Shells and Fuses Agreement" of 1915, unions permitted the introduction of unskilled labour on jobs traditionally reserved for skilled workers, in return for government and employer pledges that traditional practices would be restored in good faith after the war (Pelling 1972, p.151).

These conditions of co-operation did not extend into the interwar period, which witnessed considerable social unrest. Heavy unemployment in the early 1920s was followed by the General Strike of 1926, which lasted nine days and involved one million miners and one and a half million transport workers, printers, construction workers, and iron and steel, chemical and power industry workers (Pelling 1972, p.175). In the aftermath of the General Strike, the 1927 Trade Disputes and Trade Union Act curbed the power of unions to engage in sympathetic strike action, a measure which was resented within the labour movement as a curtailment of its rights. Furthermore, in the light of the mass high unemployment of the Great Depression, trade unions became increasingly antagonistic to the orthodox (balanced budget) fiscal policies pursued by the state. On the whole, the period between 1920 and c.1935 was not a good one for British industrial relations, characterized as it was by significant breakdowns such as the General Strike, and feelings of mutual hostility and suspicion between trade unions, employers and the government. This is partly reflected by the high value of Shalev's (1978) index of relative strike volume for 1927-32, as reported in Table 1 below.[3]

Table 1: *Index of relative strike volume in the UK 1927-72 (1968-72 = 0).*

	1927 -32	1933 -37	1948 -52	1953 -57	1958 -62	1963 -67	1968 -72
Strike volume	80	35	25	45	60	35	100

Source: Shalev (1978) p.15.

[3] Following Hibbs (1978), we can define three dimensions which make up the profile of a strike: size (the number of strikers per strike); duration (the number of man days lost per striker); and incidence (the number of strikes per thousand workers). It is widely agreed that strike volume - the product of these three dimensions - is the most appropriate aggregate indicator of strike activity (see Hibbs, 1978).

Note, however, that the volume of strike activity is an imperfect indicator of the state of industrial relations. As Naples (1981, 1986) notes, strikes are only one form of industrial action - phenomena such as absenteeism and even high quit rates may also be symptoms of labour unrest. Furthermore, a low level of strike volume may indicate trade union weakness rather than capital-labour co-operation. It is important to bear these shortcomings in mind when assessing the contents of Table 1.

b. World War II and the emergence of the "post war consensus."

Ironically, it was once again the threat of a war which began the reversal of these conditions. As early as the mid 1930s, there is evidence of a growing acceptance of trade unionism in Britain, and a consequent willingness of employers to negotiate with labour. Pelling (1972, p.211) accredits the growing "respectability" of trade unions during the 1930s to the pragmatism and constitutionalism of prominent union leaders such as Ernest Bevin. Whatever the source of this growing toleration of trade unionism, it resulted in a significant reduction in the amount of time lost due to industrial disputes in the mid 1930s relative to the preceding era. This is reflected in the fall in the Shalev index in Table 1 for 1933-37, and by the fact that between 1934 and 1939, the number of working days lost to disputes exceeded two million only once, whereas it was never less than seven million between 1919 an 1926 (Pelling 1972, p.211)

By the late 1930s, the preparations for, and especially the outbreak of World War II posed sufficient threat to national security to enforce co-operation in industrial relations. In 1940 the number of working days lost to strikes was the lowest since records began in the late nineteenth century (Pelling 1972, p.218).[4] However, co-operation was

---

[4] War and the preparations for it also drastically reduced the rate of unemployment, a major source of social strife in interwar Britain. Unemployment fell from a peak of 23% in 1932 to 9% by 1937, although it then increased again between 1938-9 to 12% "in spite of re-armament" (Pollard 1962,

never absolute, even during the War. The growing number of unofficial strikes is illustrated by the fact that more working days were lost to strikes in 1944 than in any year since 1932 (Pelling 1972, p.221).[5]

Following World War II, the task of rebuilding the British economy and the austerity measures this necessitated led to renewed co-operation between capital, labour and the state. This spirit of co-operation, borne partly of necessity and partly as a result of a rejection of the conditions of the 1930s,[6] came to characterize the social consensus which accompanied the post war "golden age" of growth and full employment. This is reflected in figures for relative strike volume between 1945 and 1960. The Shalev index in Table 1 falls to 25 and 45 during the periods 1948-52 and 1953-57 respectively, compared to an index of 80 for the strife torn years between 1927 and 1932. However, even during this era of co-operation in industrial relations, no form of "social corporatism" emerged in Britain. Trade unions remained many and varied, and no strong, centralized, representative bodies

---

p.243). It would have likely continued to increase had it not been for the outbreak of World War II, following which the unemployment fell to the extent that it had been "virtually abolished" by 1941 (Pollard 1962, p.347).

[5] The number of unofficial strikes during this period was also indicative of a problem that was to re-emerge in subsequent periods - the battle for control of the labour movement between national leadership and shop stewards.

[6] In 1945, the Labour Party won what remains the largest ever majority in the House of Commons in British parliamentary history.

of trade union and employer interests emerged to fully co-ordinate industrial relations at the national level. For example, within the labour movement, even individual unions within a specific industry frequently struggled to contain the strength, independence and militance of shop stewards at the level of the firm (Pelling, 1972). The tradition of "voluntarism" in industrial relations also remained (Pelling, 1972; Crouch, 1978), restricting the role of the state to conciliation rather than detailed coordination. The fragile nature of industrial relations, even during this era of increased cooperation, is illustrated in Table 1 by the general upward trend in relative strike volume throughout the post war era. This was accompanied by uniformly upward trends in the incidence, median duration and median size of strikes, as indicated by Table 2 below.

Table 2: ***Incidence, duration and size of strikes in the UK, 1959-73.***

	1950-58	1959-63	1964-67	1968-73
Incidence[a]	27	50	67	115
Median Duration[b]	1.3	1.5	1.7	2.6
Median Size[c]	-	58	80	115

Notes: a. Strikes per million employees, excluding mining.
b. Number of man days lost per striker.
c. Number of strikers per strike.

Source: Shalev (1978) p.17

c. The decline of the "post war consensus."

Although the objectives of British trade unions have traditionally been modest and narrow, the decade of the 1960s witnessed the emergence of new demands from the labour movement (Crouch, 1978). Whilst traditional concerns such as wage increases continued to dominate the agenda, trade unions began to show greater concern over issues such as the quantity of employment, resisting redundancies and demanding to negotiate the terms of layoffs. For example, most productivity agreements by the late 1960s were negotiated to ensure no loss of manpower, or at least that redundancies would be voluntary and compensated (Crouch, 1978).

New egalitarian concerns with the plight of low paid workers also emerged. The Transport and General Workers Union notion of "social justice bargaining," based on demands for a 16 pounds and 10 shillings minimum wage, was taken up by the Trades Union Congress (TUC) in 1969. The late 1960s also witnessed an increase in pay settlements which concentrated on the lowest paid, and "tapered" as they progressed up the income scale (Layton, 1973). Finally, the campaign for the low paid was matched by an increase in the militancy and demands of these workers - especially in cases where low paid workers comprised "new" demographic groups within the labour force, such as women.

Trade unions also began to make new demands with respect to qualitative, non-wage aspects of the labour process - in particular, in connection with control over the organization of work (Brown, 1972). For example, in 1967 the TUC began formally advocating worker representation on company boards. Finally, the labour movement began to show an increased awareness of the broader political issues related to its actions. The traditional division between the "industrial" and "political" wings of the British labour movement began to fade, as the TUC engaged in such activities as the formulation of its own economic policies, complete with growth targets and policies designed to achieve them, and policy initiatives on the welfare state.

By the late 1960s and early 1970s, the decline of the post war industrial relations consensus was becoming increasingly evident in the increased industrial conflict in the British economy. Indices of the number of strikes, the number of workers involved in strikes, and the number of working days lost all indicate an increase in industrial conflict occurring between 1966 and 1975 (Crouch 1978, pp.201-202). These trends are clearly reflected in Tables 1 and 2, which indicate that the periods 1968-72 and 1968-73 respectively were peak years for strike activity, whether measured by volume, size, incidence or duration.

The most significant response to these developments was that of the state, which Coates and Topham (1986) argue was conditioned by the rapid rise of the New Right in British politics. This led to the rejection of post war efforts to assimilate and diffuse industrial conflict in the political arena. Following the General Election of 1979, a succession of legislation including the Employment Acts of 1980 and 1982 and the Trade Union Act of 1984 imposed an unprecedented degree of state control over internal union processes, by legislating rules for the election of union leaders, the decision to strike, and the accumulation and use of political funds.[7] Legislation of this nature, deliberately aimed by the State at the allegedly unrepresentative and unaccountable power of union leaders brought about a further, dramatic breach in co-operation between labour, capital and the state, and constituted a total rejection of even the comparatively modest terms on which this cooperation had previously been based.[8]

II The evolution of the industrial relations system and its effects on British economic performance.

The outline above suggests that although there have been changes in the character of British industrial relations during the twentieth century, industrial relations in Britain have never been systematically resolved in a co-operative

---

[7] These Acts also affected other aspects of union activity, such as the legality of the Closed Shop.

[8] Recall that even during the post war years, the British economy did not develop social democratic industrial relations structures, such as those which typified the Scandinavian economies, or "plant level corporatist" institutions such as those defining capital-labour relations in Japanese firms. In fact, Britain made no radical break from the structure of the industrial relations system it had inherited from the late nineteenth/early twentieth centuries.

manner. In light of the BGW effect, according to which technological and organisational change in an economy depends on the compliance of labour in the production process, this suggests the possibility of inefficiency in the British industrial relations system. As illustrated below, industry level studies provide concrete examples of the inhibiting effects of the structure of industrial relations on technological and organisational change, and consequently on the relative efficiency of British industry. These provide evidence in favour of the notion that Britain has experienced inefficient institutional evolution in the sphere of industrial relations during the twentieth century.

a. Inefficient institutional evolution in the cotton industry.

A central feature of industrial relations in the cotton industry during the early and mid twentieth century was the system of wage lists. Negotiated at the district level, the wage structures enshrined in these lists were designed to adjust pay to compensate workers for the amount and difficulty of work done. For example, the Oldham list, first adopted in 1876, provided spinning mule minders with standard weekly earnings plus a "quick speed" allowance, which compensated them for time during which the mules were operating faster than a negotiated standard speed (Lazonick, 1990). The Oldham and Bolton lists, the most important of the district wage lists, determined 75% of mule spinners wages in 1894, and 90%

by 1945 (Lazonick, 1979).

Although the wage lists tended to reflect the local balance of power between labour and firms at the time of their negotiation, there may be some extent to which they originally constituted "efficiency wages" in the late nineteenth century. Cotton industry operatives constantly feared that higher output per man hour resulting from greater work intensity would lead management to cut piece rates. By guaranteeing piece rates for workers regardless of their productivity, the wage lists therefore reduced the incentives to shirk (Lazonick 1986, p.25).

ever, the wage lists became dysfunctional to the British cotton industry in the twentieth century. By this time, the success of the US cotton industry based on ring frame and automatic loom technology necessitated changes in the technique of production employed in British cotton mills.[9] Such changes required the dismantling of the wage lists, and the relative pay scales and division of labour enshrined therein. Instead, the wage lists remained in place until after the Second World War, so that despite its being the basis for the growth and efficiency of the American cotton industry,

> '(s)cientific management' was to find little application in the British cotton industry, where even into the 1960s management was constrained by industrial

[9] The need for technological change and its relationship to efficiency in the cotton industry was previously discussed in chapter 7.

> relations systems that had become entrenched in the late nineteenth century.
>
> (Lazonick, 1986 p.28)

Why, then, did the wage lists remain in place for so long?

One reason for their longevity is due to interrelatedness *within* these institutions, in the form of the connections which arise between individual decision making units whose behaviour is governed by the same institution. Hence it would have been difficult for any individual spinning or weaving firm to deviate from an industrial relations practice common throughout the industry, because of the conflict with organised labour that this would have provoked. This is attested by the spate of industrial disputes in the interwar period provoked by individual mill owners attempts to re-organise working conditions and rewards. For example, an attempt by one large Burnley weaving firm in early 1931 to introduce an eight loom per weaver system (departing from an industry standard of less than six) resulted in a strike and subsequently an industry wide lock out from which mill workers emerged victorious (Lazonick 1986, p.29). Hence interrelatedness within the wage lists would have deterred attempts by individual firms to change these institutions at the margin.

This still leaves us to explain why firms in the cotton industry could not change the pattern of industrial relations by acting in concert. Recall from chapter 7 that the British cotton industry was characterised by a high degree of horizontal and vertical specialisation. Although employer associations did exist, attempts to assert their will in the competitive conditions that specialisation bred were usually ineffective.[10] Hence there existed a source of interrelatedness *between* the institutions of industrial relations and industrial organisation in the cotton industry; changing the former required accompanying changes in the latter, in order to increase cotton firms' degree of central co-ordination.

As early as the beginning of the twentieth century, the industrial relations system in the cotton industry required fundamental changes in order to remain functional to the development of the industry. However, the changes required (the replacement of the wage lists) were large relative to the size of the individual decision making unit (specialised spinning and weaving firms). Because of this, the industry effectively became locked in to a system of industrial relations inherited from the nineteenth century. This system of industrial relations persisted into the twentieth century as an inefficient institution, inhibiting technological change in the cotton industry and thereby contributing to the industry's decline.

---

[10] For example, even attempts at cartelization in order to control output prices were a failure. See Lazonick (1986, pp. 31-2).

b. Inefficient institutional evolution in the shipbuilding industry.

Prior to 1914, the British shipbuilding industry was largely comprised of small, independent shipyards. The size of these shipyards affected the technological and organizational development of the industry, and hence its system of industrial relations (Lorenz and Wilkinson, 1986).[11] The development of highly specialized trades in the shipbuilding industry resulted in the steady proliferation of trade unions, each one responsible for organizing a different trade. Even during the nineteenth century heyday of British shipbuilding, this complicated pattern of labour representation hindered technological change. Changing production methods tended to blur traditional craft distinctions, something to which the many craft unions in the industry were opposed. However, these effects were compensated by the benefits of unionization, which aided the organization of work and the transmission of skills, and lent cyclical flexibility to the processes of hiring and firing (Lorenz and Wilkinson, 1986).

After 1920, and especially after World War II, the decline of British shipbuilding, which had first been evident at the turn of the century, took the industry to the brink of

---

[11] Again , the lack of any trend towards consolidation in this key British industry seems to have played an important role in its subsequent institutional and technological development. For a discussion of the relative lack of merger activity in late nineteenth and early twentieth century Britain, the reader is referred back to section IIc of chapter 7.

collapse.[12] This decline took place during a period of sustained expansion in world shipbuilding. More importantly, it took place during a period of rapid technological change, as welding, burning and prefabrication techniques made traditional methods of riveting and plate working obsolete, and so redefined the division of labour in shipbuilding. The British industry's failure to assimilate technological changes that were widely adopted by its competitors was instrumental in its decline, and this failure may in part be explained by lock in to a by now inefficient industrial relations system.

As Lorenz and Wilkinson submit, trade unions were generally antagonistic to technological change and subsequent changes in craft demarcation lines. But this was hardly surprising in an industry already characterised by high labour turnover, and where fears of structural unemployment resulting from demarcation changes consequently ran high.

Rather, the problems of the shipbuilding industry were related to its pattern of industrial organization (small scale production) and its industrial relations system (characterised by the proliferation of unions and the subsequent inflexibility in the practice of specialized trades across occupational boundaries). Interrelatedness existed *within*

---

[12] In 1890, Britain accounted for 80% of the world market in shipbuilding. This share diminished to 60% by 1900, to 35% during the interwar years, and to 15% by the 1950s, after which the continued decline of the British shipbuilding industry threatened its very survival (Lorenz and Wilkinson 1986, p.109).

these institutions, in the form of the connections which necessarily arise between individual decision making units who comply with the same institution. It would have been difficult for any individual firm to deviate from industrial relations practices common to the shipbuilding industry as a whole (i.e., practices to which all other firms conformed) because of the conflict with organised labour that such deviance would have provoked. This is illustrated by the series of demarcation disputes that resulted from attempts in the 1930s by Clydeside shipbuilders to introduce and extend welding, burning and prefabrication techniques (McGoldrick, 1982). As in the case of the cotton industry, what this implies is that the basic interrelatedness existing between individuals whose behaviour is conditioned by a common institution would have deterred marginal (i.e., single firm) changes to the industrial relations system in the shipbuilding industry. This in turn points to a source of interrelatedness *between* the patterns of industrial organisation and industrial relations in shipbuilding, which is again similar to that found in the cotton industry. Hence for shipyard owners to have changed the industry wide industrial relations system would have required accompanying changes in the pattern of industrial organisation, from small independent firms to some degree of central coordination.

Both the industrial relations system and the pattern of industrial organisation therefore required non-marginal changes if new techniques of production based on the rationalisation of existing craft methods were to be successfully introduced in the shipbuilding industry. The key problem for British shipbuilding was that the institutional units which required change (the pattern of industrial organisation and the industrial relations system) were large relative to the size of the decision making unit (the individual firm). For example, the proliferation of competing unions in the industry as a whole was a problem beyond the sphere of control of any individual shipyard. Hence the difficulties associated with co-ordinating decentralized decision making units appear to have locked the shipbuilding industry in to an inefficient institutional structure, which thwarted technological change and contributed to the dramatic competitive decline of the industry during the postwar years.

c. Inefficient institutional evolution in the "new" industries: the case of motor vehicles.

The problem of inefficient institutional evolution in the British economy has not been confined to traditional industries. "New" industries, which began emerging only after Britain's period of industrial dominance, in the late nineteenth and early twentieth centuries, inherited no early/mid Victorian institutional legacy. Nevertheless, in cases such as the motor vehicle industry, the evolution of industrial relations still proved inimical to long run dynamic efficiency. What was common to the experience of both the new and traditional industries was that initial institutional selections proved critical in their subsequent development and decline.

The key technical feature of British motor vehicle firms in the late nineteenth/early twentieth century was their manufacture of virtually the entire vehicle. This was an important point of contrast with US producers, who were largely assemblers relying on outside supplies of finished components. It created a unique problem for British producers, who had to co-ordinate production units very much more complex than the US assembly shops, and bodies of labour which were much larger per unit of output.[13] These problems in turn led to the creation of a distinct pattern of industrial relations in the motor vehicle industry, characterized by "labour independence," i.e., weak control of management over the shopfloor and the supply of effort by labour (Lewchuk, 1986; Walker, 1981). This structure of industrial relations persisted well into the post war era, until after the industry was nationalised in the 1970s.

The lack of managerial control over the labour process resulted in the failure of the British motor vehicle industry during the interwar and especially the post war years to

[13] Some illustration of this is provided by the fact that in 1913, Ford produced over two hundred times as many cars as Wolseley, the largest British producer, Whilst employing only three times as much labour (Lewchuk 1986, p.139; Saul 1962, p.25)

embrace the capital intensive, Fordist technique of production pioneered in the US. The Fordist technique of production involved use of the integrated assembly line, and an accompanying division of labour and degree of managerial control over the labour process that was antithetical to the type of craft control that emerged in the British motor vehicle industry. Hence the pattern of industrial relations created by the early evolution of the industry, and which persisted thereafter, was inefficient; it obstructed the adoption of new techniques and was thus dysfunctional to the long run dynamic efficiency of the motor vehicle industry. This is illustrated by the fact that between 1950 and 1965, a period of expansion for the world motor vehicle industry, the rate of productivity growth in the British motor vehicle industry was slower than that achieved by all of its major competitors (France, West Germany, Italy, Japan) except the USA (Pratten and Silberston, 1967).

Throughout this period, the industry remained profitable only by virtue of the strong growth of world demand for motor vehicles, and because firms retained control over the setting of piece rates and were thus able to pay low wages (Lewchuk, 1986). However, by the late 1960s, declining demand conditions had provoked a crisis of profitability in the industry. Its failure to undergo technological change in the past coupled with the hostile reaction of labour towards management efforts to finally convert to Fordism - both of which were the product of the industry's inefficient industrial relations - now created the conditions for the virtual collapse of the British motor vehicle industry after 1970. Hence inefficient institutional evolution in the British motor vehicle industry created conditions which were inimical to the dynamic efficiency of this industry, and which contributed directly to its dramatic competitive decline.

III Twentieth Century British Industrial Relations: An Analytical Interpretation.

The outline in section I above suggests that for analytical purposes, the history of British industrial relations during the twentieth century can be divided into three broad epochs. The first, running from 1918 to c. 1935, was characterized by adversarial industrial relations, accompanied by the depressed economic and social conditions of the interwar years. The second, which encompasses both the late 1930s/World War II and the postwar years up to the mid 1960s, was characterized by more co-operative industrial relations. Many of the vestiges of Britain's historically decentralized and un-coordinated industrial relations system remained. However, in the context of twentieth century British industrial relations, the degree of co-operation between trade unions, employers and successive governments suggests that this period may be referred to as one of (albeit limited) capital-labour-state consensus. It was accompanied by the macroeconomic performance of the post war "golden age," which, as was illustrated in chapter 6, saw Britain achieve historically high but relatively low (compared with its major competitors) rates of output and productivity growth.

The third epoch, which runs from the late 1960s to the present, represents a return to an adversarial pattern of industrial relations, marked initially by increased labour militancy and subsequently by the hostile response of the state to trade unions. Furthermore, at least since 1973 the macroeconomic performance of the British economy has compared unfavourably to that achieved during the preceding "golden age."[14]

Interpreting these developments in terms of the model of institutional hysteresis of chapter 5, it appears that the epochs identified above and the economic conditions associated with them have been sequentially linked, with each epoch bequeathing the conditions for its successor. Hence negative social evaluations of the economic and social conditions during the interwar years, coupled with the legacy of enforced co-operation between labour, capital and the state during World War II resulted in the move towards increased co-operation associated with the post war capital-labour-state

---

[14] Again, the reader is referred to chapter 6 for comparative historical statistics on Britain's post war dynamic macroeconomic performance.

consensus.[15] This appears to have been broadly complementary to the dynamics of the post war British economy, accompanied as it was by the macroeconomic performance of the "golden age."[16] However, the capital-labour-state consensus only ever represented a partial change in the institutions of British industrial relations. Although the *attitudes* of trade unions, employers and especially governments appear to have been more favourably inclined towards the notion of a "social partnership," there was no trend towards *systemic* change in the pattern of industrial relations after World War II. As is clear from section I and the industry level studies in section II, traditional structures and practices derived from earlier periods continued to dominate the structure of British industrial relations. Accompanying this, it is important to remember that Britain's macroeconomic performance, whilst satisfactory by historical standards, was relatively poor by contemporary international standards.

Furthermore, the economic conditions of the "golden age" appear themselves to have been responsible for generating a

---

[15] The view that the emergence of post war welfarism in Britain derived from the conditions of the preceding era can be found in the Beveridge Report, which stresses the emergence of sentiments such as "national unity" during World War II. An alternative hypothesis, forwarded by Goodin and Dryzek (1987), is that the degree of uncertainty created by the war amongst all sections of the British population generated widespread demands for "socialised risk sharing," and so to the growth of welfarism.

[16] See Cornwall (1990) for a more detailed exploration of the links between institutions and the post 1945 macroeconomic performance of Britain and other capitalist economies.

re-evaluation of and subsequent challenge to the capital-labour-state consensus by materially secure, organized labour. Many of the "new demands" made by the labour movement in the late 1960s, such as those connected with control over the organisation of work and the quantity of employment, lay beyond the terms of the original post war industrial relations consensus, which was essentially based on recognition of the rights of trade unions to bargain over their members pay, and the provision of state welfarism. This re-evaluation of the post war "social partnership," coupled with the hostile reaction of capital and especially (after 1979) the state towards trade unions as a result of increased industrial strife, signified a further institutional change which gave rise to a new epoch based on adversarial industrial relations. These institutions appear to have been broadly dysfunctional to the dynamics of the British economy, judging by its inferior macroeconomic performance since 1973 compared to the preceding quarter of a century.

This interpretation of the evolution of British industrial relations can be summarised by the following (annotated) causal chain:

CONDITIONS OF INTERWAR YEARS
- legacy of social and economic strife
- legacy of co-operation prior to and during WWII

↓

Institutions "more" functional to macroeconomic performance

POST WAR INSTITUTIONAL CHANGE
- Emergence of capital -labour-state consensus
- But institutional change only partial - *systemic* change lacking, so that traditional structures and practices persist

↓

POST WAR ECONOMIC CONDITIONS
- "Golden Age" of macro performance
- Generates new demands from labour movement beyond terms of original capital-labour state consensus

↓

Institutions "less" functional to macroeconomic performance

FURTHER INSTITUTIONAL CHANGE
- Late 1960s labour unrest, hostile response of state leads to demise of capital-labour-state consensus

↓

'70s/'80s ECONOMIC CONDITIONS
- Poor macro performance relative to prior epoch

a. The procreation and persistence of inefficient institutions.

The causal chain above indicates that in terms of the model of institutional hysteresis developed in chapter 5, we can identify two key features of the evolution of British industrial relations. First, it suggests that since World War I, the institutions of British industrial relations have followed an evolutionary time path along which they have transformed from being less functional (1918-c.1935) to more functional (late 1930s - mid 1960s) to less functional (late 1960s - present) to the dynamics of British capitalism. In other words, the trajectory of the British economy since World War I has involved the *procreation* of different institutional structures in the sphere of industrial relations, which have been more or less complementary to its dynamics.[17] This interpretation of the British economy as having experienced different "regimes" of industrial relations during the twentieth century, with each regime growing out of a set of economic conditions associated with its predecessor, has received widespread support (see, for example, Barkin, 1975; Soskice, 1978; Crouch, 1978; Cornwall, 1990).[18]

Second, and more importantly, the causal chain indicates that even during the post war "golden age" of British capitalism, the British industrial relations system was highly imperfect. The capital-labour-state consensus involved only a partial institutional transformation. Many of the practices of

[17] The discussion of the motor vehicle industry in section II c above provides an industry level example of institutional procreation. In this particular case, the process of institutional procreation proved to be of an inefficient nature.

[18] Several of these authors offer an international perspective on the rise and demise of these "regimes" in industrial relations. They have begun to receive particular attention in the context of the US economy, with accompanying indications as to how changes in industrial relations regimes have affected the dynamic efficiency of this economy (see, for example, Weisskopf, Bowles and Gordon, 1983; Naples, 1981, 1986).

the preceding period of industrial relations adversarialism, which the consensus might ideally have overcome, were retained as central features of the post war industrial relations system. The limitations of the post war capital-labour-state consensus are also illustrated by the fragility of this consensus in the face of new, non-wage demands from the labour movement, which it was unable to accommodate. Finally, although the period of adversarialism since the late 1960s has chiefly been characterised by the hostile attitude of the state towards *trade unions* during the 1980s, at least some of this hostility has been directed towards "traditional practices" in the *industrial relations system*, such as the nature and legality of the Closed Shop.[19] Hence regardless of the causes of and underlying motives for the states' hostility towards labour, it is significant to note that even as late as the 1980s, traditional practices and forms of control over the shopfloor which, in many industries, had emerged more than a century earlier, were still an important issue in British industrial relations.

What all this suggests is that during the twentieth century, the British economy experienced the *persistence* of institutions, inherited as a legacy of the past, which were no longer functional to the dynamics of British capitalism. This

[19] The reader is again reminded of the significant difference between *organizations* such as trade unions, and *institutions* which govern the relationship between organizations, such as the industrial relations system.

phenomenon is clearly illustrated by the historical case studies in section II above. These case studies suggest that interrelatedness between and within institutions, and hence lock in to prior institutional structures, helps explain the persistence of nineteenth century industrial relations practices in Britain. Most importantly, they also illustrate the inhibiting effect of nineteenth century patterns of industrial relations on technological change, and hence the dynamic efficiency of British industry in the twentieth century.

What is clear from section II is that British industrial relations throughout the twentieth century have been synonymous with the persistence in Britain of the "craft method" of production, which relies on skilled craft labour on the shopfloor, and vests a relatively high degree of control over activity on the shopfloor in labour. During Britain's initial industrialisation, the advantage of this technique was that skilled craft labour instilled discipline in the newly emerging factories. For example, in late eighteenth century spinning firms, skilled spinners were employed as subcontractors, who out of their piece rate wages paid time rate wages to assistants whom they themselves recruited and supervised (Lazonick, 1979). This gave the spinner a vested interest in intensifying the supply of effort by his assistants and maintaining shopfloor discipline. The subcontracting hierarchy was retained even after spinning became more capital intensive in the 1830s and 1840s (Lazonick, 1979). Indeed, even in the late nineteenth century, craft co-ordination of activity on the shopfloor in Britain entailed lower fixed costs than those associated with the organisational changes (heralding the advent of managerial capitalism) underway in the US (Lazonick, 1990), whilst the apprenticeship system ensured the continual (intergenerational) retraining of labour on the shopfloor.

However, the main disadvantage with the craft method of production also lies in its system of industrial relations - and this disadvantage was to become increasingly evident in British industry during the twentieth century. The craft technique vests a low degree of control over the shopfloor in management, and consequently creates a poor environment for organisational and technological change in the context of a capitalist firm. Hence throughout the twentieth century, Britain has been forced to rely largely on intensifying the supply of effort by labour rather than undergoing technological change in order to remain competitive. Locked in to nineteenth century institutional structures, the British economy has been forced to grow *within* the craft technique of production. It has proved incapable of expanding *between* techniques, as evidenced by the failure of key British industries to convert to mass production during the middle of the twentieth century.[20]

IV Conclusion.

In this chapter, the evolution of British industrial relations in the twentieth century has been interpreted in terms of the model of institutional hysteresis developed in chapter 5. Evidence has been provided to suggest that during the twentieth century, Britain has experienced inefficient institutional evolution in the sphere of industrial relations, in particular as a consequence of the persistence of industrial relations practices inherited as a legacy of the nineteenth century. This has been shown to have had a detrimental effect on the dynamic efficiency of key industries in the British economy, by undermining the ability of these industries to undergo technological and organisational change.

It is important to bear in mind that the nineteenth century craft technique of production and the system of industrial relations associated with it were not brought about by trade union power. On the contrary, as illustrated earlier in this chapter, eighteenth and nineteenth century factory proprietors were only too willing to utilise craft workers to hire and train labour, and maintain discipline on the shopfloor. Following from this, it is important to emphasise

[20] For the view that Fordist mass production has since been superseded by the Japanese flexible manufacturing system, and that the latter now constitutes the technique of production within which the most dynamically successful economies are accumulating, see Lazonick (1990).

that it is not trade unions per se that have posed problems for the British economy, but rather the structure of the industrial relations system (i.e., the institutional setting) in which these organizations have found themselves operating.

The evidence presented in this chapter illustrates that, as in the case of technological evolution discussed in chapter 7, Britain has experienced inefficient institutional evolution since its nineteenth century industrial heyday. In terms of the model developed in Part II, the development over time of key labour market institutions in the British economy can be interpreted as having proved dysfunctional to Britain's long run growth dynamics. This suggests that inefficient institutional evolution has been instrumental in Britain's continued and accelerated relative economic decline during the twentieth century.

Conclusion.

I Understanding the evolution of capitalist economies.

The discussion in this thesis has grown out of a discontent with the traditional equilibrium approach to economic modelling, especially as applied to the dynamics of long run growth and development in capitalist economies. The methodology that has been adopted - based on the importance of self reinforcing mechanisms and in particular the concept of hysteresis - represents a radical departure from equilibrium analysis, especially in so far as it calls into question the usefulness of the steady state concept when studying the evolution of an economic system.

Yet if an approach based on hysteresis constitutes a *radical* departure from traditional equilibrium analysis, several features of the preceding analysis suggest that it is nevertheless a *sensible* departure, in the sense that modelling with hysteresis provides a superior "metaphor" for understanding capitalist dynamics than the concept of the steady state.

a. The importance of real historical time.

As noted in the introduction to Part I, traditional equilibrium economics - whether of a formally static or dynamic nature - treats time as a mathematical space. The approach adopted here, on the other hand, attempts to take account of the fact that economies exist in real historical time. Economic events occur in a uni-directional sequence, along which any event occurring in the "present" takes place in the context of a series of prior events corresponding to the periods which comprise the past.

The importance of conceiving economic dynamics in this way is that it rules out any approach to economic analysis which insists on interpreting the economy as a *tabular rasa* in each period. Rather, we are forced to take account of the elementary fact that in each period, the economy inherits the legacy of its own past. Once this is admitted, we are faced with the possibility that "the past matters," in the precisely defined sense that it may systematically influence subsequent configurations of the economy. If this is the case - and the analysis in this thesis based on hysteresis, cumulative causation and lock in is strongly suggestive of this possibility - then long run economic outcomes cannot be seen as determinate, depending only on exogenous "data" imposed upon the economic system from without. Instead, they must be seen as path dependent, which places emphasis on the nature of adjustment processes endogenous to the economy in explaining these outcomes. The notion that "the past matters" - that the *historical configurations* of an economic system can affect its evolution in a manner unforeseen by traditional equilibrium analysis - is essence of what this thesis seeks to assert.

b. The "unsmooth" process of economic change.

A second important feature of the preceding analysis is its emphasis of the idea that economic change is not a smooth and continuous process. On the contrary, it is suggested that there may be important difficulties associated with effecting a transition between, for example, different technological or institutional structures. In part, this is again due to the influence of the past on current economic activity. However, it is also caused by the existence of important indivisibilities in economic variables, which frequently render change a non-marginal process. In this thesis, the possibility of interrelatedness between technological and institutional components of the production process has been shown to be an important source of indivisibilities. As illustrated in chapters 4 and 5, interrelatedness can make incremental changes at the margin a practical impossibility, giving rise to the phenomenon of lock in. The elementary idea that not all of the activities or components which comprise the production process are continuously divisible, which implies that change is frequently a non-marginal and discontinuous process, can therefore have a major influence on the evolution of a decentralised economic system.

c. The potential inefficiency of economic outcomes.

What follows naturally from consideration of the constraints of history and the difficulties associated with economic change is the possibility that economic outcomes in any period may be inefficient - that they could be improved upon to the net benefit of society as a whole. Allowing for the possibility of inefficient outcomes seems intuitively appealing - especially when compared with the alternative of arguing that anything *but* an efficient outcome is impossible. Yet this is something of a departure from at least some of the more casual interpretations of traditional equilibrium analysis - in spite of the association of branches of this methodology with the problems of public goods and externalities (neoclassical welfare economics) and involuntary unemployment (conventional Keynesian macroeconomics). In the hands of its less careful practitioners, traditional equilibrium analysis can degenerate into a series of Panglossian assertions to the effect that the assumption of individual rationality *demands* that market outcomes be resolved in an efficient manner.

This assertion of the necessity of efficient market outcomes is highly unsatisfactory. As Elster (1989b, pp.147-8) argues, we can only be truly confident that a particular outcome has been brought about by virtue of its efficiency if a process can be demonstrated which specifies how the good consequences of an outcome act so as to both give rise to, and maintain it over time. In short, "(i)mputations of optimality require hard work, not just armchair speculation" (Elster 1989b, p.150).

An obvious corollary of Elster's argument is, of course, that imputations of inefficiency cannot be made casually, either. In view of this, taking as its point of departure the fundamentally historical nature of economic processes, one of the main aims of this thesis has been to carefully specify processes which may create and sustain inefficient economic outcomes.

d. Capitalism as an evolving system.

Ultimately, the analysis in this thesis can be seen as working towards a conception of capitalism as an endogenously evolving system which, whilst continually subject to genuinely external influences, is nevertheless comprised of mechanisms whose very operation can influence the long run configuration of the system as a whole. This conception is in keeping with the views expressed by Robinson (1974a, 1980) and Kaldor (1933, 1970, 1972, 1985) on the role of history in endogenously determining current economic outcomes, and following Kregel (1976), may even be seen as close to the views of Keynes himself. It suggests that there are no determinate "laws of history" which govern the fate of a capitalist economy independently of its prior trajectory. Instead, it is held that "(t)he only truly exogenous factor is *whatever exists at a given moment of time*, as a heritage of the past" (Kaldor 1985, p.61 - italics in original), a heritage from which current and future economic activities will arise as part of a historical sequence of path dependent events.

II Growth and development in the "real world."

To what extent does this conception of the historical process underlying capitalism contribute to our understanding of the growth and development experience of actual capitalist economies? Part III above indicates the potential applicability of a hysteretic model of growth and development to the relative rise and decline of the British economy. In particular, the concept of lock in to past technological and institutional structures - an important source of adjustment asymmetries in the model developed in Part II - has been found useful for explaining why Britain's initial phase of relative economic success bequeathed a subsequent phase of relative economic decline. Industry level studies suggest the persistence in Britain of nineteenth century techniques of production long after these had been rendered obsolete, and the contribution made by this inefficient technological and institutional evolution to declining relative efficiency and competitiveness.

a. Is the British case representative?

However, it is important to acknowledge that only so much can be learnt from studying the experience of one specific country. The evidence presented in Part III is therefore suggestive rather than conclusive. Whether the hysteretic model developed in Part II can explain the experience of other advanced industrial nations, and whether it is therefore more generally applicable to the evolution of capitalism than the evidence provided by the British experience is able to suggest, remains to be seen. To this end, it is perhaps reasonable to speculate that the trajectory of the US economy, which to some observers appears to have followed a developmental path not unlike that encountered by Britain before it,[1] is worthy of substantial future investigation. The tremendous contemporary significance of the decline of the US economy, as remarked in the introduction to this thesis, only provides further motivation for an investigation of this nature.

b. The neglected role of policy.

It is also important to note that one vitally important contemporary "exogenous" variable that has not been discussed in the course of this thesis is public policy. For some authors writing on the East Asian experience, policy can be seen as a vital component in ensuring that the technological and especially the institutional preconditions for cumulative high growth emerge and are maintained over time (see, for example, Amsden 1989, pp.63-4). Researching the influence of policy on self reinforcing growth dynamics would appear to be

[1] See, for example, Elbaum and Lazonick (1986).

imperative if the British experience is to be successfully reversed. It may also lend vital insights into what is necessary to prevent economies such as the US following similar trajectories. Following authors such as Amsden (1989), it may prove fruitful for research of this nature to examine more closely the example of the "East Asian model" of growth and development. Whatever, the challenge for policy that this thesis presents is clear. We have seen how cumulative high growth dynamics can endogenously break down, and how a period of relative economic success can consequently bequeath a subsequent period of relative economic failure. Is it possible, then, that policy can play a role in preventing such a breakdown, by successfully managing the fragile dynamic of cumulative high growth? Indeed, does the further possibility exist of utilising public policy to reverse a given historical breakdown in this dynamic, permitting the renewed rise of a "fallen," but once relatively successful economy?

XII Summary.

Taking as its starting point a methodology based on hysteresis, the principle aim of this thesis has been to construct a long run model of growth and development which takes account of the fact that economic activity occurs as a uni-directional sequence in real historical time. The model that has been developed is of a simple form, and can by no means claim to account for all the features of growth and development in a capitalist economy. Furthermore, the model raises as many important questions about growth as it provides potential answers to others (in particular, the vital conundrum surrounding the role of public policy in influencing the course of capitalist development). What the preceding analysis can claim, however, is to have provided a step away from the theoretical straitjacket of the steady state, a concept which attempts to characterise the history of capitalism - a history rich with change - as intrinsically unchanging. In so doing, the model that we have developed points towards the possibility of a rigorous analysis of economic dynamics "as if history mattered," which takes account of both the nature and significance of change in the evolution of economic systems.

Appendix A: Growth as transformation, and the existence of large scale increasing returns.

The purpose of this appendix is to provide a formal illustration of the way in which movement through a commodity hierarchy in the course of economic growth can provide a continual source of increasing returns. This is performed in the context of the model developed by Sundrum (1990).

Consider a simple economy in which the commodity hierarchy consists of the output of two industries or sectors, a and b.[1] Following Sundrum, we can write

$$\dot{q} = \frac{(q' - q)}{q} = \frac{(\lambda_a' q_a' - \lambda_a q_a + \lambda_b' q_b' - \lambda_b q_b)}{q}$$

where λ_i denotes the share of labour of the i^{th} sector, q_i is the average labour productivity of the i^{th} sector, q is the economy wide productivity growth rate, and q is the average productivity of the economy. A prime denotes the final value of a variable. This expression can be rewritten as:

$$\dot{q} = \frac{(\lambda_a' q_a' - \lambda_a q_a' + \lambda_a q_a' - \lambda_a q_a + \lambda_b' q_b' - \lambda_b q_b' + \lambda_b q_b' - \lambda_b q_b)}{q}$$

[1] The assumption that there are only two sectors can be made without loss of generality. See Sundrum (1990).

$$\Rightarrow \dot{q} = \frac{\lambda_a \cdot q_a}{q} \cdot \frac{(q_a' - q_a)}{q_a} + \frac{\lambda_b \cdot q_b}{q} \cdot \frac{(q_b' - q_b)}{q_b}$$

$$+ \frac{q_a'(\lambda_a' - \lambda_a) + q_b'(\lambda_b' - \lambda_b)}{q}$$

$$\Rightarrow \dot{q} = k_a \dot{q}_a + k_b \dot{q}_b + \frac{q_a'(\lambda_a' - \lambda_a) + q_b'(\lambda_T - \lambda_a' - \lambda_T + \lambda_a)}{q}$$

where $k_i = \lambda_i q_i / q$ represents the output share of the i^{th} sector and $\lambda_T = \lambda_a + \lambda_b$. Hence finally, we arrive at:

$$\dot{q} = k_a \dot{q}_a + k_b \dot{q}_b + \frac{(\lambda_a' - \lambda_a)(q_a' - q_b')}{q} \qquad [A.1]$$

Equation [A.1] suggests that the economy wide rate of growth of productivity depends not only on the sectoral productivity growth rates and the initial sectoral output shares, but also on the rate at which labour is shifting between sectors which have unequal productivity levels.

Defining the income elasticity of demand for the output of the i^{th} sector as:

$$e_i = \frac{(\lambda_i' q_i' - \lambda_i q_i)}{\lambda_i q_i q}$$

we can obtain an expression for the transfer of labour between sectors thus:

$$\lambda_i' - \lambda_i = \frac{\lambda_i(\dot{q}e_i - \dot{q}_i)}{(1 + \dot{q}_i)} \qquad [A.2]$$

and substituting [A.1] into [A.2] we arrive at:

$$\lambda_a' - \lambda_a = \frac{\lambda_a \lambda_b(\dot{q}_b e_a - \dot{q}_a e_b)}{\lambda_a e_a(1 + \dot{q}_b) + \lambda_b e_b(1 + \dot{q}_a)} \qquad [A.3]$$

Equation [A.3] suggests that labour will shift into a sector, such as a, if $\dot{q}_b e_a > \dot{q}_a e_b$, or alternatively (for non-zero rates of productivity growth) if $e_a/\dot{q}_a > e_b/\dot{q}_b$. Furthermore, from equation [A.1] it is apparent that in the limit, the economy wide productivity growth rate will converge to the productivity growth rate of the labour absorbing sector. Hence suppose that at the current level of per capita income we have $\dot{q}_a > \dot{q}_b$ due to the exhaustion of scale economies in sector b. But if $e_a > e_b$ because a is above b in the commodity hierarchy, it is possible that overall we will have

$e_a/\dot{q}_a > e_b/\dot{q}_b$. Suppose also that the average level of productivity is highest in sector a ($q_a > q_b$). Then equations [A.1] and [A.3] above suggest that the exhaustion of scale economies in sector b will not adversely affect the overall rate of growth of productivity in the economy. The transfer of resources to sector a (equation [A.3]) coupled with the exploitation of scale economies in this sector ($\dot{q}_a > \dot{q}_b$) will ensure that the economy wide rate of growth of productivity is converging to the new, higher level $\dot{q}_a$ over time (equation [A.1]). The sequential exploitation of increasing returns in different sectors of the economy in the course of movement through the commodity hierarchy will therefore cause the economy as a whole to act as if it is able to continuously realise economies of scale. The result of this is that the scope for increasing returns can, at any given point in time, be considered large relative to the size of the economy.

Appendix B: Labour transfer and productivity growth - an illustrative example.

Following Sundrum (1990), the dynamic analog of the productivity equation:

$$q_i = \lambda_{in}.q_{in} + \lambda_{io}.q_{io} \quad , \qquad i = A,B \qquad [B.1]$$

can be written as

$$\dot{q}_i = k_{in}.\dot{q}_{in} + k_{io}.\dot{q}_{io} + \frac{(\lambda'_{in} - \lambda_{in})(q'_{in} - q'_{io})}{q_i} \qquad [B.2]$$

where a dot above a variable represents its rate of growth, a prime denotes the final value of a variable, and $k_{in} = (\lambda_{in}.q_{in})/q_i$ and $k_{io} = (\lambda_{io}.q_{io})/q_i$ represent the output shares of the old and new techniques of production respectively. As indicated in Appendix A, equation [B.2] above expresses aggregate productivity growth in region i as a function of the weighted average rates of productivity growth corresponding to techniques n and o, and the rate at which labour is being transferred between these techniques, which have different average levels of productivity ($q_{in}' - q_{io}' \neq 0$).

In order to illustrate the influence of labour transfer between techniques (and hence the rate of adoption of the new technique) on aggregate productivity growth in regions A and

B, equation [B.2] can be simulated on the basis of the assumption that $\lambda_{An}' - \lambda_{An} > \lambda_{Bn}' - \lambda_{Bn}$. To isolate this resource transfer effect, we also assume that $q_{Bn} = q_{An} = q_n$, $q_{Bo} = q_{Ao} = q_o$, $\dot{q}_{Bn} = \dot{q}_{An} = \dot{q}_n$, and $\dot{q}_{Bo} = \dot{q}_{Ao} = \dot{q}_o$. The parameters $\dot{q}_n$ and $\dot{q}_o$ are set at 0.050 and 0.005 respectively, and the initial values (in period t = 0) of the variables λ_{in}, q_n and q_o are as follows:

$$\lambda_{in} = 0.100 \quad , \quad i = A,B$$
$$q_n = 200$$
$$q_o = 80$$

The evolution of the "endogenous" right hand side variables (i.e., those sensitive to λ_{in}, q_n and/or q_o in any period) is described as follows:

$$k_{in} = \lambda_{in}.q_n/q_i \quad , \quad i = A,B$$
$$k_{io} = \lambda_{io}.q_o/q_i \quad , \quad i = A,B$$
$$q_n' = (1 + 0.4)q_n$$
$$q_o' = (1 + 0.2)q_o$$
$$\lambda_{in}' = \lambda_{in} + [\phi_i.(1 - \lambda_{in})] \quad , \quad i = A,B$$
$$\Rightarrow \quad \lambda_{in}' - \lambda_{in} = \phi_i.(1 - \lambda_{in})$$

where $(1 - \lambda_{in}) = \lambda_{io}$, and ϕ_i, the rate at which labour is being transferred from o to n in region i, is 0.010 for region B and 0.015 for region A.

The results of this simulation are reported in Table 1 below:

Table 1: *Simulated values of $\dot{q}_B$ and $\dot{q}_A$ for ten periods.*

Period	$\dot{q}_B$	$\dot{q}_A$
t = 1	0.0282	0.0347
2	0.0298	0.0366
3	0.0314	0.0385
4	0.0329	0.0403
5	0.0345	0.0421
6	0.0360	0.0437
7	0.0374	0.0453
8	0.0388	0.0467
9	0.0402	0.0481
10	0.0415	0.0494

Table 1 illustrates that as a result of its faster transfer of resources towards the higher productivity technique of production, productivity growth is faster in region A than in region B throughout the ten periods of the simulation. Indeed, even if region A is a late starter in the adoption process, it may still succeed in overtaking region B in terms of productivity growth performance. This is illustrated in Table 2 below, where we assume that region A begins transferring labour towards the new technique in period

6.[2]

Table 2: *Simulated values of $\dot{q}_B$ and $\dot{q}_A$ when A is a "late starter" in adoption.*

Period	$\dot{q}_B$	$\dot{q}_A$
t = 1	0.0282	0.0151
2	0.0298	0.0155
3	0.0314	0.0158
4	0.0329	0.0162
5	0.0345	0.0166
6	0.0360	0.0347
7	0.0374	0.0366
8	0.0388	0.0385
9	0.0402	0.0403
10	0.0415	0.0421

Even in the "late starter" scenario captured in Table 2, region A is able to catch up with region B's rate of productivity growth by virtue of its faster labour transfer, and overtakes region B's rate of productivity growth in period 9.

Referring back to equation [B.1], note that when $\lambda_{in} = 1$ for i = A,B:

---

[2] We assume that in periods 1-5, $\lambda_{in} = 0.1$. Note that during this period, $\dot{q}_A$ evolves according to the expression:

$$\dot{q}_A = k_{An}.q_n + k_{Ao}.q_o$$

and since neither k_{An} nor k_{Ao} are constant, this implies some change in $\dot{q}_A$, as reflected in Table 2.

$$q_1 = q_n$$

$$\Rightarrow \quad \dot{q}_1 = \dot{q}_n$$

This implies that, in the limit, we will observe $\dot{q}_B = \dot{q}_A = \dot{q}_n$.[3] However, this result should not be given too literal an interpretation. First, recall that we are only considering one adoption process in isolation. If there are others which region A is subsequently faster to embark upon, its rate of productivity growth may remain permanently above that of region B simply as a result of its greater ability to transfer resources towards high productivity growth techniques of production. Second, and most importantly, we are assuming that $\dot{q}_{nB} = \dot{q}_{nA}$ in order to isolate the resource transfer effect on productivity growth. However, if we reintroduce the notion of dynamic returns to scale, differences in the rate of adoption between regions A and B may enhance and impair $\dot{q}_{nA}$ and $\dot{q}_{nB}$ respectively, and these effects will become self reinforcing in the process of cumulative causation. The implications of this for the competitive positions of regions A and B are analysed in chapter 4.

[3] Note from Figure 1 in chapter 4 that this limit value of $\dot{q}_B$ and $\dot{q}_A$ will be achieved in periods t_B and t_A respectively, the difference being accounted for by the dissimilar rates of labour transfer in regions A and B. Hence the result $\dot{q}_B = \dot{q}_A = \dot{q}_n$ holds for all periods $t \geq t_B$.

References.

Akerlof, G. (1970) "The market for 'lemons': quality uncertainty and the market mechanism," Quarterly Journal of Economics, 84, 488-500.

Akerlof, G. (1976) "The economics of caste and race and other woeful tales," Quarterly Journal of Economics, 90, 599-617.

Alchian, A. (1950) "Uncertainty, evolution and economic theory," Journal of Political Economy, 58, 211-21.

Aldcroft, D.H. (1964) "The entrepreneur and the British economy, 1870-1914," Economic History Review, 17, 113-34

Aldcroft, D.H. (ed.) (1968) The Development of British Industry and Foreign Competition, 1875-1914, London, Allen and Unwin.

Aldcroft, D.H. (1968) "Introduction: British industry and foreign competition, 1875-1914," in D.H. Aldcroft (1968) op. cit.

Amsden, A. (1989) Asia's Next Giant, Oxford, Oxford University Press.

Arrow, K. (1959) "Towards a theory of price adjustments," in M. Abramovitz et al, The Allocation of Economic Resources, Stanford, Stanford University Press.

Arrow, K. (1962) "The economic implications of learning by doing," Review of Economic Studies, 29, 155-73.

Arthur, W.B. (1988) "Self reinforcing mechanisms in economics," in P.W. Anderson, K. Arrow and D. Pines (eds.) The Economy as an Evolving Complex System, Reading, Mass., Addison - Wesley.

Arthur, W.B. (1988) "Competing technologies: an overview," in G. Dosi et al (eds.) Technical Change and Economic Theory, New York, Pinter.

Ayres, C. (1962) The Theory of Economic Progress, New York, Schocken Books.

Bardhan, P. (1989) "The new institutional economics and development theory: a brief critical assessment," World Development, 17, 1389-95.

Barkin, S. (1975) "The third post war decade (1965-75): progress, activism and tension," in S. Barkin (ed.) Worker Militancy and its Consequences, 1965-75, New York, Praeger.

Basu, K., E. Jones and E. Schlicht (1987) "The growth and decay of custom: the role of the new institutional economics in economic history," Explorations in Economic History, 24, 1-21.

Baumol, W.J., S.A.B. Blackman and E.N. Wolff (1989) Productivity and American Leadership: The Long View, Cambridge, Mass., The MIT Press.

Bowles, S., D.M. Gordon and T.E. Weisskopf (1990) After the Wasteland, Armonk, N.Y., M.E. Sharpe.

Boyer, R. (1987) "Regulation," in J. Eatwell, M. Milgate and P. Newman (eds.) The New Palgrave: A Dictionary of Economics, London, Macmillan.

Boyer, R. (1990) The Regulation School: A Critical Introduction, New York, Columbia University Press.

Bromley, D. (1989) Economic Interests and Institutions, Oxford, Basil Blackwell.

Brown, W. (1972) "A consideration of 'custom and practice,'" British Journal of Industrial Relations, 10.

Brunner, K. (1987) "The perception of man and the conception of society: two approaches to the understanding of society," Economic Inquiry, 25, 367-88.

Cairncross, A. "The post war years, 1945-77," in R.C. Floud and D.N. McCloskey (eds.) op. cit.

Chick, V. (1983) Macroeconomics After Keynes, Cambridge, The MIT Press.

Church, R.A. (1975) The Great Victorian Boom, 1850-1873, London, Macmillan.

Clark, G. (1987) "Why isn't the whole world developed? Lessons from the cotton mills," Journal of Economic History, 47, 141-73.

Clower, R. (1965) "The Keynesian counterrevolution: an appraisal," in F.H. Hahn and F.P.R. Brechling (eds.) The Theory of Interest Rates, London, Macmillan.

Coates, K. and T. Topham (1986) Trade Unions and Politics, Oxford, Basil Blackwell.

Coddington, A. (1972) "Positive economics," Canadian Journal of Economics, 5, 1-15.

Coddington, A. (1975) "The rationale of general equilibrium theory," Economic Inquiry, 13, 539-58.

Commons, J.R. (1961) Institutional Economics, Madison, University of Wisconsin Press.

Cornwall, J. (1972) Growth and Stability in a Mature Economy, London, Martin Robertson.

Cornwall, J. (1977) Modern Capitalism, Oxford, Martin Robertson.

Cornwall, J. (1990) The Theory of Economic Breakdown, Oxford, Basil Blackwell.

Cornwall, J. (1991) "Prospects for unemployment in the 1990s with hysteresis," in J. Cornwall (ed.) The Capitalist Economies: Prospects for the 1990s, Aldershot, Edward Elgar.

Cornwall, J. and W. Cornwall (1987) "The political economy of stagnation," Journal of Economic Issues, 25, 535-50.

Cornwall, J. and W. Cornwall (1992) "Export led growth: a new interpretation," in W. Milberg (ed.) The Megacorp and Macrodynamics, Armonk, New York, M.E. Sharpe.

Cross, R. (1987) "Hysteresis and instability in the natural rate of unemployment," Scandinavian Journal of Economics, 89, 71-89.

Cross, R. (1990) "On the foundations of hysteresis in economic systems," University of Strathclyde Working Paper.

Cross, R. (1991) "The NAIRU: Not an interesting rate of unemployment?" University of Strathclyde Working Paper.

Cross, R. and A. Allen (1988) "On the history of hysteresis," in R Cross (ed.) Unemployment, Hysteresis and The Natural Rate Hypothesis, Oxford, Basil Blackwell.

Crouch, C. (1978) "The intensification of industrial conflict in the United Kingdom," in C. Crouch and A. Pizzorno (eds.) The Resurgence of Class Conflict in Western Europe Since 1968 Volume 1, New York, Holmes and Meier.

Dixit, A. (1989) "Hysteresis, import penetration and exchange rate pass through," Quarterly Journal of Economics, 104, 205-28.

Doeringer, P. and M.J. Piore (1971) Internal labour markets and Manpower Analysis, Lexington, Mass., Heath.

Dopfer, K. (1991) "Toward a theory of economic institutions: synergy and path dependence," Journal of Economic Issues, 25, 535-50.

Dornbusch, R. (1987) "Exchange rate economics: 1986," Economic Journal, 97, 1-18.

Dowie, J.A. (1968) "Growth in the interwar period: some more arithmetic," Economic History Review, 21, 93-112.

Dunlop, J. (1958) Industrial Relations Systems, New York, Holt.

Eatwell, J. (1984) Whatever Happened To Britain? New York, Oxford University Press.

Elbaum, B. (1990) "Cumulative or comparative advantage? British competitiveness in the early twentieth century," World Development, 18, 1255-72.

Elbaum, B. and W. Lazonick (eds.) (1986) The Decline of the British Economy, Oxford, The Clarendon Press.

Elbaum, B. and W. Lazonick (1986) "An institutional perspective on British decline," in B. Elbaum and W. Lazonick (eds.) op. cit.

Elster, J. (1976) "A note on hysteresis in the social sciences," Synthese, 33, 371-91.

Elster, J. (1985) Making Sense of Marx, Cambridge, Cambridge University Press.

Elster, J. (1989a) "Social norms and economic theory," Journal of Economic Perspectives, 3, 99-117.

Elster, J. (1989b) The Cement of Society, Cambridge, Cambridge University Press.

Farnie, D.A. (1979) The English Cotton Industry and the World Market 1815-1896, Oxford, Oxford University Press.

Feinstein, C.H. (1972) National Income, Expenditure and Output of the United Kingdom 1855-1965, Cambridge, Cambridge University Press.

Feinstein, C.H., R.C.O. Matthews and J.C. Odling Smee (1982) "The timing of the climacteric and its sectoral incidence in the UK, 1873-1913," in C.P. Kindleberger and G. di Tella (eds.) Economics in the Long View Volume 2: Applications and Cases Part I. New York, New York University Press.

Field, A.J. (1984) "Microeconomics, norms and rationality," Economic Development and Cultural Change, 32, 683-711.

Floud, R.C. (1981) "Britain 1860-1914: a survey," in R.C. Floud and D.N. McCloskey (eds.) op. cit.

Floud, R.C. and McCloskey, D.N. (eds.) The Economic History of Britain Since 1700 Volume 2: 1860 to the 1970s, Cambridge, Cambridge University Press.

Frankel, M. (1955) "Obsolescence and technological change in a maturing economy," American Economic Review, 45, 296-319.

Franz, W. (1990) "Hysteresis in economic relationships: an overview," Empirical Economics, 15, 109-25.

Friedman, M. (1953) "The methodology of positive economics," in Essays in Positive Economics, Chicago, Chicago University Press.

Friedman, M. (1968) "The role of monetary policy," American Economic Review, 43, 1-17.

Georgescu-Roegen, N. (1971) The Entropy Law and the Economic Process, Cambridge, Mass., Harvard University Press.

Giavazzi, F. and C. Wyplosz (1985) "The zero root problem: a note on the dynamic determination of the stationary equilibrium in linear models," Review of Economic Studies, 52, 353-57.

Goodin, R.E. and J. Dryzek (1987) "Risk sharing and social justice: the motivational foundations of the post war welfare state," in R.E. Goodin and J. LeGrand, Not Only The Poor, London, Allen and Unwin.

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

Gordon, R.J. (1989) "Hysteresis in history: was there ever a Phillips curve?", American Economic Review, 79, 220-5.

Gordon, W. (1980) Institutional Economics, Austin, Texas, University of Texas Press.

Grossman, G.M. and E. Helpman (1991) Innovation and Growth in the Global Economy, Cambridge, Mass., The MIT Press.

Hahn, F. (1970) "Some adjustment problems," Econometrica, 38, 1-17. Reprinted in Equilibrium and Macroeconomics, Oxford, Basil Blackwell, 1984.

Hahn, F. (1974) "On the notion of equilibrium in economics," inaugural lecture, Cambridge University, Cambridge University Press. Reprinted in Equilibrium and Macroeconomics, Oxford, Basil Blackwell, 1984.

Hahn, F. (1982) "Reflections on the invisible hand," Lloyds Bank Review, 1-21. Reprinted in Equilibrium and Macroeconomics, Oxford, Basil Blackwell, 1984.

Handbook Of International Economic Statistics, Faculty of Economics, Cambridge University, 1986.

Hargreaves Heap, S.P. (1980) "Choosing the wrong natural rate: Accelerating inflation or decelerating employment and growth," Economic Journal, 90, 611-20.

Harley, C.K. and D.N. McCloskey (1981) "Foreign trade: competition and the expanding international economy," in R.C. Floud and D.N. McCloskey (eds.) op. cit.

Hayek, F.A. (1945) "The use of knowledge in society," American Economic Review, 35, 519-30.

Henin, P.Y. (1986) Macrodynamics: Fluctuations and Growth, London, Routledge and Kegan Paul.

Hibbs, D. (1978) "On the political economy of long run trends in strike activity," British Journal of Political Economy, 8, 153-76.

Hirsch, F. (1976) Social limits to Growth, Cambridge, Mass., Harvard University Press.

Hodgson, G.M. (1986) "Behind methodological individualism," Cambridge journal of Economics, 10, 211-24.

Hodgson, G.M. (1991) "Economic evolution: intervention contra Pangloss," Journal of Economic Issues, 25, 519 - 33.

Hont, I. (1983) "The 'rich country-poor country' debate in Scottish classical political economy," in I. Hont and M. Ignatieff (eds.) Wealth and Virtue: The Shaping of Political Economy in the Scottish Enlightenment, Cambridge, Cambridge University Press.

Jenkinson, T. (1987) "The natural rate of unemployment?" Oxford review of Economic Policy, 3, 20-6.

Kaldor, N. (1933) "A classificatory note on the determinateness of equilibrium," Review of Economic Studies, 1, 122-36.

Kaldor, N. (1970) "The case for regional policies," Scottish Journal of Political Economy, 18, 337-48. Reprinted in Further Essays on Economic Theory, New York, Holmes and Meier, 1978.

Kaldor, N. (1972) "The irrelevance of equilibrium economics," Economic Journal, 82, 1237-55.

Kaldor, N. (1981) "The role of increasing returns, technical progress and cumulative causation in the theory of international trade and economic growth," Economie Appliquee, 34, 593-617. Reprinted in Further Essays on Economic Theory and Policy, New York, Holmes and Meier, 1989.

Kaldor, N. (1985) Economics Without Equilibrium, Cardiff, University College Cardiff Press.

Keynes, J.M. (1936) The General Theory of Employment Interest and Money, New York, Harcourt Brace and Company.

Kitson, M. and S. Solomou (1989) "The macroeconomics of protectionism: the case of Britain in the 1930s," Cambridge Journal of Economics, 13, 155-69.

Krasnosel'skii, M.A. and A.V. Pokrovskii (1989) Systems With Hysteresis, Berlin, Springer-Verlag.

Kregel, J.A. (1976) "Economic methodology in the face of uncertainty: the modelling methods of Keynes and the Post-Keynesians," Economic Journal, 86, 209-25.

Krugman, P.R. (1989) Exchange Rate Instability, Cambridge, The MIT Press.

Lamfalussy, A. (1961) Investment and Growth in Mature Economies, New York, Macmillan.

Lamfalussy, A. (1963) The UK and the Six, Homewood, Irwin.

Langlois, R.N. (1986) "The new institutional economics: an introductory essay," in R.N. Langlois (ed.) Economics as a Process, Cambridge, Cambridge University Press.

Layton, D. "Low pay and collective bargaining," in F. Field (ed.) Low Pay, London, Acton Society Trust.

Lazonick, W. (1979) "Industrial relations and technical change: the case of the self acting mule," Cambridge Journal of Economics, 3, 231-62.

Lazonick, W. (1981) "Factor costs and the diffusion of ring spinning in Britain prior to World War I," Quarterly Journal of Economics, 96, 89-109.

Lazonick, W. (1983) Industrial organization and technological change: the decline of the British cotton industry," Business History Review, 57, 195-236.

Lazonick, W. (1986) "The cotton industry, " in B. Elbaum and W. Lazonick (eds.) op. cit.

Lazonick, W. (1990) Competitive Advantage on the Shopfloor, Cambridge, Mass., Harvard University Press.

Lee, C.H. (1986) The British Economy Since 1700: A Macroeconomic Perspective, Cambridge, Cambridge University Press.

Lewchuk, W. (1986) "The motor vehicle industry," in B. Elbaum and W. Lazonick (eds.) op. cit.

Lucas. R.E. (1975) "An equilibrium model of the business cycle," Journal of Political Economy, 83, 1113-44.

Lucas, R.E. (1976) "Econometric policy evaluation: a critique," in K. Brunner and A.H. Meltzer (eds.) The Phillips Curve and the Labour Market, Amsterdam, North Holland.

Maddison, A. (1964) Economic Growth in the West, New York, Twentieth Century Fund.

Maddison, A. (1982) Phases of Capitalist Development, Oxford, Oxford University Press.

Matthews, R.C.O., C.H. Feinstein and J.C. Odling-Smee (1982) British Economic Growth 1856-1973, Stanford, Stanford University Press.

Mayergoyz, I.D. (1985) "Hysteresis models from the mathematical and control theory points of view," Journal of Applied Physics, 57, 3803-3805.

Mayergoyz, I.D. (1991) Mathematical Models of Hysteresis, Berlin, Springer-Verlag.

McCloskey, D.N. (1983) "The rhetoric of economics," Journal of Economic Literature, 21, 481-517.

McGoldrick, J. (1982) "Crisis and the division of labour: Clydeside shipbuilding in the Interwar period," in A Dickson (ed.) Capital and Class in Scotland, Edinburgh, John Donald and Co..

Miller, E.S. (1978) "Institutional Economics: philosophy, methodology and theory," Social Science Journal, 15, 13-25.

Naples, M.I. (1981) "Industrial conflict and its implications for productivity growth," American Economic Review, 71, 36-41.

Naples, M.I. (1986) "The unravelling of the union-capital truce and the US industrial productivity crisis," Review of Radical Political Economics, 18, 110-131.

OECD Historical Statistics, Paris, OECD, 1990.

Payne, P.L. (1974) British Entrepreneurship in the Nineteenth Century, London, Macmillan

Pelling, H. (1972) A History of British Trade Unionism, London, Macmillan.

Phelps, E.S. (1972) Inflation Policy and Unemployment Theory, London, Macmillan.

Pollard, S. (1962) The Development of the British Economy 1914-50, London, Edward Arnold.

Pollard, S. (1989) Britain's Prime and Britain's Decline, London, Edward Arnold.

Pratten, C. and A. Silberston (1967) "International comparisons of productivity in the automobile industry, 1950-1965," Oxford Bulletin of Economics and Statistics, 29, 373-94.

Ramaswamy, R. and R.E. Rowthorn (1991) "Efficiency wages and wage dispersion," Economica, 58, 501-14.

Richardson, G.B. (1959) "Equilibrium, expectations and information," Economic journal, 49, 223-37.

Richardson, H.W. (1962) "The basis of economic recovery in the 1930s: a review and a new interpretation, Economic History Review, 15, 344-63.

Richardson, H.W. (1965) "Over-commitment in Britain before 1930," Oxford Economic Papers, 17, 237-62.

Richardson, H.W. (1967) Economic Recovery in Britain 1932-9, London, Weidenfeld and Nicolson.

Richardson, H.W. (1968) "Chemicals," in D.H. Aldcroft (ed.) op. cit.

Ricoy, C.J. (1987) "Cumulative causation," in J. Eatwell, M. Milgate and P. Newman (eds.) The New Palgrave: A Dictionary of Economics, London, Macmillan.

Robinson, J. (1974a) "History versus equilibrium," Thames Papers in Political Economy, London, Thames Polytechnic. Reprinted in Collected Economic Papers Vol. V, Cambridge, Mass., The MIT Press, 1980.

Robinson, J. (1974b) "The abdication of neoclassical economics," in A. Mitra (ed.) Economic Theory and Planning: Essays in Honour of A.K. DasGupta, Calcutta, Oxford University Press. Reprinted in Collected Economic Papers Vol. V, Cambridge, Mass., The MIT Press, 1980.

Robinson, J. (1980) "A lecture delivered at Oxford by a Cambridge economist," in Collected Economic Papers Vol. IV, Cambridge, Mass., The MIT Press.

Romer, P.M. (1990) "Endogenous Technical Change," Journal of Political Economy, 98, S71-S102.

Rowthorn, R.E. and J.R. Wells (1987) De-industrialization and Foreign Trade, Cambridge, Cambridge University Press.

Rutherford, M. (1989) "What is wrong with the new institutional economics (and what is still wrong with the old)?" Review of Political Economy, 1, 299-318.

Samuelson, P.A. (1965) "Some notions on causality and teleology in economics," in D. Lerner (ed.) Cause and Effect, New York, The Free Press. Reprinted in R.C. Merton (ed.) The Collected Scientific Papers of Paul Samuelson Vol. 3, Cambridge, Mass., The MIT Press, 1972.

Sandberg, L. (1974) Lancashire in Decline, Columbus, Ohio State University Press.

Sandberg, L. (1981) "The entrepreneur and technological change," In R.C. Floud and D.N. McCloskey (eds.) op. cit.

Saul, S.B. (1962) "The motor industry in Britain to 1914," Business History, 5, 22-44.

Saul, S.B. (1968) "The engineering industry," in D.H. Aldcroft (ed.) op. cit.

Saul, S.B. (1969) The Myth of the Great Depression, London, Macmillan.

Schotter, A. (1981) The Economic Theory of Social Institutions, Cambridge, Cambridge University Press.

Schumpeter, J. (1942) Capitalism, Socialism and Democracy, New York, Harper.

Scitovsky, T. (1956) "Economies of scale and European integration," American Economic Review, 46, 71-91.

Setterfield, M.A. (1992a) "Towards a long run theory of effective demand: modelling macroeconomic systems with hysteresis," Journal of Post Keynesian Economics, forthcoming.

Setterfield, M.A. (1992b) "A model of institutional hysteresis," Journal of Economic Issues, forthcoming.

Shackle, G.L.S. (1958) Time in Economics, Amsterdam, North Holland.

Shalev, M. "Lies, damned lies and strike statistics," in C, Crouch and A. Pizzorno (eds.) The Resurgence of Class Conflict in Western Europe Since 1968 Vol. 1, New York, Holmes and Meier.

Shubik, M. (1981) "Game theory models and methods in political economy," in K.J. Arrow and M.D. Intriligator (eds.) Handbook of Mathematical Economics, Amsterdam, North Holland Pub. Co..

Skott, P. (1983) "An essay on Keynes and general equilibrium theory," Thames Papers in Political Economy, London, Thames Polytechnic.

Skott, P. (1985) "Vicious circles and cumulative causation," Thames Papers in Political Economy, London, Thames Polytechnic.

Soskice, D. (1978) "Strike waves and wage explosions 1969-70: an economic interpretation," in C. Crouch and A. Pizzorno (eds.) The Resurgence of Class Conflict in Western Europe Since 1968 Vol. 2, New York, Holmes and Meier.

Soskice, D. and W. Carlin (1989) "Medium run Keynesianism: hysteresis and capital scrapping," in P. Davidson and J. Kregel (eds.) Macroeconomic Problems and Policies of Income Distribution, Aldershot, Edward Elgar.

Stiglitz, J. (1985) "Equilibrium wage distributions," Economic Journal, 95, 595-618.

Stiglitz, J.E. (1987) "The causes and consequences of the dependence of quality on price," Journal of Economic Literature, 25, 1-48.

Sundrum, R.M. (1990) Economic Growth in Theory and Practice, London, Macmillan.

Svennilson, I. (1983) Growth and Stagnation in the European Economy, New York, Garland.

Targetti, F. and B. Kinda Hass (1982) "Kalecki's review of Keynes' General Theory," Australian Economic Papers, 21, 244-60.

Taylor, A.J. (1961) "Labour productivity and technological innovation in the British coal industry, 1850-1914," Economic History Review, 14, 48-70.

Taylor, A.J. (1968) "The coal industry," in D.H. Aldcroft (ed.) op. cit.

Thirlwall, A.P. (1980) Balance of Payments Theory and the UK Experience, London, Macmillan.

Thirlwall, A.P. (1983) "A plain man's guide to Kaldor's growth laws," Journal of Post Keynesian Economics, 5, 345-58.

Tyson, R.E. (1968) "The cotton industry," in D.H Aldcroft (ed.) op. cit.

Tyszynski, H. (1951) "World trade in manufactured commodities, 1899-1950," The Manchester School of Economic and Social Studies, 19, 272-304.

Varian, H.R. (1984) Microeconomic Analysis, New York, W.W. Norton and Co.

Veblen, T. (1975) The Theory of Business Enterprise, Clifton: Augustus M. Kelley.

Walker, J. (1981) "Markets, industrial processes and class struggle: the evolution of the labour process in the UK engineering industry," Review of Radical Political Economy, 12, 46-69.

Weisskopf, T.E., S. Bowles and D.M. Gordon (1983) "Hearts and minds: a social model of US productivity growth," Brookings Papers on Economic Activity, 2, 381-441.

Weitzman, M.L. (1989) "A theory of wage dispersion and job market segmentation," Quarterly Journal of Economics, 54, 121-37.

Wilber, C.K. and R.S. Harrison (1978) "The methodological basis of institutional economics: pattern model, storytelling, and holism," Journal of Economic Issues, 12, 61-89.

Williamson, O.E. (1985) The Economic Institutions of Capitalism, New York, The Free Press.

Witt, U. (1990) "Reflections on the present state of evolutionary economic theory," paper presented at the meetings of the European Association for Evolutionary Political Economy, Florence, Italy, Nov. 1990

Wyplosz, C. (1987) "Comments," in R. Layard and L. Calmfors (eds.) The Fight Against Unemployment, Cambridge, Mass., The MIT Press.

Young, A. (1928) "Increasing returns and economic progress," Economic Journal, 38, 527-42.